Axioms for Social Welfare Orderings*

Claude d'Aspremont†

1 Introduction

The theoretical literature on social organizations has always been concerned with the determination of effective institutions or common decision criteria that integrate, in some way or another, the different participating units. In fact this "aggregation problem" has become the fundamental subject of the formal analysis of political and economic organizations. This is clear in the considerable development of the theory of social choice, as initiated by K. Arrow [1951,1963]. But this problem is also at the foundation of the theory of games, which has grown extensively since the book of J. von Neumann and O. Morgenstern [1947]. Furthermore it motivates essentially the revival of interest in a "formal" approach to ethics that has been recently stimulated by Rawls's [1972] criticism of utilitarianism and by Harsanyi's [1965, 1977] defense of this doctrine.

The intention of this essay is to give an introduction to that part of social choice theory that has strong ethical implications. This part is particularly related to the area of economics that is called, after Pigou [1920], welfare economics. It is less directly related to the area of political science, which analyzes the various methods of election.

As a deductive system, social choice theory has made clear the difficult issues that the various interpretations of its basic terms introduce. It has made clear that for each particular

*Reprinted from Leonid Hurwicz, David Schmeidler and Hugo Sonnenschein (eds.), *Social Goals and Social Organization. Essays in Memory of Elisha Pazner.* Cambridge, Cambridge University Press, 19-75, 1985.

I am much obliged to Louis Gevers, Peter Hammond and John Weymark for helpful comments and suggestions.

†CORE, Université catholique de Louvain

"model" a specific set of axioms should be constructed. My objective is thus to discuss the specific conditions that determine the welfare interpretation of social choice theory. The results surveyed in the following are only representative of this welfare model.[1]

In welfare terms, the aggregation problem can be formulated as the problem of deriving, for some collectivity of individuals and some set of social alternatives, a social judgment about those alternatives, which is based on their evaluation by each individual. In other words, the social evaluation of the different alternatives should be determined by their "utility" for each individual. The basic terms involved are the individuals themselves, their utilities, and the social alternatives together with their social evaluation. Even in welfare economics these terms have been understood quite differently and the relevance of the results presented thereafter depends crucially on these differences.

1.1 Utility and social alternatives

The term *utility* appeared in the eighteenth-century moral philosophy, characterized by its concern for human "happiness" and fundamentally inspired by the success of the natural sciences. It is as an integrated part of this general effort to develop a social science oriented toward human happiness that the "classical utilitarianism" doctrine was built up. This doctrine probably starts with Hutcheson [1725] but is best known through Bentham [1789]. As well emphasized by Little ([1957], p. 7), the notion of utility was then seen as an intrinsic property of objects, that of generating satisfaction. Thus the happiness of an individual was simply the sum of his satisfactions and the happiness of society the sum total of the happiness of all the individuals in society. In the utilitarian approach the set of social alternatives is usually taken to be very comprehensive. At a first level, where we shall speak of "social states", it may include the level of all sorts of goods, not only economic goods and services, but, to take a Rawlsian terminology, all "social primary goods" including "rights and liberties, powers and opportunities, income and wealth".

[1]Some more general results are given in K. Roberts thesis (or K. Roberts [1980a-c]). See also Sen [1977] and Blackorby, Donaldson, and Weymark [1984] and Moulin [1982]. A general survey of social choice theory can be found in Sen [1979b].

However, adopting a consequentialist viewpoint, the set of alternatives may also be defined at a higher level and comprise all individual actions resulting in some more or less good social state (see Sen [1979a]). At even higher levels one may introduce the institutions or rules governing these actions or, in a more subtle way, the patterns of motivations or personal dispositions that are most useful to obtain some social states. These different interpretations of the set of social alternatives give rise successively to different kinds of utilitarianism: outcome utilitarianism, act utilitarianism, rule utilitarianism, motive utilitarianism, and so on. The reason for this comprehensiveness is to maintain the project of a unified scientific approach to the understanding of mankind, both for descriptive and prescriptive purposes, and, hence, to take the principle of "the greatest amount of happiness on the whole" (Sidgwick [1907]) as the objective criterion for morality and to base on such a maximum principle the edification of a *mécanique sociale*, which "may one day take her place along with *mécanique céleste*" (Edgeworth [1881]).[2]

Contrary to this unifying utilitarian project, another tendency has been to try to separate economics and other social sciences from ethics in some way or another. This other tendency was already present at the rise of the utilitarian doctrine, but it received its main impetus from Pigou [1920], for whom the economics of welfare were to be distinguished from ethics since it was a scientific discipline dealing with measurable quantities. These measurable quantities were the individual utilities or "satisfactions", namely that part of total welfare which "can be brought directly or indirectly into relation with the measuring rod of money" (Pigou [1920]).[3] More than a reinterpretation of the notion of utility, this represents a limitation imposed on the set of social alternatives: welfare economics should only be concerned with social states differing in the amount of goods and services relevant to describe the production and exchange process. A reinterpretation of the notion of utility was definitely involved though in the development of what has been called later the new welfare economics (Stigler [1943], Samuelson [1947]) and which centered on the collective optimality concept introduced by Pareto [1909]. This has led to the conclusion that utility should only be an ordinal representation of an individual preference

[2]See the 1967 Reprint p. 1, quoted in Little [1957], p. 8.

[3]See the 1962 Reprint p. 11.

defined on some set of possibilities. These preferences should in turn be determined by reference to some (hypothetical) choice situation; an individual prefers some possibility to some other if he would choose it rather than the other. In other words, an individual utility is based on his (hypothetical) behavior, which is supposed to respect some general consistency conditions. No underlying notion of satisfaction or happiness is now required. More generally, the motivations underlying the choices (e.g., to get some pleasure, to do his duty, etc.) are not relevant as long as these choices are consistent.[4] A characteristic of this new approach was the avoidance, or for some authors like Robbins [1932] the rejection, of any sort of interpersonal comparisons of utilities. The Pareto optimality conditions – as also developed by Lerner [1934], Barone [1935], Hotelling [1938] and others – were realized to be valid even if interpersonal comparisons were not possible. Moreover, later on the great multiplicity of Pareto-optimal points was emphasized and "compensation tests" were devised[5] to extend Pareto's definition of an increase in social welfare without making interpersonal comparisons and thus avoid ethical considerations. It is the merit of Bergson's [1938] and Samuelson's [1947] concept of social welfare function to have made clear

> that it is not literally true that the new welfare economics is devoid of *any* ethical assumptions. Admittedly, however, its assumptions are more general and less controversial, and it is for this reason that it gives incomplete necessary conditions, whose full significance emerges only after one has made interpersonal assumptions. To refuse to take the last step renders the first two steps nugatory; like pouring out a glass of water and then refusing to drink ... (Samuelson [1947], p. 249).

We shall turn subsequently to this concept of social welfare function. In fact, one may argue that Arrow's [1951, 1963] theorem on the impossibility of a social welfare function may be viewed, in

[4]However, if the ordinality of the utility representation is combined with some limitation imposed on the set of alternatives (quantities of goods, for example) then notions like satisfactions, tastes, and needs can still be used, but more as a way to describe "expected" experiences then ex-post feelings.

[5]These were given by Kaldor [1939], Hicks [1939], Scitovsky [1941-42]. See also Baumol [1946], Little [1957], and Samuelson [1947]. For a survey see Graaff [1957].

some sense, as the explicit formalization of Samuelson's statement, leading to the same negative conclusion concerning the compensation principle.

1.2 Interpersonal comparisons and the identity of individuals

It is not because the introduction of interpersonal comparisons reinforces the ethical character of the presuppositions involved in any kind of welfare judgment, that these comparisons should be denied any empirical significance. As stressed by Little [1957], interpersonal comparisons do rest on observation or introspection. In fact many economists have attempted to measure marginal utility. The early methods were criticized by Vickrey [1945]:

> Most attempts to determine marginal utility hitherto have been based on the assumptions first that some utility function can be found that will thus make the marginal utility of some commodity independent of the quantities of all other commodities, and second that this function when found has some special validity for the purposes at hand. The methods suggested by Irving Fisher [1927], and the isoquant, quantity-variation, and translation methods of Ragnar Frisch [1932] all involve such assumptions.

Hence these methods do not seem to lead to a plausible way of making interpersonal comparisons. Arrow [1951, 1963] criticizes three different analytic approaches to derive empirically (and ethically) meaningful interpersonal comparisons. The first is due to Dahl [1956] and bases an interpersonal measure of preference intensity on the disutility of the act of voting. The second is due to Goodman and Markowitz [1952] and uses the psychological notion of "just-noticeable-difference" between alternatives: each individual has a fine number of indifference levels called levels of discretion and "a change from one level to the next represents the minimum difference which is discernible to an individual" (p. 259). Their fundamental assumption then is to use, for every individual, the number of discernible discretion levels between a pair of alternatives as the common measure of the strength of his preference. With this assumption and other conditions that are very similar to conditions used by Milnor [1954] and to conditions that shall be introduced below to characterize utilitarianism, they obtain that the social evaluation of an alternative is the sum of the individual utilities associated to that alternative. However, the difficulty in using levels of discretion as a means for interpersonal comparisons is well put in an example given by Luce and Raiffa [1957]:

> Consider two individuals, s_1 and s_2 who have to select one of two candidates A_1 or A_2; candidate A_1 is preferred by s_1 and A_2 by s_2. To resolve the strength of preference problem, these voters are also asked to rank some nonavailable candidates, $A_3, A_4, \cdots, A_{100}$. Voter s_2 is very discerning, and he ranks the candidates A_2 over A_3 over A_4 $\cdots$ over A_{99} over A_{100} over A_1. On the other hand, voter s_1 is dedicated to a single issue and he divides all candidates into two camps, namely, those who are "for" it and those who are "against" it; he is indifferent among the "for's" and indifferent among the "against's". In this case s_2 will have his way since he can discern 99 difference levels between A_2 and A_1, whereas s_1 can only discern one indifference level between them. But who is to say that s_2 feels more strongly than s_1. (pp. 347–348)

The discussion has been pursued recently by Ng [1975]. He argues that empirical studies have shown that thresholds for various feelings almost coincide for different individuals and that the number of just-noticeable-differences should not vary too much from one to the other. This argument, however, does not provide a fundamental reason to use his method. Such a reason may better be looked for in his attempt to ground his conventions on an approach in terms of expected utility.[6] This type of justification, based on assumptions about rational behavior in the face of risk (or uncertainty) has been advocated by several authors and will be discussed in Section 4. This type of justification is also linked to the third approach, purporting to found

[6] Historically the approach based on discrimination levels can be traced back to Borda [1781] and Edgeworth [1881]. It was analyzed more extensively by Armstrong [1951]. See also the discussion in Rothenberg [1961] and Arrow ([1951, 1963], p. 115). Concerning the discussion (and references) on the expected utility approach and the problem of measuring a single person's von Neumann-Morgenstern utility function, see Luce and Raiffa ([1957], p. 34). For a critique and an alternative model see Kahneman and Tversky [1979].

interpersonal comparisons mentioned by Arrow, which has received much attention. This is based on a principle of "extended sympathy" (Suppes [1966], sen [1970], Chs. 9 and 9*). The idea is to adopt an individual viewpoint and consider comparisons of the form: I prefer to be in the position of individual j under alternative x than of individual k under alternative y. As we immediately see, this amounts to expand the set of alternatives on which all individuals can, hypothetically, exert their capacity for choice. Hence it is compatible with the choice interpretations of individual utility. However, this approach amounts to making interpersonal comparisons from the point of view of a single person. In other words, the difficulty of such comparisons is brought back again to the difficulty of measuring a single individual utility, but of a new kind. In addition, if the viewpoint of a single individual is adopted, then we must assume some special status for the individual (such as being an ethical observer) or, more fundamentally, we must introduce some way of identifying the viewpoints of all individuals (each being considered as a different ethical observer). This could be based either on the ethical justification of an "identity principle" for the moral person or on the ethical recognition of a "principle of cooperation" leading to some sort of consensus. Here the work of Pazner [1979] should be recalled. He well stressed the fact that the Arrow impossibility theorem was also applicable at the level of extended preferences. Indeed he showed (among other things) that reasonable conditions analogous to Arrow's conditions would imply a dictatorial ethical observer. And as he said, "even dictatorship in the present sense is troublesome, especially if one wishes to interpret extended sympathy orderings in terms of value judgments of a higher order" (p. 164, fn. 2).

Moreover, this identity (or consensus) problem is present in any empirical approach to the measurement of individual utility (be it of the first kind or of the second kind) and any attempt at solving it should be related to the use this measurement is to have. This is due to the fact that in all empirical studies only "typical" individuals can be considered. Whatever the method of interpersonal comparison used – whether the ones of Fisher and Frisch or the just-noticeable-difference method or, to take another example, the use of "proxies" like suicide rates, proposed by Simon [1974] – it is always assumed, for practical reasons, that the individuals are "representative individuals" or "average individuals" in some relevant classes.[7] That this identity problem should arise in any empirical study is even clearer if one realizes that, practically, the difficulty of the interpersonal comparisons is not very different from the difficulty of the comparisons of a single individual utility at various times (i.e., intrapersonal but intertemporal comparisons).[8]

In spite of this problem it seems that the "extended sympathy" approach has been formally the most fruitful. In Section 2, I introduce Sen's concept of "social welfare functional" and the related concept of "social welfare ordering", which may be viewed as particular formal representations of such an extended preference. In Section 3 various types of interpersonal comparisons are defined for this concept of social welfare ordering. Finally, in Section 4, I come back to the present discussion by considering the ethical theories of Rawls and Harsanyi. What is needed for the presentation is to start from a single extended preference. The examination of some aspects of these two, essentially distinct, theories will eventually lead to some ethical justification for such a procedure.

2 Welfarism

Since Arrow's [1951, 1963] particular formalization of the concept of social welfare functions, several other formalizations have been proposed. In this section we shall review some of these and try to relate them to the original Bergson-Samuelson concept. This review will provide some understanding of the issues involved in basing social welfare judgments on individual utility alone which, by definition, forms the welfarist point of view.

[7]On this point see Vickrey [1960], p. 522, or Little [1957], p. 49. It seems that it is also for practical purposes that Rawls introduces the notion of representative individual (see Rawls [1972], p. 128). Notice that Little is also conscious that this notion may help to solve the intertemporal and intergenerational problem. On this last problem and in relation to Rawls see Arrow [1973b], Dasgupta [1974], and Solow [1974].

[8]See the recent discussion of Mirrlees [1982] and his notion of isomorphic individuals.

2.1 From constitutions to social welfare orderings

We start with Arrow's definition of a social welfare function, which, following his suggestion, we shall call a "constitution". First we are given a (finite) set of individuals $N = \{1, 2, \cdots, i, \cdots, n\}$ and a set X of possible social alternatives. For our purpose we shall assume that the number of individuals is at least two and the number of alternatives at least three. A priori, however, the set of social alternatives is not more specified, but we shall see that some of the conditions introduced later imply some additional restrictions concerning X. Second we consider all possible preference orderings $\mathcal{R}$ over X and the set $\mathcal{R}^{\mathcal{N}}$ of all n-tuples $(R_1, R_2, \cdots, R_n)$ of preference orderings in $\mathcal{R}$, called the set of "individual preference profiles". As usual (see Sen [1970], Ch. 1), a preference ordering R in $\mathcal{R}$ will be assumed to be a reflexive, transitive, and complete binary relation on X. Also, for any $R, R', \overline{R}, \cdots$, in $\mathcal{R}$, we shall denote by $P, P', \overline{P}, \cdots$, the associated strict preference relation and by $I, I', \overline{I}, \cdots$, the associated indifference relation. Then a *constitution* is a function F from some subset of $\mathcal{R}^{\mathcal{N}}$ to $\mathcal{R}$, associating to each admissible individual profile some admissible ordering called the social ordering. This definition can be weakened to take into account the fact that the social choice of some alternative among a finite set of social alternatives does not require a social ordering of X. Indeed, the range of the function F has only to be the set of reflexive, complete, and acyclic preference relations on X, and F is then called a *social decision function* (Sen [1970], Ch. 1).

The second definition I want to introduce with the view to formalizing social welfare functions is due to Sen [1970]; it is motiviated by the need to consider interpersonal comparisons. It reintroduces in social choice individual utility functions. More precisely, instead of considering a domain in the set $\mathcal{R}^{\mathcal{N}}$ of all individual preference profiles it considers a domain in the set $\mathcal{U}$ of all real-valued functions on $X \times N$. Accordingly, for each i in N, the real-valued function $U(\cdot, i)$ defined on X is to be interpreted as i's utility function, and any U in $\mathcal{U}$ may be called an "individual utility profile". Also a *social welfare functional* (SWFL) can be simply defined to be a function F from some subset of $\mathcal{U}$ to $\mathcal{R}$. Finally, as before, if the range of F may include any reflexive, complete, and acyclic preference relation, then F is called a *social decision functional.*

To the extent that individual preference orderings can be represented by individual utility functions (i.e., xR_iy if and only if $U(x,i) \geq U(y,i)$), Sen's framework contains Arrow's approach to social welfare functions. Hence the conditions imposed on constitutions can be formulated in terms of SWFLs. However, to be equivalent to Arrow's original conditions some precision should be given about the significance of the individual utility indicators. This question is taken up in the next section.

To simplify the analysis performed here, we want to impose some basic conditions on the SWFLs. These conditions are very similar to some of Arrow's conditions. The first is a condition of

Unrestricted domain for SWFLs (UD): The SWFL F is defined for every U in $\mathcal{U}$.

This condition is broad and imposes implicitly some restriction on X. For instance, if we suppose that X is the distribution of some fixed total income, then it may seem unreasonable to assume, for some individual i in N, that his utility is decreasing with the income level. However one might adopt a more abstract viewpoint, from which the investigation is not limited to a specific social choice problem with some given particular set of social alternatives. Then X may be any set of social alternatives and no individual utility profile can be discarded a priori.

The next condition is the key condition for introducing the simplifying approach to be used in the following. The main justification for this restriction is practical. All the results that will be surveyed here were obtained under conditions as restrictive or nearly as restrictive as this one. This condition says that the evaluation of some social alternative x should be based on its corresponding welfare vector $(U(x,1), U(x,2), \cdots, U(x,n))$ and should not take into account nonwelfare characteristics of the alternatives themselves. More precisely, we define[9]

Strong neutrality (SN): For any U^1 and U^2 in $\mathcal{U}$ and any two pairs of social alternatives $\{a,b\}$ and $\{c,d\}$ if

$$(U^1(a,1), U^1(a,2), \cdots, U^1(a,n)) = (U^2(c,1), U^2(c,2), \cdots, U^2(c,n))$$

[9]Strong neutrality is defined here as in Sen [1977]. On neutrality see G. Th. Guilbaud [1952], May [1952], Guha [1972], and Blau [1976]. In relation to welfarism see also Sen [1979c] and Ng [1981].

and

$$U^1(b,1), U^1(b,2), \cdots, U^1(b,n)) = (U^2(d,1), U^2(d,2), \cdots, U^2(d,n))$$

then aR^1b if and only if cR^2d, where $R^1 = F(U^1)$ and $R^2 = F(U^2)$.

To illustrate the strength of this condition, Sen [1979a] considers the following example. He considers a utility matrix involving the evaluation of two social outcomes by two individuals:

	Outcome 1	Outcome 2
Individual r	10	8
Individual p	4	7

In a first case individual 1 is a rich person r and individual 2 a poor person p, outcome 2 is a social state b where some redistributive taxation is adopted (but r remains richer), and outcome 1 the status quo, say state a. Now for the second case, Sen gives the following short parable.

> Let r be the rider of a motor cycle – joyful, rich, in good health and resilient – while p is a pedestrian – morose, poor, ill in health and frustrated. In state d (outcome 1 here) the rider gleefully goes by; in state c (outcome 2) he falls inadvertently into a ditch, breaking his bike and getting bruised badly. The rider is worse off in c than in d, while the pedestrian, who has not caused the accident in any way, thoroughly enjoys the discomfiture and discomfort of the rider (I could kill myself laughing looking at that crestfallen Angel!). The utility values of r and p are the same in this case ... (p. 477).

Strong neutrality would imply that b (taxation) should be socially preferred to a (no taxation) if and only if d (accident) is socially preferred to c (no accident). However the moral intuition of many people could be hurt by the pedestrian taking pleasure from the rider's discomfiture and discomfort. The argument against the condition is that some nonutility information should be used to assert the morality of outcomes. Moreover, this nonutility information should not be introduced simply to replace utility information when this one is insufficient, but to bring in essential elements like personal motivations and values (liberty, historical rights) or general moral principles. This argument concerns directly the way the utility functions are interpreted and the determination of the sets of social outcomes that are to be ranked. Clearly, the more the utility measures are interpreted in descriptive terms (introducing such notions as "pleasure" or "desires) – as opposed to normatively constrained interpretations – and the more arbitrary is the set X of admissible social states, the stronger will this argument be.[10]

We shall return to this issue later and analyze now the consequences of these basic conditions. Indeed, a third formalization of the concept of social welfare functions results from these basic conditions. Specifically, the following theorem shows that a SWFL satisfying UD and SN generates an ordering on the Euclidean space E^N of welfare vectors (E denotes the real line and E^N the n-dimensional space indexed by the names of the individuals). Such an ordering has been called by Gevers [1979] a *social welfare ordering.* Formally we shall identify *welfarism* to the existence of such a social welfare ordering. This is justified by the following theorem (see also Theorem 2.3 below).

Theorem 2.1.1 *(welfarism theorem): Let F be a SWFL satisfying UD and SN. Then there is a social welfare ordering R^* such that, for all x and y in X and U in $\mathcal{U}$,*

$$uR^*v \text{ if and only if } xRy$$

where $u = (U(x,1), U(x,2), \cdots, U(x,n)), v = (U(y,1), U(y,2), \cdots, U(y,n)),$ and $R = F(U)$.

Proof: By UD, for any u and v in E^N there are U in $\mathcal{U}$, a and b in X such that $u = (U(a,1), U(a,2), \cdots, U(a,n))$ and $v = (U(b,1), U(b,2), \cdots, U(b,n))$. Hence, we let uR^*v if and only if aRb, for $R = F(U)$. However, by SN, for any other V in $\mathcal{U}$ and any c and d in X, such that $(V(c,1), V(c,2), \cdots, V(c,n)) = u, (V(d,1), V(d,2), \cdots, V(d,n)) = v$ and $R' = F(V)$, we have $cR'b$ if and only if aRb if and only if uR^*v. Hence R^* is well defined. The fact R^* is reflexive and complete results immediately from the reflexivity and completeness of $F(U)$

[10] For instance, Sen [1979a] excludes the definition of X as a set of actions, or rules of actions (the use of a SWFL is not to establish the "rightness" of actions), but he insists on the descriptive content the utility measures should have.

(for any U). To prove transitivity take any u, v and w in E^N such that uR^*v and vR^*w. By UD, there are a, b, and c in X and U in $\mathcal{U}$ such that $u = (U(a,1), U(a,2), \cdots, U(a,n))$, $v = (U(b,1), U(b,2), \cdots, U(b,n))$, $w = (U(c,1), U(c,2), \cdots, U(c,n))$, and, letting $R = F(U)$, aRb and bRc. Since R is an ordering we get aRc and, hence, uR^*w. The result follows. ■

2.2 The single-profile approach

If we come back now to the original concept of a Bergson-Samuelson social welfare function, Arrow's constitution may be directly related to this concept, if interpreted as being a social ordering of X. Indeed a constitution may be viewed as associating to each individual preference profile a certain Bergson-Samuelson social welfare function. Furthermore, if we write the Bergson-Samuelson function as it is sometimes, but ambiguously, written: $W = f(U_1, U_2, \cdots, U_n)$ then the SWFL or the SWO formalization seems to be adequate. As explained by Sen [1977]

> If W and U_i are taken not be welfare *numbers* but *functions* defined over X, then welfarism is not, in fact, implied. (Indeed the $f(\cdot)$ will be very like a social welfare *functional* SWFL defined [above], with W being a real-valued representation of the social R determined by a SWFL.) However, it appears that this "functional" interpretation of $f(\cdot)$ was not intended in the formulations in question (see the operations of Samuelson (1947, p. 246, Eq. (31)) and Graaff (1957, Fig. 7(b)). And if $(U_1, \cdots, U_n)$ is simply a vector of individual utilities, then welfarism will follow $\cdots$ (p. 1,566)

However the usefulness of these formalizations have been greatly contested by some authors (see recently Samuelson [1977]). These authors insist on the fact that traditional (new) welfare economics are only dealing with social choice for only one given individual preference profile (a utility profile): "it could be *any* one, but it is *only* one" (Samuelson [1967], p. 49). Therefore all conditions imposed on either constitutions or SWFLs that involve comparisons for different individual profiles and are at the basis of most major results in social choice theory, and, above all, Arrow's impossibility theorem, should be rejected. In particular the SN condition is not admissible and therefore the welfarism theorem is not relevant. The conclusion seems to be rather ambiguous: in some way the welfarist formalism appears to be very close to welfare economics presentations and in another it has inadmissible foundations.

A new perspective in this debate has been brought by contributions of Parks [1976], Kemp and Ng [1976], Hammond [1976b], Roberts [1977], and Pollak [1979]. They developed a social choice theory within a single-profile approach obtaining, for instance, analogues to Arrow's [1951, 1963] impossibility theorem, Sen's [1970] theorem characterizing the Pareto extension rule, and May's [1952] theorem characterizing majority voting. In particular this development provides a basis to come back to welfarism, even if we want to restrict it to a single individual profile, by giving conditions similar to conditions UD and SN. The idea is to strengthen the structural implications of these conditions on the set of admissible social alternatives. (For other arguments, see Rubinstein [1979] and Bordes [1980].) Now the term *single profile* is somewhat misleading if we talk about individual utility profile. In that case utility functions are taken as representations of individual preferences and they are unique up to some admissible transformation. For example, in the original setup of the new welfare economists, each individual utility function is unique up to a strictly monotone transformation. In general to any individual utility profile $\overline{U}$ in $\mathcal{U}$ one can associate the set $[\overline{U}]$ of all individual utility profiles that are to be considered as equivalent representations, so that $F(U) = F(\overline{U})$ for all U in $[\overline{U}]$. We shall return to this unicity problem later. For the moment the problem is to impose conditions on some SWFL that are relative to some particular given individual utility profile $\overline{U}$ in $\mathcal{U}$. Following Roberts [1977] we define UP.

Unrestricted individual utility profile (UP): The individual utility profile $\overline{U}$ in $\mathcal{U}$ is such that, for any u, v and w in E^N there exist distinct x, y and z and X for which

$$(U(x,1), U(x,2), \cdots, U(x,n)) = u$$
$$(U(y,1), U(y,2), \cdots, U(y,n)) = v$$

and

$$(U(z,1), U(z,2), \cdots, U(z,n)) = w \text{ for some } U \text{ in } [\overline{U}].$$

It is clear that the satisfaction of this condition depends on the nature of the social alternatives considered in a particular situation. In an ordinal context this is denoted U^{*3} by Pollak [1979], who provides the following examples.

> If we think of social states as "social consumption vectors" specifying (among other things) the quantity of one or more continuously divisible goods that each selfish individual is to receive then U^{*3} is highly plausible. But neither selfishness, the presence of many consumption goods, nor their continuous divisibility is required for U^{*3}. For example, if there is a subset of social states that differ only in each individual's consumption of a single good, and if each individual's allotment of that good can assume at least three distinct values (small, medium, and large), and if each individual's ordering of the states within the subset depends only on his own consumption, then U^{*3} is automatically satisfied.

The other condition is the analogue of SN, but here it is relative to $\overline{U}$, the given individual utility profile. Hence the SWFL has only to be defined on $[\overline{U}]$.

Relative neutrality (RN): Given $\overline{U}$ in $\mathcal{U}$, for any U in $[\overline{U}]$ and any two pairs of social alternatives $\{a, b\}$ and $\{c, d\}$ if

$$(\overline{U}(a,1), \overline{U}(a,2), \cdots, \overline{U}(a,n)) = (U(c,1), U(c,2), \cdots, U(c,n))$$

and

$$(\overline{U}(b,1), \overline{U}(b,2), \cdots, \overline{U}(b,n)) = (U(d,1), U(d,2), \cdots, U(d,n)),$$

then $a\overline{R}b$ if and only if $c\overline{R}d$, where $\overline{R} = F(\overline{U}) = F(U)$.

This condition is subject to the same kind of objections as strong neutrality in the multiprofile approach. The interesting fact, proved below, about these conditions is that given some $\overline{U}$ in $\mathcal{U}$ and some SWFL F defined and constant on $[\overline{U}]$ and a set of social alternatives such that UP and RN are both satisfied then it is possible to extend F in a natural way such that both UD and SN are satisfied (for this and related results, see Roberts [1977, 1980c]). This result in turn allows to construct a SWO R^* on E^N by application of the welfarism theorem.

Theorem 2.2.2 *Suppose $\overline{U}$ in $\mathcal{U}$ and a SWFL $\overline{F}$ defined on $[\overline{U}]$ such that $\overline{F}(U) = \overline{F}(\overline{U}) = \overline{R}$ for all U in $[\overline{U}]$. Assume UP and RN are satisfied. For any U in $\mathcal{U}$ and any x and y in X let xRy and $F(U) = R$ if and only if, for some V in $[\overline{U}]$ and a and b in X,*

$$(U(x,1), U(x,2), \cdots, U(x,n)) = (V(a,1), V(a,2), \cdots, V(a,n))$$
$$(U(y,1), U(y,2), \cdots, U(y,n)) = (V(b,1), V(b,2), \cdots, V(b,n))$$

and $a\,\overline{R}\,b$.

Then F satisfies UD and SN and $F(U) = \overline{F}(U)$ for all U in $[\overline{U}]$.

Proof: We start by showing that for any U in $\mathcal{U}$, for which $F(U) = R$ is defined, then R is an ordering of X. To prove the transitivity of R take any x, y, z in X such that xRy and yRz. Then there are a, b, c in X and V in $[\overline{U}]$ such that

$$(U(x,1), \cdots, U(x,n)) = (V(a,1), \cdots, V(a,n))$$
$$(U(y,1), \cdots, U(y,n)) = (V(b,1), \cdots, V(b,n))$$
$$(U(z,1), \cdots, U(z,n)) = (V(c,1), \cdots, V(c,n))$$
$$a\,\overline{R}\,b \quad \text{and} \quad b\,\overline{R}\,c.$$

By the transitivity of $\overline{R}$ we get $a\overline{R}c$ and, hence, xRy. Completeness and reflexivity can be proved similarly. Now UP implies directly that for any U in $\mathcal{U}$, any x and y in X, there is some V in $[\overline{U}]$ and some a and b in X such that

$$(U(x,1), \cdots, U(x,n)) = (V(a,1), \cdots, V(a,n))$$

and

$$(U(y,1), \cdots, U(y,n)) - (V(b,1), \cdots, V(b,n)).$$

Hence F is a SWFL defined on all of $\mathcal{U}$; that is, it satisfied UD. To see that it satisfies SN, take any U^1 and U^2 in $\mathcal{U}$, and any two pairs $\{a, b\}$ and $\{c, d\}$ satisfying the antecedent of this condition. Then by construction of F there is V in $[\overline{U}]$ and $\bar{a}$ and $\bar{b}$ in X such that

$$(U^1(a,1), \cdots, U^1(a,n)) = (U^2(c,1), \cdots, U^2(c,n)) = (V(\bar{a},1), \cdots, V(\bar{a},n)),$$
$$(U^1(b,1), \cdots, U^1(b,n)) = (U^2(d,1), \cdots, U^2(d,n)) = (V(\bar{b},1), \cdots, V(\bar{b},n)),$$

and $a\,R^1\,b$ if and only if $\bar{a}\bar{R}\bar{b}$ if and only if $c\,R^2\,d$.

Therefore aR^1b if and only if cR^2d and the result follows. ■

This theorem is interesting not only because it brings us back to welfarism but also because it shows how properties or conditions imposed in a single-profile SWFL approach have their analogues in the multiprofile SWFL approach, which in turn have their analogues in the social welfare ordering approach. As another example take the Pareto condition as used by Arrow. In the single-profile approach it can be stated as follows.

Weak Pareto (for a single profile $\bar{U}$): For any x and y in X, if $\bar{U}(x,i) > \bar{U}(y,i)$, for all i in N, then $x\bar{P}y$, for $\bar{P}$ associated to $\bar{R} = F(\bar{U})$.

Whereas in the SWFL approach it becomes the following.

Weak Pareto (WP): For any U in $\mathcal{U}$ and any x and y in X, if $U(x,i) > U(y,i)$, for all i in N, then xPy, for P associated to $R = F(U)$.

The analogous condition for SWOs will be given in the next section.

2.3 Neutrality reconsidered

Now let us return to the debate on Bergson-Samuelson social welfare functions. In a reply to Kemp and Ng [1976], Samuelson [1977]asserts that the Bergson-Samuelson social welfare function reflects the judgment of an ethical observer and that Arrow's constitution notion should only be considered as an attempt to formalize the design of political processes. In other words, "mathematical politics" should be clearly separated from "mathematical ethics". As well explained in Pollak [1979], Samuelson's final objection against most of social choice theory is not so much that it is limited to a multiprofile approach, since it can be reformulated in a single-profile approach, but that this reformulation requires a neutrality condition. Hence there is a convergence of opinions against neutrality: Samuelson, like Sen, considers that nonwelfare characteristics should intervene in a crucial way in the judgments of an ethical observer.

An this, of course, goes against the implications of strong neutrality and unrestricted domain as shown by Theorem 2.1. Moreover, under unrestricted domain, strong neutrality is equivalent to two other conditions which have been extensively used in social choice theory.[11] These conditions for SWFLs are as follows.

Pareto indifference (PI): For any U in $\mathfrak{U}$ and any x and y in X, if $U(x,i) = U(y,i)$, for all i in N, then xIy, for I associated to $R = F(U)$.

Independence of irrelevant alternatives for pairs (IR): For any U^1 and U^2 in $\mathfrak{U}$ and any *pair* x and y in X, if $U^1(x,i) = U^2(x,i)$ and $U^1(y,i) = U^2(y,i)$, for all i in N, then xR^1y if and only if xR^2y, where $R^1 = F(U^1)$ and $R^2 = F(U^2)$.

In Sen [1977], it is shown that, under UD and PI, for SWFLs, IR is equivalent to SN. Here we provide the following (logically close) statement.

Theorem 2.3.3 *For any SWFL satisfying UD, the conditions PI and IR hold if and only if SN holds.*

Proof: That SN implies PI and IR is easy to see. Indeed, for PI, it is enough to take, in SN, $U^1 = U^2$, $a = d = x$ and $b = c = y$. For IR it is enough to put $a = c$ and $b = d$ in SN.

To prove the converse, suppose U^1 and U^2 in $\mathfrak{U}$ and two pairs $\{a, b\}$ and $\{c, d\}$ (not necessarily distinct) are such that

$$(U^1(a,1), \cdots, U^1(a,n)) = u = (U^2(c,1), \cdots, U^2(c,n))$$
$$(U^1(b,1), \cdots, U^1(b,n)) = v = (U^2(d,1), \cdots, U^2(d,n)).$$

Since UD holds and $|X| \geq 3$, we may take e in X, and U^3, U^4 and U^5 in $\mathfrak{U}$ such that if

[11] For example, for PI, Samuelson [1947] (condition (5), p. 223): "If any movement leaves an individual on the same indifference curve, then the social welfare function is unchanged ..." For IR, see Arrow [1951, 1963].

$\{a,b\} = \{c,d\}$ then $a \neq e \neq b$, and

$$
\begin{aligned}
&(U^3(a,1),\cdots,U^3(a,n)) = (U^3(e,1),\cdots,U^3(e,n)) = u,\\
&(U^3(b,1),\cdots,U^3(b,n)) = v;\\
&(U^4(c,1),\cdots,U^4(c,n)) = (U^4(e,1),\cdots,U^4(e,n)) = u,\\
&(U^4(d,1),\cdots,U^4(d,n)) = v;\\
&(U^5(e,1),\cdots,U^5(e,n)) = u,\\
&(U^5(b,1),\cdots,U^5(b,n)) = (U^5(d,1),\cdots,U^5(d,n)) = v.
\end{aligned}
$$

Then, assuming IR, we get

$$
\begin{aligned}
&aR^1b \text{ if and only if } aR^3b, &&cR^2d \text{ if and only if } cR^4d,\\
&eR^3b \text{ if and only if } eR^5b, \text{ and } &&eR^4d \text{ if and only if } eR^5d.
\end{aligned}
$$

Also, assuming PI, we have

$$a\,I^3\,e, \quad c\,I^4\,e, \quad \text{and } b\,I^5\,d.$$

Combining all these relations, finally we get

$$a\,R^1\,b \text{ if and only if} \, c\,R^2\,d.$$

■

This result means that the welfarist approach amounts to imposing the conditions of unrestricted domain, Pareto indifference, and independence of irrelevant alternatives on SWFLs.

Roberts [1980b], in a similar approach, does not impose Pareto indifference and hence allows SWFLs for which nonwelfare characteristics may play a role. However, since he still imposes UD, IR, and WP, his results still have strong "welfarist" implications. In fact he proves that under these conditions, a "weak neutrality" property holds. This is done by introducing, for an admissible SWFL, a notion of "strong strict social preference" over pairs of social alternatives that cannot be affected by nonwelfare characteristics in the same way as, with Pareto indifference, the whole social ordering could not be affected. Then, as we shall do in the following, but with strong neutrality, he analyzes the possibilities for social orderings under different informational bases.

On the other hand, the negative implications for social choice of the neutrality features underlying the weak Pareto conditions, combined with unrestricted domain, seem to be already explicitly and extensively exploited in the social choice literature. Indeed these features play a crucial role for two well-known paradoxes: Arrow's [1951, 1963] theorem on *the impossibility of a social welfare function* and Sen's [1970] theorem on *the impossibility of a Paretian liberal.* Since we shall return to the former theorem in the next section, let us now examine the latter as well as the extensive discussion it has triggered. However, this we shall do briefly; for a complete reappraisal of the paradox, see Sen [1976a].

2.4 The Paretian liberal and the multiplicity of social choice problems

The "Paretian liberal" paradox is due to the consideration of social alternatives that should be viewed as belonging to some individual personal domain of decisions. For instance, if two social alternatives differ from each other only in some matter that is considered part of individual i's privacy – for example, that i sleeps on his back or on his belly – then the respect of his liberty would require that i alone should decide in the choice between x and y. More specifically, Sen [1970] introduces a "weak libertarianism" condition requiring that for each individual i there is at least one pair of social alternatives x and y such that if i prefers x to y (resp. y to x), then society should prefer x to y (resp. y to x). Formally Sen [1970] has established the impossibility of finding a social decision function satisfying simultaneously the conditions (as stated for social decision functions) of unrestricted domain, weak Pareto, and weak libertarianism. We shall not prove this theorem. However, the essence of this paradox is well explained by Sen's original example.

There are three social alternatives involving the reading of a copy of D.H. Lawrence's book *Lady Chatterley's Lover* by a "prude" Mr. A or by a "lascivious" Mr. B or by neither. A prefers most that no one read the book but would rather read it himself then let B read it. On the other hand, B enjoys above all that A read this book and, of course, prefers to read it himself than having no one reading it. Hence, by weak libertarianism, B reading the book should be socially preferred to no one reading it. Now both prefer A to read the book rather than B so that, by weak Pareto, A reading the book should be socially preferred to B reading it. Finally, since A prefers no one reading the book to reading it himself, if one applies weak libertarianism again then one gets a social preference cycle.

Many authors have discussed this paradox and proposed some way out of the difficulty by relaxing one condition or another. Sen [1980] argues that the plurality of the possible solutions "shows how a variation of non-utility description can precipitate different moral judgments" and that "this is, of course, contrary to the essence of welfarism. Nonutility information relating to *how* 'personal' the choices are, what *motivation* the persons have behind their utility rankings, whether the interdependence arises from liking or disliking the others' physical acts (in this case the reading of the book) or from the *joys and sufferings* of the others, etc. $\cdots$, may well be found to be relevant in deciding which way to resolve the conflict." It seems, though, that more fundamentally it is the question of the plurality of the interpretation of the "aggregation problem" itself that is concerned. Even if one restricts one's interpretation to social welfare evaluation, several particular models of social choice can be studied involving different sets of alternatives.

This interpretation question is typically involved in Nozick's [1974] proposed solution, where the set of alternatives is so restricted that individual rights are already taken into account: "Rights do not determine a social ordering but instead set the constraints within which a social choice is to be made, by excluding certain alternatives, fixing others, and so on. $\cdots$" (p. 165). This suggests that, contrary to the universality of the original utilitarian project, one should introduce the problem of social choice at different levels, the choices made at some level being the norms constraining the choices to be made at the next. The levels would be determined by the nature of the social alternatives. In any case the importance of the nature of the set of alternatives is already et either xplicit in Sen's starting distinction between choices that are personal and those that are not, and justifies Seidl's [1975] criticism of the existential form of the libertarianism condition. This is also explicit in one of the Farrell's [1976] proposed solutions, by which the set of social alternatives is partitioned into socially equivalent subsets of alternatives differing in matters that are private to some individual.

However, it is not only the nature of the alternatives that is involved; the nature of the preferences or the interpretation of the individual utilities is also essential. This kind of paradox arises because the preferences of individuals are such that "the individuals are meddling in each other's affairs" – to use Blau's [1975] terminology – or from "one person's taking a perverse interest in the affairs of another" (Gibbard [1974], p. 398). Hence to solve the conflict one may on the one hand, as Blau and Gibbard do, consider that libertarian rights are alienable in some way or another.[12] On the other hand, one may consider either that the preferences themselves should be amended (an alternative solution of Farrell [1976]) or that only some part of it should count (Sen [1976a]). However, to choose between these two types of solutions seems to be more than to choose between accepting and rejecting the welfarist implications of the Pareto condition. It indicates the fundamental difficulty of a unified and universal approach to social choice.

3 Axioms for social welfare orderings

We have seen that, under some conditions, the derivation of a collective preference ordering over the set X, from the knowledge of individual utility functions, can be adequately simplified by the introduction of a social ordering on the n-dimensional Euclidean space E^n indexed by the name of the individuals. Thus instead of introducing axioms on SWFLs and then showing their equivalent formulation for SWOs, we shall introduce them directly for SWOs. It should be kept in mind, however, that this simplification can be avoided and even should be avoided if non-welfare considerations are to be introduced.

3.1 A first approach to utilitarianism and leximin

As mentioned in the introduction, the most traditional rules to evaluate alternative economic policies are utilitarian, even though the justifications for such rules have greatly evolved in

[12]See also Suzumura [1980] for a critical analysis of the related Gibbard [1974] and Kelly [1976] system of alienable rights.

parallel to the evolution of the concept of utility. Given following is the definition of classical utilitarianism as a particular kind of SWOs.

Utilitarianism: The *pure utilitarian* SWO R^* is such that, for any u and v in E^N, uR^*v if and only if $\sum_{i=1}^n u_i \geq \sum_{i=1}^n v_i$.

Even though Rawls's theory of justice seems to be rather in opposition to the present pure welfarist approach, it remains that the "difference principle" it advocates can be expressed in similar terms.

The difference principle: For any u and v in E^n, if $\min_i u_i > \min_i v_i$ then uP^*v.

However, this principle, known also as the maximin rule, does not give as such a SWO. Instead we shall concentrate on Sen [1970] proposed modification, as an alternative SWFL, which is simply a lexicographic completion of it. Before presenting it, as a SWO, we need the following notation. For any $u \in E^N$, let us define a function $i(\cdot)$ from N to itself such that, for every h and k in N,

$$u_{i(h)} < u_{i(k)} \text{ implies } h < k.$$

In other terms, for every $u \in E^N$, we have a particular function $i(\cdot)$ such that

$$u_{i(1)} \leq u_{i(2)} \leq \cdots \leq u_{i(n)};$$

that is, the function $i(\cdot)$ gives the ranking of the components of u. When there are ties, they may be broken arbitrarily. Note that $i(k)$ does not necessarily denote the same individual in u and in v. We may now write the leximin rule.

The leximin rule: Let R^* be the SWO such that, for any u and v in E^N, uP^*v if and only if

$$\begin{aligned} &\text{for some } k \text{ in } N, \quad && u_{i(k)} > v_{i(k)} \\ &\text{for all } h < k, && u_{i(h)} = v_{i(h)}. \end{aligned}$$

In the first part of Section 3 we shall mainly concentrate on these two rules and reserve for the end the introduction of other rules. We start by examining the basic properties of such

SWOs. Only such an examination could give some kind of justification for using one of these principles as an evaluation procedure for social choice. The first set of properties that can be introduced is very general and the properties will not allow discrimination between these two rules. We shall see that they are pure ordinal properties. The first is the weakest and we stated it for SWFLs.

Weak Pareto (WP)*: For any u and v in E^N, if, for all i in N, $u_i > v_i$, then uP^*v.

In other words, this says that if some welfare vector u strictly dominates some other welfare vector v, with respect to every individual component, than it should be strictly preferred by the collectivity. A stronger form of this Paretian principle, which corresponds to the Pareto-dominance concept used by most economists, is the following.

Strong Pareto (SP)*: For any u and v in E^N, if for all i in N, $u_i \geq v_i$, then uR^*v; if, in addition, $u_j > v_j$, for some $j \in N$, then uP^*v.

On the other hand, once welfarism is accepted the two axioms seem reasonable. Indeed, if it is granted that utility indices take into account all elements that should (or should not) influence social choice (and, for some, these would include interpersonal feelings as envy and pity, or even moral considerations about fairness, liberty, equality, or all kinds of rights), then the concomitant betterment of all such indices should be admitted as a collective improvement.

In the same spirit, we may state another property of the two foregoing rules as acceptable. It is a separability condition which says that the ordering of two welfare vectors should be independent of the welfare of all unconcerned individuals, namely those individuals who have the same utility level in the two vectors. Formally we get a strong property first proposed by Fleming [1952] in his characterization of utilitarianism (see Young [1974], Strasnick [1975], and Arrow [1977]).

Separability (SE)*: For any u^0, v^0, u^1 and v^1 in E^N, if every individual i in M, a subset of N, is such that $u_i^0 = u_i^1$ and $v_i^0 = v_i^1$ but if every other individual j in $N \setminus M$ is such that $u_j^0 = v_j^0$

and $u_j^1 = v_j^1$, then $u^0 R^* v^0$ if and only if $u^1 R^* v^1$.

The first set of properties does not raise any question concerning the comparability of the utility measurements. However utilitarianism and leximin as such do raise this kind of question. In particular they raise the question of interpersonal utility comparisons because they treat all the individual symmetrically. Indeed, both satisfy an axiom of anonymity.

Anonymity (A)*: If σ is a permutation of N (i.e., a bijection from N to itself) and if u and v in E^N are such that $u_i = v_{\sigma(i)}, i = 1, 2, \cdots, n$, then uI^*v.

Further, they both satisfy a condition that is a generalization both of A* and SP* and which was proposed by Suppes [1966] as follows.

The grading principle: For any u and v in E^N, if for some permutation σ of N and for all i in N, $u_i \geq v_{\sigma(i)}$, then uR^*v; and if, moreover, $u_j > v_{\sigma(j)}$ for some j, then uP^*v.

It is important to note that the grading principle implies a partial ordering of utility vectors. A more general formulation of this principle which we shall use later is the following.

The m-person grading principle: A SWO R^* satisfies the m-grading principle, $1 \leq m \leq n$, if, for every subset M of m individuals and all u and v in E^N, with $u_h = v_h$ for every h not in M, and for any permutation σ of M,

$$uR^*v \text{ whenever } u_i \geq v_{\sigma(i)} \text{ for all } i \text{ in } M,$$

and

$$uP^*v \text{ whenever } u_i \geq v_{\sigma(i)} \text{ for all } i \text{ in } M \text{ and } u_j \geq v_{\sigma(j)} \text{ for some } j \text{ in } M.$$

This will permit us to demonstrate a characteristic of this principle that is true of many other equity principles and that we shall use recurrently in the sequel: to require it in general it is sufficient to require it in situations where only two persons are nonindifferent. Paraphrasing Sen [1979b], for the leximin, we may state the following lemma (see Hammond [1979]).

Lemma 3.1.1 *(The grading principle from inch to ell): The 2-grading principle implies A^* and SP^* that imply the m-grading principle, $1 \leq m \leq n$.*

Proof: Suppose u and v in E^N are such that $u_i = v_{\sigma(i)}$, for all i in N and some permutation σ of N. Now a result on permutations permits the assertion of the existence of a sequence of pairs $N_1, N_2, \cdots, N_p$ in N, a sequence of permutations $\sigma_1, \sigma_2, \cdots, \sigma_p$ on N and a sequence of welfare vectors $u^1, u^2, \cdots, u^p$ in E^N such that, for $1 \leq k \leq p$,

$$\begin{aligned}
&\sigma_k(i) = i \text{ for } i \text{ not in } N_k,\\
&\sigma_p(\sigma_{p-1}(\cdots(\sigma_1(i))\cdots)) = \sigma(i),\\
&u^0 = u, u_i^{k-1} = u^k_{\sigma_k(i)} \text{ for all } i \text{ in } N.
\end{aligned}$$

Hence, applying the 2-grading principle p times, we get

$$u = u^0\, I^*\, u^1\, I^* \cdots I^*\, u^p = v.$$

So that A^* holds.

To prove SP*, take any u and v in E^N such that $u_i \geq v_i$ for all i in N (consider simultaneously the eventuality that, in addition, $u_j > v_j$ for some j). Take $v^0, v^1, \cdots, v^n$ such that $v^0 = v$, $v_k^k = u_k$ and $v_i^k = v_i^{k-1}$, for all $i \neq k$ and $k = 1, 2, \cdots, n$. Then, by the 2-grading principle, $v^k R^* v^{k-1}$ (and $v^k P^* v^{k-1}$ if $u_k > v_k$) and by transitivity uR^*v (and eventually uP^*v), since $v^n = u$.

Hence the 2-grading principle implies A* and SP*. To prove that A* and SP* imply the m-grading principle, for $1 \leq m \leq n$, consider any u and v in E^N such that $u_i = v_i$, for $m < i \leq n$, and for some permutation σ of $\{1, 2, \cdots, m\}$, $u_i \geq v_{\sigma(i)}, 1 \leq i \leq m$. Clearly, if $u_i = v_{\sigma(i)}$ for all i in $\{1, 2, \cdots, m\}$, then uI^*v by A*, since we can extend σ to a permutation τ of N for which $\tau(i) = i, m < i \leq n$. Otherwise we may define w in E^N such that $w_i = v_{\tau(i)}$, for all i in N. Then we get, by A*, wI^*v and, by SP*, uP^*w. Hence uP^*v and the proof is complete. ■

However may authors have proposed to reformulate utilitarianism in a generalized form so that it does not imply anonymity. For an $S \subset N$, let $E_+^S = \{u \text{ in } E^S : u_i \geq 0,\ \text{all } i \text{ in } S \text{ and } u \neq 0\}$.

Generalized utilitarianism: For λ in E_+^N, a *utilitarian* SWO R^* is such that: for any u and v in E^N,

$$uR^*v \text{ if and only if } \sum_{i=1}^{n} \lambda_i u_i \geq \sum^{n} \lambda_i v_i.$$

Hence a pure utilitarian SWO is simply a utilitarian SWO for which λ has all its components equal to the same positive number. However, the choice of the weights to assign to each individual utility index remains and, so, the problem of interpersonal comparability. This is linked also to the measurement scale that is used for each individual utility. The way to specify these measurability-comparability requirements is to determine the class of numerical transformations of the utility indices that leave the SWO invariant. This invariance reflects the type of measurement used to quantify all individual information that is available. It is therefore very crucial to notice that utilitarian rules as well as leximin are not compatible with any kind of informational base. More generally the type of invariance that is associated to the available information determines drastically the type of SWOs that are possible.

In the context of this discussion one way to interpret Arrow's impossibility theorem ([1951, 1963]) is particularly interesting. First Arrow used a property much weaker than anonymity that remains ordinal and involves no interpersonal comparisons.

Nondictatorship (ND)*: There is no i in N such that: for any u and v in E^N, if $u_i > v_i$ then uP^*v.

Second if one wants to keep the requirements of the new welfare economics, where utility indices are to be interpreted ordinally as reflecting the preferences of the individuals as they would be expressed by some choice behavior, then these utility indices would have meaning only up to a monotone transformation. This gives our first invariance axiom.

Ordinality and noncomparability (ON)* : Let $\varphi_1, \varphi_2, \cdots, \varphi_n$ be any strictly increasing numerical functions. Then, for any u and v in E^N,

$$uR^*v \text{ if and only if } \varphi(u)R^*\varphi(v),$$

where $\varphi(u) = (\varphi_1(u_1), \varphi_2(u_2), \cdots, \varphi_n(u_n))$, and $\varphi(v) = (\varphi_1(v_1), \varphi_2(v_2), \cdots, \varphi_n(v_n))$.

However, in a welfarist approach, such a property appears to be prohibitively restrictive as it is well known[13] from Arrow's theorem

Theorem 3.1.1 *If a SWO satisfies WP* and ND* then it cannot satisfy ON*.*

Therefore to build up some SWO having even the weakest Paretian property and the minimal egalitarian property of nondictatorship, an ordinal information base, which excludes interpersonal comparisons, is insufficient. In the next section we introduce different informational bases allowing for more discrimination.

3.2 Invariance axioms

The axioms we are going to list (nonexhaustively) below all restrict the kind of transformations that can be applied to the individual utility indices without affecting the social ordering. The analysis of the implications of these informational hypotheses for the equity content of collective choice was initiated by Sen. With respect to ON*, either cardinality or some interpersonal comparability is introduced.

The first introduces only cardinality by allowing all positive affine transformations.

Cardinality and noncomparability (CN)*: Let $a_1, a_2, \cdots, a_n$ be any numbers and $b_1, b_2, \cdots, b_n$ be any positive numbers. Then, for any u and v in E^N, uR^*v if and only if

$$(a_1 + b_1u_1, a_2 + b_2u_2, \cdots, a_n + b_nu_n)R^*(a_1 + b_1v_1, a_2 + b_2v_2, \cdots, a_n + b_nv_n).$$

However, in the present welfarist framework the introduction of cardinality does not change the situation for the following reason.[14]

[13] A proof for this theorem will be provided below (p. 52). For versions of this theorem with specific economic domains see Kalai, Muller, and Satterthwaite [1979] and Border [1983]. This last author, as Wilson [1972], drops the Pareto condition.

[14] See d'Aspremont and Gevers [1977]. This cardinality statement was first considered by Sen [1970] but in a limited case (for binary independence of irrelevant alternatives).

Lemma 3.2.1 *A SWO R^* satisfies CN^* if and only if it satisfies ON*.*

Proof: We shall only prove that CN* implies ON*, the other direction being trivial. Suppose u and v in E^N are such that uR^*v. Consider n strictly increasing numerical functions $\varphi_1, \varphi_2, \cdots, \varphi_n$. We have to show that, by CN*, $\varphi(u)R^*\varphi(v)$, where $\varphi(u) \equiv (\varphi_1(u_1), \cdots, \varphi_n(u_n))$ and $\varphi(v) \equiv (\varphi_1(v_1), \cdots, \varphi_n(v_n))$. However it is simple to find $a \in E^N$ and $b \in E_+^N$ such that $b_i > 0$,

$$a_i + b_i u_i = \varphi_i(u_i) \quad \text{and} \quad a_i + b_i v_i = \varphi_i(v_i), \qquad i = 1, 2, \cdots, n.$$

Hence $\varphi(u) = (a_1 + b_1 u_1, \cdots, a_n + b_n u_n), \varphi = (a_1 + b_1 v_1, \cdots, a_n + b_n v_n)$ and the result follows.

■

Cardinality and origin comparability (COC)*: Let a and $b_1 > 0$, $b_2 > 0, \cdots, b_n > 0$ be $n+1$ numbers. Then, for any u and v in E^N, uR^*v if and only if $(a + b_1 u_1, \cdots, a + b_n u_n)R^*(a + b_1 v_1, \cdots, a + b_n v_n)$.

Lemma 3.2.2 *A SWO R^* satisfied CN* if and only if it satisfied COC*.*

Proof: Again we need only prove that COC* implies CN*. Suppose u and v in E^N such that uR^*v and choose $a, b \in E^N$, $b > 0$. We have to show that, by COC*, $(a_1 + b_1 u_1, \cdots, a_n + b_n u_n)R^*(a_1 + b_1 v_1, \cdots, a_1 + b_n v_n)$. Choose, for that purpose, $\beta \in E_+^N$ and $\theta \in E$ such that

$$a_1 - \frac{b_1}{\beta_1} = a_2 - \frac{b_2}{\beta_2} = \cdots = a_n - \frac{b_n}{\beta_n} = \theta < \min_i a_i.$$

Then, by COC*, $u^1 R^* v^1$ with $u_i^1 = 1 + \beta_i u_1$ and $v_i^1 = 1 + \beta_i v_i$, $i = 1, 2, \cdots, n$. Also, by COC*, $u^2 R^* v^2$ with $u_i^2 = \theta + (a_i - \theta)u_i^1$ and $v_i^2 = \theta + (a_i - \theta)v_i^1, i = 1, 2, \cdots, n$. The result follows since, by construction, for every i in N

$$u_i^2 = a_i + b_i u_i \quad \text{and} \quad v_i^2 = a_i + b_i v_i.$$

■

Using Lemma 3.2.1 (resp. Lemma 3.2.2) we immediately get the following generalization of Theorem 3.1 first given by Sen [1970].

Theorem 3.2.2 *If a SWO satisfies WP* and ND* then it cannot satisfy CN* (resp. COC*).*

It seems that, in order to get possibility results, we need to introduce more comparability. Of course this can be done simply by supposing

Co-cardinality (or cardinality and comparability) (CC)*: Let a and $b > 0$ be two numbers. Then, for any u and v in E^N, uR^*v if and only if $(a + bu)R^*(a + bv)$.

Here both individual units and individual origins of the utility indices are common. Hence both interpersonal comparisons of welfare gains and interpersonal comparisons of welfare levels are permitted. Although this is not the strongest informational setup, it is already very demanding. Some would argue that even if interpersonal comparisons are introduced, an ordinal approach should be kept. This leads to the axiom of co-ordinality.

Co-ordinality (or ordinality and comparability) (OC)*: Let φ be a strictly increasing numerical function. Then, for any u and v in E^N, uR^*v if and only if

$$(\varphi(u_1), \varphi(u_2), \cdots, \varphi(u_n))R^*(\varphi(v_1), \varphi(v_2), \cdots, \varphi(v_n)).$$

If the ordering R^* results from some SWFL F defined on $X \times N$, this property means that U is imply an ordinal representation of a social preference over pairs (x, i) in $X \times N$. To say that a pair (x, i) is to be preferred to a pair (x, j) is to say that individual i is better off in state x than individual j in state y. Hence, in utility terms, welfare levels are comparable but not the welfare gains. The reverse situation is obtained by the property of

Cardinality and unit-comparability (CU)*: Let $a_1, a_2, \cdots, a_n$ and $b > 0$, be $n+1$ numbers. Then, for any u and v in E^N, uR^*v if and only if

$$(a_1 + bu_1, \cdots, a_n + bu_n)R^*(a_1 + bv_1, \cdots, a_n + bv_n).$$

This invariance axiom, allowing only for a common utility unit, may be linked to the utilitarian tradition. There the importance was attached to interpersonal comparisons of *marginal utilities* in a context of income distribution: does an additional unit of income increase more the welfare of individual i than the welfare of individual j? This is why invariance with respect to individual origins of utility was indifferent. However, even this may appear to be too strong a restriction to decide on some welfare issues, like a normative analysis of poverty or of population size. Making the origin both common and nonarbitrary is achieved through

Ratio scale and comparability (RC)*: Let b be any positive number. Then, for any u and v in E^N, uR^*v if and only if buR^*bv.

As well noted by Blackorby and Donaldson [1982] "this restriction allows negative and positive utilities to be treated in a qualitatively different way. Thus the origin may represent an interpersonally significant welfare position such as a poverty line" (p. 253).

On the other hand it may also be interesting from a welfare point of view to distinguish between an increase in everyone's welfare and a simple reduction of the unit of measurement of each utility index. But this would mean that it is possible to define a "natural" unit of measurement. In that case one could enunciate an axiom of

Difference comparability (DC)*: Let a be any number. Then, for any u and v in E^N, uR^*v if and only if $(u+\bar{a})R^*(v+\bar{a})$ where $\bar{a} = (a, a, \cdots, a) \in E^N$.

Of course to combine (RC*) with (DC*) would amount to reduce invariance to nothing, but would be extremely demanding with respect to welfare information: it would require having both a "natural" origin and a "natural" unit of measurement.

The last invariance axiom we shall define is given here to show how one can multiply the possibilities by varying the combinations and because it will be used in the sequel. It is called by Gevers [1979]

Almost co-cardinality (ACC)*: Let $a_1, a_2, \cdots, a_n$ and $b > 0$ be any numbers and let φ be any strictly increasing numerical function. Then, for any u and v in E^N, uR^*v if and only if $u^1R^*v^1$, where, for every i in N, $u_i^1 = a_i + bu_i, v_i^1 = a_i + bv_i$ and also $u_i^1 = \varphi(u_i)$ and $v_i^1 = \varphi(v_i)$.

This axiom combines OC* with CU* and as such prohibits interpersonal comparisons of utility gains within a particular utility vector.

3.3 A characterization of utilitarian rules and lexical individual dictatorship

After examining alternative information bases for social choice we shall return now to the two rules introduced earlier, namely utilitarianism and leximin. The axiomatization we shall review here will put forward the fact that it is mainly the type of discrimination that is admitted in the ordering of all utility vectors that determines the type of welfare rule that is admissible. In this section we present an axiomatic characterization of utilitarianism and as a by-product a proof of Arrow's theorem as stated previously. In the next we shall do the same exercise for leximin. In both cases the presentation will be much simplified by the welfarist approach we have adopted. For both rules we shall proceed in two steps: first we show that if the rule holds for utility vectors where only two persons are concerned then it holds for all utility vectors; second we characterize the rule for two-person situations. The approach we adopt here is, in some sense, intermediate between two basic approaches to utilitarianism already known in decision theory, one represented by Blackwell and Girschik [1954] and the other by Milnor [1954]. These are far from being the only axiomatic elucidation of utilitarianism. Others will be discussed in subsequent sections.

In order to proceed this way for utilitarianism we need some more definitions:

Generalized m-person utilitarianism: A SWO R^* is called m-utilitarian, $1 \leq m \leq n$, if for every subset M of m individuals there is some λ in E_+^M such that, for all u and v in E^N with $u_h = v_h$ for every h not in M,

$$uR^*v \text{ if and only if } \sum_{i \in M} \lambda_i u_i \geq \sum_{i \in M} \lambda_i v_i.$$

Weak m-person utilitarianism: A SWO R^* is called *weakly m-utilitarian*, $1 \leq m \leq n$, if for every subset M of m individuals there is some λ in E_+^M such that, for all u and v in E^N with $u_h = v_h$ for every h not in M,

$$\sum_{i \in M} \lambda_i u_i > \sum_{i \in M} \lambda_i v_i \text{ implies } uP^*v.$$

Then, again paraphrasing Sen [1977], we may state

Lemma 3.3.1 *(Utilitarianism from inch to ell): If a SWO is 2-utilitarian, then it is m-utilitarian for all m, $1 < m \leq n$.*

Proof: The argument goes by induction. More precisely, it is enough to show that 2-utilitarianism and m-utilitarianism for some m, $2 \leq m < n$, implies $(m+1)$-utilitarianism. Without loss of generality take $M = \{1, 2, \cdots, m\}$. By m-utilitarianism we may associate to M some λ^0 in E_+^M and some j in M such that $\lambda_j^0 > 0$. By 2-utilitarianism we may associate to $\{j; m+1\}\lambda^1$ in $E_+^{\{j,m+1\}}$ and suppose λ^0 is appropriately normalized (λ_j^1 must be positive, otherwise a contradiction would arise on u and v such that $u_h = v_h, h \neq j$), so that $\lambda_j^0 = \lambda_j^1$. Let λ in $E_+^{M\cup\{m+1\}}$ be equal to $(\lambda_1^0, \lambda_2^0, \cdots, \lambda_m^0, \lambda_{m+1}^1)$ and take any u and v in E^N such that $u_h = v_h$ for $h = m+2, m+3, \cdots, n$. We may easily find w in E^N such that

$$\begin{aligned} &\lambda_j w_j + \lambda_{m+1} w_{m+1} = \lambda_j u_j + \lambda_{m+1} u_{m+1} \\ &w_{m+1} = v_{m+1} \\ &w_k = u_k,\ j \neq k \neq m+1. \end{aligned}$$

By 2-utilitarianism, wI^*u; by m-utilitarianism, wR^*v if and only if $\sum_{i=1}^m \lambda_i w_i \geq \sum_{i=1}^m \lambda_i v_i$. Equivalently we get uR^*v if and only if $\sum_{i=1}^{m+1} \lambda_i u_i = \sum_{i=1}^{m+1} \lambda_i w_i \geq \sum_{i=1}^{m+1} \lambda_i v_i$. The result follows. ∎

Lemma 3.3.2 *(Weak utilitarianism from inch to ell): If a SWO is weakly 2-utilitarian, then it is weakly m-utilitarian for all m, $1 < m \leq n$.*

Proof: As above take $M = \{1, 2, \cdots, m\}$ and λ^0 in E_+^M with $\lambda_j > 0$, $j \in N$, such that weak m-utilitarianism applies. Consider, first, the case where, as above we may associate to $\{j, m+1\}$ some λ^1 in $E_+^{\{j,m+1\}}$ with $\lambda_j^1 > 0$ and appropriately normalized so that $\lambda_j^1 = \lambda_j^0$. Then, similarly, we define $\lambda = (\lambda_1^0, \lambda_2^0, \cdots, \lambda_m^0, \lambda_{m+1}^1)$ in $E_+^{M\cup\{m+1\}}$ and, for any u and v in E^N with $u_h = v_h$, $h > m+1$, and $\sum_{i=1}^{m+1} \lambda_i u_i - \sum_{i=1}^{m+1} \lambda_i v_i > 0$, we may construct w in E^N such that, for some k

in $\{1, 2, \cdots, m\}$,

$$\begin{aligned}
&\lambda_j w_j + \lambda_{m+1} w_{m+1} = \lambda_j u_j + \lambda_{m+1} u_{m+1} - \varepsilon, \text{ for } 0 < \varepsilon < \sum_{i=1}^{m+1} \lambda_i (u_i - v_i) \\
&w_{m+1} = v_{m+1} \\
&w_h = u_h, \ j \neq h \neq m+1.
\end{aligned}$$

By weak 2-utilitarianism, uP^*w; moreover, by construction

$$\sum_{i=1}^{m+1} \lambda_i w_i = \sum_{i=1}^{m+1} \lambda_i u_i - \varepsilon > \sum_{i=1}^{m+1} \lambda_i v_i$$

and, since $w_{m+1} = v_{m+1}$, we get

$$\sum_{i=1}^{m} \lambda_i w_i > \sum_{i=1}^{m} \lambda_i v_i$$

which, by weak m-utilitarianism, implies wP^*v and hence uP^*v.

A trickier case arises when $\lambda_j^0 > \lambda_j^1 = 0$ (hence, $\lambda_{m+1}^1 > 0$). In this case we take $\lambda = (0, 0, \cdots, 0, \lambda_{m+1}^1)$ in $E_+^{M \cup \{m+1\}}$. Indeed, for any u and v in E^N, with $u_h = v_h$, $h > m+1$, and $u_{m+1} > v_{m+1}$, there exist $\varepsilon > 0$ sufficiently small and w in E^N such that

$$\begin{aligned}
&\sum_{i=1}^{m} \varepsilon \lambda_i^0 u_i + \lambda_{m+1}^1 u_{m+1} > \sum_{i=1}^{m} \varepsilon \lambda_i^0 v_i + \lambda_{m+1}^1 v_{m+1} \\
&w_{m+1} = v_{m+1} \\
&\varepsilon \lambda_j^0 w_j + \lambda_{m+1}^1 w_{m+1} = \varepsilon \lambda_j^0 u_j + \lambda_{m+1}^1 u_{m+1} \\
&w_h = u_h, \ j \neq h \neq m+1.
\end{aligned}$$

Then, by weak 2-utilitarianism, uP^*w (since $u_{m+1} > w_{m+1}$); also,

$$\begin{aligned}
\sum_{i=1}^{m} \varepsilon \lambda_i^0 w_i + \lambda_{m+1}^1 w_{m+1} &= \sum_{i=1}^{m} \varepsilon \lambda_i^0 u_i + \lambda_{m+1}^1 u_{m+1} \\
&> \sum_{i=1}^{m} \varepsilon \lambda_i^0 v_i + \lambda_{m+1}^1 v_{m+1}
\end{aligned}$$

which implies $\sum_{i=1}^{m} \varepsilon \lambda_i^0 w_i > \sum_{i=1}^{m} \varepsilon \lambda_i^0 v_i$. Therefore, by weak m-utilitarianism, wP^*v and, hence, uP^*v. ∎

The following theorem is a weaker version of Blackwell and Girschik [1954], Theorem 4.3.1, (see also Theorem 2 by Roberts [1980b]).

Theorem 3.3.3 *A SWO R^* is weakly n-utilitarian whenever SP* and CU* are satisfied.*

Proof: Take any SWO R^* satisfying SP* and CU*. By Lemma 3.3.2 we only have to show that it is weakly 2-utilitarian. Consider the space $\tilde{E}^2 = \{u \in E^N : u_i = 0, 1 \neq i \neq 2\}$. By CU* it is clear that we only have to show that, for some $\lambda \in E^2_+$, if u and v in $\tilde{E}^2$ are such that $\lambda_1 u_1 + \lambda_2 u_2 > \lambda_1 v_1 + \lambda_2 v_2$ then uP^*v. We shall distinguish two cases.

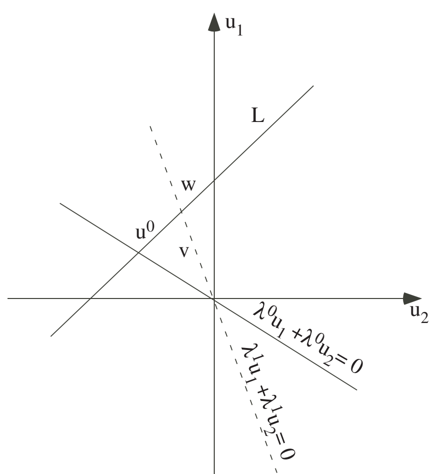
Case 1: There exist u^0 and v^0 in $\tilde{E}^2$ such that $u^0\, I^*\, v^0$ and $u^0 \neq v^0$. Suppose without loss of generality that $u_1^0 - v_1^0 > 0$. So, by SP*, $u_2^0 - v_2^0 < 0$ and, by CU*, $w^0\, I^*\, \mathbf{0}$, for $w^0 = (u^0 - v^0)$ and $\mathbf{0} = (0, \cdots, 0) \in \tilde{E}^2$. Take $\lambda^0 \in E^2_+$ such that $\lambda_1^0 w_1^0 + \lambda_2^0 w_2^0 = 0$. Then

$$\left(1, -\frac{\lambda_1^0}{\lambda_2^0}, 0, \cdots, 0\right) = \left(\frac{w_1^0}{w_1^0}, \frac{w_2^0}{w_1^0}, 0, \cdots, 0\right) I^*\mathbf{0},$$

using CU*. Also for any u in $\tilde{E}^2$ such that $u_1 > 0$, $u_2 \leq 0$, and $\lambda_1^0 u_1 + \lambda_2^0 u_2 = 0$, we must have $uI^*\mathbf{0}$. Otherwise we would have either $\mathbf{0}P^*(u_1, u_2, 0, \cdots, 0)$ or $(u_1, u_2, 0, \cdots, 0)P^*\mathbf{0}$. Since $(1, -\lambda_1^0/\lambda_2^0) = (u_1/u_1, u_2/u_1)$ this is equivalent, by CU*, of having $\mathbf{0}P^*(1, -\lambda_1^0/\lambda_2^0, 0, \cdots, 0)$ or $(1, -\lambda_2^0/\lambda_2^0, 0, \cdots, 0)P^*\mathbf{0}$, which gives us a contradiction. Moreover, for any u in $\tilde{E}^2$ such that $u_1 < 0, u_2 > 0$ and $\lambda_1^0 u_1 + \lambda_2^0 u_2 = 0$, we get $-uI^*\mathbf{0}$ and, by CU*, $\mathbf{0}I^*u$. Also, for any v in $\tilde{E}^2$, if $\lambda_1^0 v_1 + \lambda_2^0 v_2 > 0$ then there is u in $\tilde{E}^2$ such that $u_1 < v_1, u_2 < v_2$ and $\lambda_1^0 u_1 + \lambda_2^0 u_2 = 0$, so that $uI^*\mathbf{0}$ and, by SP*, $vP^*\mathbf{0}$. Similarly, for any v in $\tilde{E}^2$, if $\lambda_1^0 v_1 + \lambda_2^0 v_2 < 0$, $\mathbf{0}P^*v$. In other terms for any u and v in $\tilde{E}^2$ such that $\lambda_1^0 u_1 + \lambda_2^0 u_2 > \lambda_1^0 v_1 + \lambda_2^0 v_2$, we get $(u - v)P^*\mathbf{0}$ or, by CU*, uP^*v. In this first case we even have more than we need since if $\lambda_1^0 u_1 + \lambda_2^0 u_2 = \lambda_2^0 v_1 + \lambda_2^0 v_2$, then $(u - v)I^*\mathbf{0}$, or, by CU*, uI^*v. We get 2-utilitarianism.

Case 2: For any distinct, u and v in $\tilde{E}^2$ either uP^*v or vP^*u. Define in $\tilde{E}^2$ the line $L = \{u \in \tilde{E}^2 : u_1 - u_2 = 1\}$ and two subsets of this line: $I_1 = \{u \in L : uP^0\mathbf{0}\}$ and $I_2 = \{u \in L : \mathbf{0}P^*u\}$. Each of these subsets is a connected subset of L, since if u is in I_1 (resp. is in I_2) and v in L is such that $v_1 > u_1$ (resp. $v_1 < u_1$) then, by SP*, v must also belong to I_1 (resp. to I_2). Moreover, for any u in L, SP* implies that $uP^*\mathbf{0}$, whenever $u_1 \geq 0$ and $u_2 \geq 0$, and $\mathbf{0}P^*u$ whenever $u_1 \leq 0$ and $u_2 \leq 0$. Therefore there must be a point u^0 in L, $u_1^0 \geq 0$, $u_2^0 \leq 0$, such that, for all u in L: either $uP^*\mathbf{0}$ whenever $u_1 \geq u_1^0$ and $\mathbf{0}P^*u$ otherwise or $uP^*\mathbf{0}$ whenever $u_1 > u_1^0$ and $\mathbf{0}P^*u$ otherwise.

Accordingly define λ^0 in E^2_+ to be such that $\lambda^0_1 u^0_1 + \lambda^0_2 u^0_2 = 0$. Suppose first that $u^0 P^* \mathbf{0}$. Then, by arguments similar to Case 1, for any u in $\tilde{E}^2$ such that $u \neq 0$ and $\lambda^0_1 u_1 + \lambda^0_2 u_2 = 0$, we must have $uP^*\mathbf{0}$ whenever $u_1 \geq 0$ and $u_2 \leq 0$ (and $\mathbf{0}P^*u$ whenever $u_1 \leq 0$ and $u_2 \geq 0$). Similarly if we have $\mathbf{0}P^*u^0$, then for any u in $\tilde{E}^2$ such that $\lambda^0_1 u_1 + \lambda^0_2 u_2 = 0$ and $u \neq 0$, we must have $\mathbf{0}P^*u$ whenever $u_1 \geq 0$ and $u_2 \leq 0$ (and $uP^*\mathbf{0}$ whenever $u_1 \leq 0$ and $u_2 \geq 0$). Moreover, for any v in $\tilde{E}^2$ with $v_1 \geq 0$, $v_2 \leq 0$ and $\lambda^0_1 v_1 + \lambda^0_2 v_2 > 0$, there is λ^1 in E^2_+ and w in L such that $\lambda^1_1 v_1 + \lambda^1_2 v_2 = \lambda^1_1 w_1 + \lambda^1_2 w_2 = 0$ (see Figure 1) and, since $w_1 > u^0_1$, we must have $wP^*\mathbf{0}$. Hence by the same reasoning as above $vP^*\mathbf{0}$. Similarly, for any v in $\tilde{E}^2$ with $v_1 \geq 0$, $v_2 \leq 0$ and $\lambda^0_1 v_1 + \lambda^0_2 v_2 < 0$ we must have $\mathbf{0}P^*v$. All v in $\tilde{E}^2$ such that $v_1 \leq 0$ and $v_2 \geq 0$ can be treated in the same way since, by CU*, we may simply consider $-v$. Any other v in $\tilde{E}^2$, $v \neq 0$, is either such that $v_1 \geq 0$ and $v_2 \geq 0$ or such that $v_1 \leq 0$ and $v_2 \leq 0$; by SP* we get either $vP^*\mathbf{0}$ or $\mathbf{0}P^*v$. In conclusion we see that, in this second case, we obtain the following: for any u and v in $\tilde{E}^2$, if $\lambda^0_1 u_1 + \lambda^0_2 u_2 > \lambda^0_1 v_1 + \lambda^0_2 v_2$ then uP^*v. The SWO R^* satisfying SP* and CU* is weakly 2-utilitarian. ∎



The first version of Theorem 3.3.3, given by Blackwell and Girschick [1954], rests on a weaker Pareto condition and its proof is based on a supporting hyperplane theorem (also proved in their book).

However in the proof just given it is interesting to consider each of the two cases. The second case is the more particular. It occurs when R^* is a simple order on $\tilde{E}^2$. This simple order is completely described in the proof and can be of two types: one that privileges the first individual, along the lines determined by the coefficients λ_1^0, λ_2^0, and one that privileges the other individual along these lines.

The first case considered in the proof can be obtained under very reasonable conditions, since it only requires that there be two indifferent vectors in $\tilde{E}^2$. Such a condition is anonymity and we then obtain a characterization of pure utilitarianism.

Theorem 3.3.4 *A SWO R^* is pure utilitarian if and only if it satisfies SP*, CU* and A*.*

Proof: For any u^0 in $\tilde{E}^2$ such that $u_1^0 \neq u_2^0$, if $v^0 = (u_2^0, u_1^0, 0, \cdots, 0)$ then $v^0 I^* u^0$. Therefore in the proof of Theorem 3.3.3 we may consider only Case 1 and instead of using Lemma 3.3.2, use Lemma 3.3.1 and A^*. The converse is easy to verify. ∎

Another proof of this theorem can be based on an argument used by Milnor [1954] in the characterization of the Laplace criterion. However to characterize n-utilitarianism a much weaker condition than anonymity can be used such as the following.

Weak anonymity: For all i and j in N, there are u and v in E^N such that $u_i > v_i$, $u_j < v_j$, $u_h = v_h$, $i \neq h \neq j$, and uI^*v.

By exactly the same reasoning we get the following alternative to Theorem 3.3.4.

Theorem 3.3.5 *A SWO R^* is n-utilitarian if and only if it satisfies SP*, CU* and WA*.*

Note however that weak anonymity is still stronger than nondictatorship. We shall now return to nondictatorship and replace CU* by CN* (or ON*). Since all transformations allowed by CU* are allowed by CN* all the arguments in the proof of Theorem 3.3.3 still hold. This permits us to consider a particular aspect of Theorem 3.3.3. Indeed as such, weak utilitarianism does not satisfy ND*: it suffices to let $\lambda_i > 0$ for some i in N and $\lambda_j = 0$ for all $j \neq i$, which gives us dictatorship of individual i. However dictatorship of some individual is not sufficient to specify completely R^*. The following specification and the resulting theorem were suggested by Luce and Raiffa [1957] (as based on the result by Blackwell and Girschick).

Lexical individual dictatorship: There exists a permutation σ of N such that for any u and v in E^N,

$$uP^*v \quad \text{if and only if} \quad \begin{array}{ll} u_{\sigma(j)} > v_{\sigma(j)} & \text{for some } j \text{ in } N \\ u_{\sigma(i)} = v_{\sigma(i)} & \text{for some } i < j. \end{array}$$

As well stated by Gevers [1979], "this aggregation principle thus rests on an exogenously given hierarchy among individuals, which hinges only on their names, and social preference always endorses the strict preference of the individual who stands highest in the hierarchy."

Theorem 3.3.6 *A SWO R^* is lexical individual dictatorship if and only if SP* and CN* (or ON*) holds.*

Proof: Since CN* implies CU*, we may use Theorem 3.3.3; i.e., there exists λ^1 in E^N_+ such that, for all u and v in E^N,

$$\sum_{i\in N} \lambda^1_i u_1 > \sum_{i\in N} \lambda^1_i v_i \text{ implies } uP^*v \text{ and } \lambda^1_{i_1} > 0 \text{ for some } i_1 \text{ in } N.$$

Consider some $\bar{u}$ and $\bar{v}$ in E^N such that $\bar{u}_{i_1} > \bar{v}_{i_1}, \bar{u}_i < \bar{v}_i, i \neq i_1$, and $\sum_{i\in N} \lambda^1_i(\bar{u}_i - \bar{v}_i) > 0$ Hence $\bar{u}P^*\bar{v}$ and, by CN*, for every positive scalar c,

$$\lambda^1_{i_1}(\bar{u}_{i_1} - \bar{v}_{i_1}) \geq c \sum_{i\neq i_1} \lambda^1_i(\bar{v}_i - \bar{u}_i) > 0$$

which gives a contradiction unless $\lambda^1_i = 0$ for all $i \neq i_1$.

Now considering again the argument of the proof of Theorem 3.3.3 and Lemma 3.3.2, we may repeat the argument for the set of all u in E^N with u_{i_1} maintained fixed. Then we get $(n-1)$-weak-utilitarianism, and there exists λ^2 in $E^{N\setminus\{i_1\}}_+$ such that, for all u and v in E^N with $u_{i_1} = v_{i_1}$, $\sum_{i\neq i_1} \lambda^2_i u_i > \sum_{i\neq i_1} \lambda^2_i v_i$ implies uP^*v, $\lambda^2_{i_2} > 0$ for some i_2 in N, $i_2 \neq i_1$. Then by the same argument as above, $\lambda^2_i = 0$ for all $i, i_1 \neq i \neq i_2$. Repeating this procedure n times we construct a sequence $\lambda^1_{i_1}, \lambda^2_{i_2}, \cdots, \lambda^n_{i_n}$ of positive scalars such that $\{i_1, i_2, \cdots, i_n\} = N$ and such that, for any $j, 1 \leq j \leq n$, and any u and v in $E^{N \setminus \{i_1, i_2, \cdots, i_{j-1}\}}$ with $u_i = v_i$ for all i in $\{i_1, i_2, \cdots, i_{j-1}\}$ we get: $\lambda_{i_j} u_{i_j} > \lambda_{i_j} v_{i_j}$ implies uP^*v.

This gives lexical individual dictatorship. The reverse direction is left to the reader. ■

This theorem appears to imply a weaker version of Arrow's theorem, where WP* would be replaced by SP*. In Gevers [1979], Theorem 3.3.6 is proved as an implication of the following theorem.

Theorem 3.3.7 *If R^* satisfies SP* and ACC*, there exists a partition of N in s subsets $S_1, S_2, \cdots, S_s$ with strictly positive associated weights $\lambda^1, \lambda^2, \cdots, \lambda^s$, respectively in $E^{S_1}, E^{S_2}, \cdots$, E^{S_s} such that, for any v and w in $\{u \in E^N : i < j$ implies $u_i < u_j\}$, if, for some integer $r \leq s, v_i = w_i$ for all i in $S_p, p < r$, and*

$$\sum_{i \in S_r} \lambda_i^r (v_i - w_i) > 0$$

then uP^*w. Moreover, if R^* also satisfies OC, the partition of N consists only of singletons.*

Proof: See Gevers [1979].

The last statement in the theorem uses the condition OC*. In the following section we shall analyze the consequences of this invariance axiom more extensively.

3.4 A characterization of leximin and rank dictatorship

We shall now give an axiomatic presentation of leximin. In fact this presentation which, as above, will proceed in two steps – first leximin "from inch to ell" and, second, a characterization of leximin for two-person situations – leads to several of the numerous axiomatic derivations of leximin that have been considered. This is the approach adopted in Strasnick [1976], Sen [1977, 1979], and Hammond [1976a, 1979].

To confine the application of leximin to situations where only a small number of persons are nonindifferent seems to respond to an objection often raised. It consists of considering the case where an improvement (possibly enormous) in the welfare of a great number of persons should be rejected because of a deterioration (possibly almost imperceptible) of the worst off. However, this response appears to be insufficient as we shall see. First we introduce m-person leximin.

m-Person leximin: A SWO R^* is called m-leximin, $1 \leq m \leq n$, if, for every subset M of m individuals and all u and v in E^N with $u_h = v_h$ for every h not in M, uP^*v if and only if

$$\begin{aligned} &u_{i(k)} > v_{i(k)} \text{ for some } k \text{ such that } i(k) \text{ is in } M \\ &u_{i(h)} = v_{i(h)} \text{ for all } h < k. \end{aligned}$$

Lemma 3.4.1 *(Leximin from inch to ell): If a SWO is 2-leximin, then it is m-leximin for all $m, 1 \leq m \leq n$.*

Proof: First we may verify that 2-leximin implies the 2-gradient principle. Indeed take any u and v in E^N such that $u_h = v_h$, for $i \neq h \neq j$, and some i and j in N. Suppose either

$$u_i \geq v_i \text{ and } u_j \geq v_j$$

or

$$u_i \geq v_j \text{ and } u_j \geq v_i.$$

In both cases 2-leximin implies uR^*v and if, in addition, one inequality is strict, then uP^*v.

Thus, by Lemma 3.1.1, we know that A* and SP* hold and so 1-leximin is immediate. To derive m-leximin ($m \geq 3$), suppose a contrario that $(m-1)$-leximin, $(m-2)$-leximin, $\cdots$, 1-leximin hold but that m-leximin does not. Using A*, this means that we can find u and v in E^N such that

$$\begin{aligned} &u_h = v_h,\ m < h \leq n \\ &u_1 \leq u_2 \leq \cdots \leq u_m \\ &v_1 \leq v_2 \leq \cdots \leq v_m \\ &u_1 > v_1 \text{ and } u_h \neq v_h,\quad 1 < h \leq m \end{aligned}$$

and

$$vR^*u.$$

By SP*, there must be some k in $\{2, 3, \cdots, m\}$ such that $u_k < v_k$ (otherwise we would have a contradiction). Now if $v_2 = v_1$, then $u_2 > v_2$ and taking w in E^N such that $w_1 = v_1$ and

$w_h = u_h$, for $h \neq 1$, one gets uP^*w, by SP*, and wP^*v, by $(m-1)$-*leximin* (since $w_h = v_h$ for $h = 1,\ m+1, m+2, \cdots, n$ and $w_m \geq w_{m-1} \geq w_2 > v_2 \leq v_3 \leq \cdots \leq v_m$). Or, if $v_2 > v_1$ then we may construct w in E^N such that $w_k = u_k$, $v_1 < w_1 < \min\{v_2, u_1\}$ and $w_h = v_h$, $1 \neq h \neq k$. Then again we get uP^*w, but this time using $(m-1)$-leximin (since $w_h = u_h$ for $h = k,\ m+1, m+2, \cdots, n$ and $w_m \geq w_{m-1} \geq \cdots \geq w_{k+1} \geq w_{k-1} \geq \cdots \geq w_2 = v_2 > w_1 < u_1 \leq u_2 \leq \cdots \leq u_{k-1} \leq u_{k+1} \leq \cdots \leq u_m$) we obtain a contradiction, i.e., uP^*vR^*u. The result follows. ■

This lemma allows us to prove a theorem, which by Lemma 3.1.1, is equivalent to Theorem 5 in Hammond [1979] and which leads to an alternative proof of Theorem 7.2 in Hammond [1976a]. All these theorems are based on a strong equity condition concerning two-person situations and which Hammond introduced with the objective of generalizing the "weak equity axiom" introduced by Sen [1973] for the case of income distribution.

Two-person equity (Hammond's equity: HE*): If u and v in E^N are such that, for some i and j and all h in N, $i \neq h \neq j$,

$$v_i < u_i < u_j < v_j \text{ and } u_h = v_h$$

then uR^*v.

Theorem 3.4.1 *A SWO R^* is leximin if and only if it satisfies SP*, A*, and HE*.*

Proof: We know that leximin satisfied SP* and A* and it is easy to see that it satisfies HE*. So, by the preceding lemma, it remains to show that SP*, A*, and HE* imply 2-leximin. For that purpose take any u and v in E^N such that, for some i and j in N, $u_h = v_h$, $i \neq h \neq j$. By A*, we may suppose that $\{i, j\} = \{1, 2\}$ and that $u_1 \leq u_2$ and $v_1 \leq v_2$. Clearly, if $u_1 = v_1$, then uRv if and only if $u_2 \geq v_2$. So consider the case where $u_1 > v_1$. If $u_2 \geq v_2$ then uP^*v by SP*. Therefore it remains only the subcase where $v_2 > u_2 \geq u_1 > v_1$. For that take w in E^n such that

$$w_h = u_h = v_h,\ 1 \neq h \neq 2$$

and

$$v_2 > u_2 \geq u_1 > w_2 > w_1 > v_1.$$

By SP*, uP^*w and, by HE*, wR^*v (since $v_2 > w_2 > w_1 > v_1$). Hence uP^*v. The result follows by Lemma 3.4.1. ■

The characterization of leximin provided by this theorem does not rely explicitly on any invariance axiom. We shall now turn to this problem. As it is stressed by Rawls in his book, an advantage of the difference principle (and of leximin) is that it requires only an ordinal informational basis: "it suffices that the least favored person can be identified and his rational preference determined" (see Rawls [1971], p. 77). It is therefore natural to consider first co-ordinality. The other characterization we shall offer relies strongly on the following interesting lemma, based on both OC* and SE*.

Lemma 3.4.2 *(Equity-inequity lemma): Suppose a SWO R^* satisfies SP*, A*, SE*, and OC*. Then one and only one of the following two conditions arises.*

(i) If u and v in E^N are such that, for some i and j and all h in N, $i \neq h \neq j$,

$v_i < u_i < u_j < v_j$ *and* $u_h = v_h$

then uP^*v.

(ii) If u and v in E^N are such that, for some i and j and all h in N, $i \neq h \neq j$,

$v_i < u_i < u_j < v_j$ *and* $u_h = v_h$

then vP^*u.

Proof: Consider u^0, u^1, v^0 and v^1 in E^N such that

$$\text{for some } i \text{ and } j, \quad v_i^0 < u_i^0 < u_j^0 < v_j^0 \text{ and } u_h^0 = v_h^0, i \neq h \neq j,$$
$$\text{for some } k \text{ and } \ell, \quad v_k^1 < u_k^1 < u_\ell^1 < v_\ell^1 \text{ and } u_h^1 = v_h^1, k \neq h \neq \ell.$$

We want to showo first that $u^0R^*v^0$ if and only if $r^1R^*v^1$. Clearly, there exist u^2 and v^2 in E^N and a permutation σ of N such that

$$u_h^2 = u_{\sigma(h)}^1 \text{ and } v_h^2 = v_{\sigma(h)}^1, \text{ for all } h \in N,$$
$$\text{and both } \sigma(i) = k \text{ and } \sigma(j) = \ell.$$

Then, by A*, we know that $u^2 I^* u^1$ and $v^2 I^* v^1$. Now, using OC*, we may construct a strictly increasing function φ such that

$$u_i^0 = \varphi(u_i^2), \quad u_j^0 = \varphi(u_j^2), \quad v_i^0 = \varphi(v_i^2), \quad \text{and } v_j^0 = \varphi(v_j^2).$$

Taking u^3 and v^3 in E^N such that $u_h^3 = \varphi(u_h^2)$ and $v_h^3 = \varphi(u_h^3)$, for all h in N, we get $u^3 I^* u^2$ and $v^3 I^* v^2$, by OC*. Since, also, $u_i^3 = u_i^0, u_j^3 = u_j^0$, $v_i^3 = v_i^0, v_j^3 = v_j^0$, $u_h^0 = v_h^0$ and $u_h^3 = v_h^3$, $i \neq h \neq j$, we may apply SE* and get $u^0 R^* v^0$ if and only if $u^3 R^* v^3$, which is equivalent to $u^0 R^* v^0$ if and only if $u^1 R^* v^1$.

To complete the proof there remains to eliminate the possibility of indifference. Suppose, a contrario, that $u^0 I^* v^0$ and take u^1 and v^1 in E^N such that

$$v_i^0 < u_i^1 < u_i^0 \text{ and } u_h^1 = u_h^0, \text{all}\, h \neq i, \text{ but } v^1 \equiv v^0.$$

Then, by above, $u^1 I^* v^1$; that is, $u^1 I^* v^0$. Hence $u^0 I^* u^1$, which contradicts SP*. Therefore we must have $u^0 P^* v^0$ or $v^0 P^* u^0$ and the result follows. ■

It is interesting to remark the similarity of condition (i) in the lemma with HE*. In fact condition (i) is stronger than HE*, but also satisfied by leximin. So in Theorem 3.4.2 we could replace HE* by condition (i). Moreover condition (ii) appears as an "inequity condition" dual to condition (i): the better-off individual wins in all two-person situations of the kind described. Applying the reasoning of Theorem 3.4.2 in an obvious way we get a SWO that is dual to leximin and may be called *leximax*. It says that, if the best-off individual is nondifferent, then let him decide; but, if he is indifferent, then let the second best-off individual decide; and so on. We may take as a condition the negation of condition (ii), namely minimal equity.

Minimal equity (ME)*: For some u and v in E^N and i and j in N,

$$v_i < u_i < u_j < v_j, \quad u_h = v_h, \quad i \neq h \neq j, \quad \text{and } uR^*v.$$

Then we get the following results (from Lemma 3.4.3 and Theorem 3.4.2).

Theorem 3.4.2 *A SWO R^* satisfying SP*, A*, SE*, and OC* is either the leximin or the leximax.*

Theorem 3.4.3 *A SWO R^* is the leximin if and only if it satisfies SP*, A*, SE*, OC*, and ME*.*

This last theorem is another characterization of the leximin where HE* has been replaced by three conditions SE*, OC*, and ME*. It is now interesting to compare Theorem 3.4.3 to Theorem 3.3.4. Since SE* and ME* are clearly satisfied by pure utilitarianism we arrive at the disturbing conclusion that the difference between this SWO and leximin can be entirely explained by invariance axioms. So we are led to require even more discrimination than the one allowed by either CU* or OC*. This we do in the next section. However, before that, one may wonder what type of SWO would result from the three conditions SP*, A*, and OC*. First, in our welfarist approach it is clear that, once the anonymity condition A* has been introduced, individual names do not matter, as such, anymore to compare any two welfare vectors. Only the ranking position of the individual components (as given by the function $i(\cdot)$ defined in Section 3.1) are important. Focusing on ranks we see that the difference principle, the leximin and the leximax have a common property: they privileged one rank (the worst-off position or the best-off position). This allows us to extend the notion of dictatorship from individuals to ranks and to state the following theorem (see Gevers [1979] or Roberts [1977]).

Theorem 3.4.4 *(The rank-dictatorship theorem): If a SWO R^* satisfies SP*, A*, and OC*, then there exists an integer r in $\{1, 2, \cdots, n\}$, that is, a rank, such that, for any u and v in E^N,*

$$u_{i(r)} > v_{i(r)} \text{ implies } uP^*v.$$

This is in a co-ordinal, anonymous (and welfarist) framework an analogue to Arrow's theorem.

However leximin and leximax go further. They both define a hierarchy of the set of ranks, such that each rank becomes a "dictator" whenever all lower ranks in the hierarchy are indifferent. In other words we should extend lexical individual dictatorship to lexical rank dictatorship. As shown by Gevers [1979], this is implied by SP*, A*, and OC* only on the subset W of vectors in E^N such that no two individual components are equal (i.e., $u \in W$ if and only if $u_i \neq u_j$, all

i and j in N). Gevers [1979] provides a counterexample for welfare vectors in which there are ties.

Theorem 3.4.5 *(Lexical rank-dictatorship theorem): If a SWO R^* satisfies SP*, A*, and OC*, there exists a permutation σ of $\{1, 2, \cdots, n\}$ such that, for any u and v in W, uP^*v if and only if*

$$u_{i(\sigma(k))} > v_{i(\sigma(k))} \textit{ for some } k \textit{ in } \{1, 2, \cdots, n\}$$

and

$$u_{i(\sigma(h))} > v_{i(\sigma(h))} \textit{ for all } h \textit{ in } \{1, 2, \cdots, k-1\}.$$

Proof: The proof of this theorem is based on Lemma 3.1.1 and Theorem 3.3.7.

3.5 Other social welfare orderings and inequality measures

The purpose of this section is to investigate the class of SWOs that result from other types of invariance allowing for more discrimination than either CU* or OC*. Most of the results will be presented without proofs.

3.5.1 Joint characterization of utilitarianism and leximin

The first type of invariance that it seems natural to introduce now is co-cardinality: it allows both comparisons of individual welfare levels and comparisons of individual welfare gains. The first result we state in this respect is due to Roberts [1980b].

Theorem 3.5.1 *If a SWO R^* satisfies WP* and CC* then there exists a numerical function g, homogeneous of degree 1, such that for any u and v in E^N,*

$$\bar{u} + g(u - \bar{u}) > \bar{v} + g(v - \bar{v}) \textit{ implies } uP^*v$$

where

$$\bar{u} = \frac{1}{n}\sum_{i=1}^{n} u_i \textit{ and } \bar{v} = \frac{1}{n}\sum_{i=1}^{n} v_i.$$

Roberts [1980b] gives the following interesting example of such a numerical function g. Let

$$g(u) \stackrel{\text{def}}{=} \alpha \min_{i \in N} u_i, \quad u \in E^N, \quad 0 \leq \alpha \leq 1.$$

Then the SWO R^* is such that for any u and v in E^N

$$\bar{u} + \alpha \min_{i \in N}\{u_i - \bar{u}\} > \bar{v} + \alpha \min_{i \in N}\{v_i - \bar{v}\} \text{ implies } uP^*v.$$

This ordering is then weakly utilitarian for $\alpha = 0$ and satisfies the difference principle for $\alpha = 1$. When α is between 0 and 1 then R^* satisfies a combination of these two principles with respective weights $(1 - \alpha)$ and α.

However, the preceding result still gives a large class of possible SWOs. The following theorem due to Deschamps and Gevers [1978] introduces more conditions and restricts considerably the class of possible SWOs.

Theorem 3.5.2 *For $n \geq 3$, a SWO R^* satisfying SE*, SP*, A*, ME*, and CC* is either the leximin or weakly utilitarian.*

This theorem gives, a posteriori, some argument, in addition to the historical reasons, for our focusing on utilitarianism and leximin in our structural investigation of SWOs. It is clear also that the role played by minimal equity (ME*) is to discard leximax. Now if the other conditions are coupled with the condition of continuity (C*), then using Debreu's [1960] theorem on additive separability, Maskin [1978] shows that leximin and leximax are both discarded and the SWO must be pure utilitarianism.[15]

Theorem 3.5.3 *For $n \geq 3$, a SWO R^* satisfying SP*, A*, C*, and CC* is pure utilitarian.*

3.5.2 Global means and Kolm-Pollak functions

The two other types of invariance properties we shall investigate, and which introduce still more comparability than co-cardinality does, are ratio-scale comparability and difference comparabil-

[15] This argument may also be founded on Myerson [1978] approach based on either a linearity or a concavity condition. Even more different arguments are given in Yaari [1981] and Pazner and Schmeidler [1976].

ity. As stressed by Blackorby and Donaldson [1982], the crucial question resulting from ratio-scale comparability is the interpretation of the interpersonally significant welfare level taken as origin. Consequently, negative and positive utilities belong to two different categories separated by an interpersonally recognized norm, such as "poverty line". To avoid the difficulties involved by this natural origin, a way to proceed is to restrict utilities to the nonnegative orthant of E^N. For SWFLs this would amount to the introduction of some domain restriction. Such an approach is used by Roberts [1980b] and can be justified by taking the contractualist viewpoint adopted by the theory of bargaining. In this theory, the origin may be interpreted as a "status quo" or a "disagreement point" that obtains only if the negotiation breaks down. In this context SWOs, as restricted to the nonnegative orthant of E^N, may be viewed as "arbitration schemes".[16] The next chapter of this volume is devoted to the bargaining problem.

The first of the results that we shall quote and which are due to Blackorby and Donaldson [1982], does not, however, introduce any domain limitation.

Theorem 3.5.4 *For $n \geq 3$, a SWO R^* satisfying SP*, C*, SE*, and RC* is defined by a global mean of order $r > 0$. Namely, for any u and v in E^N, uR^*v, if and only if, for some $r > 0$, $G_r(u) \geq G_r(v)$ where, for any $w \in E^N$,*

$$\begin{aligned} G_r(w) &= \left[\sum_{i=1}^{n} \frac{\alpha(w_i)}{\alpha^+} |\beta_i(w_i) w_i|^r\right]^{1/r} \quad \textit{if} \quad \sum_{i=1}^{n} \alpha(w_i) |\beta_i(w_i) w_i|^r \geq 0 \\ &= -\left[\sum_{i=1}^{n} \frac{\alpha(w_i)}{\alpha^-} |\beta_i(w_i) w_i|^r\right]^{1/r} \quad \textit{if} \quad \sum_{i=1}^{n} \alpha(w_i) |\beta_i(w_i) w_i|^r \leq 0 \end{aligned}$$

with $\alpha^+ > 0$, $\alpha^- < 0$, $\beta_i(\cdot)$ positive and constant both for all $w_i \geq 0$ and all $w_i < 0$,

$$\begin{aligned} \alpha(w_i) &= \alpha^+ \quad \textit{if } w_i \geq 0, \\ &= \alpha^- \quad \textit{if } w_i < 0, \end{aligned}$$

[16] However, here interpersonal comparisons are introduced, which is not the case, for instance, in Nash's bargaining theory [1950]. Moreover, through the welfarist approach, we assume implicitly some kind of independence of irrelevant alternatives that is clearly violated in Nash's theory (e.g., see the discussion in Sen [1974a]). However, from another viewpoint, as well stressed by Pazner [1979], the extended sympathy framework can be presented (by Arrow [1951, 1963], p. 135) as one type of use of irrelevant alternatives ("irrelevant" because it is not feasible for someone to become somebody else).

and all these parameters chosen so that $G_r(1,1,\cdots,1) = -G_r(-1,-1,\cdots,-1) = 1$.

The second result simplifies a lot the foregoing description of the SWO involved by restricting its domain to $\bar{E}^N_+$ (the nonnegative orthant of E^N).

Theorem 3.5.5 *For* $n \geq 3$*, a SWO* R^* *with its domain restricted to* $\bar{E}^N_+$ *and satisfying there SP*, C*, SE*, and RC* is defined by a generalized mean of order* r*. Namely, for any* u *and* v *in* $\bar{E}^N_+$*,* uR^*v*, if and only if* $g_r(u) \geq g_r(v)$ *where, for any* $w \in \bar{E}^N_+$

$$\begin{aligned} g_r(w) &= \left[\sum_{i=1}^{n} \alpha_i w_i^r\right]^{1/r} && \text{if } r \neq 0 \\ &= \prod_{i=1}^{n} w_i^{\alpha_i} && \text{if } r = 0 \end{aligned}$$

with every $\alpha_i > 0$ *and* $\sum_{i=1}^n \alpha_i = 1$.

The proofs of these theorems are based on results on functional equations of Eichhorn [1978]. It is interesting to note that g_0 in the Theorem 3.5.5 coincides with the nonsymmetric Nash solution to the bargaining problem.

Turning now to difference comparability and using exponential transformations of the expressions in Theorem 3.5.5, the SWO involved becomes now what Blackorby and Donaldson [1980] have called a Kolm-Pollak function.[17] This is

Theorem 3.5.6 *For* $n \geq 3$*, a SWO* R^* *satisfying SP*, C*, SE*, and DC* is defined by a Kolm-Pollak function. Namely, for any* u *and* v *in* E^N*,* uR^*v*, if and only if* $K_r(u) \geq K_r(w)$*, where, for any* $w \in E^N$*,*

$$\begin{aligned} K_r(w) &= \frac{1}{r} \log \left[\sum_{i=1}^{n} \alpha_i e^{r w_i}\right] && r \neq 0 \\ &= \sum_{i=1}^{n} \alpha_i w_i && r = 0 \end{aligned}$$

with every $a_i > 0$ *and* $\sum_{i=1}^n \alpha_i = 0$.

[17] Blackorby and Donaldson ([1980a], p. 116) show that the reference-level free absolute index suggested by Kolm [1976a,b] coincides with the index derived from a social evaluation function that is additively separable and homothetic to minus infinity in the sense of Chipman and Pollak.

Letting every $\alpha_i = 1/n$, then we may compute that, for $r = 0$, K_r becomes pure utilitarianism and that for r approaching $-\infty$, K_r approaches the maximin function.[18]

An important application of these SWOs is the justification of various measures of economic inequality. The next subsection provides some remarks concerning the relationship between inequality measures and social welfare orderings.

3.5.3 Social welfare functions and inequality indices

The measurement of inequality, and more specifically the measurement of inequality of income distribution by a single index, has been the object of many economic studies. Early in this century several economists had already proposed various ways to evaluate the change in economic inequality resulting from changes in the distribution of incomes. The Lorenz curve, the Gini coefficient, as well as contributions by Pigou and by Dalton are well-known examples. (On this subject see the book by Sen [1973].) The problem in measuring inequality is that indices that look reasonable – for example the relative mean deviation, the variance, the coefficient of variation, or the relative mean difference – may give contradictory indications. It is therefore crucial to study the various properties and the ethical implications of the indices one wants to use. One way to achieve this objective is to relate each inequality measure to the social welfare function, which may be viewed as being implicit in this measure. This idea, which can be traced back to Dalton [1920], has recently been given precise formulation with the notion of an *equally distributed equivalent income* (see Kolm [1969], Atkinson [1970], and Sen [1973]). This notion can be easily presented in our welfarist framework and so we shall do, introducing simultaneously the distinction between relative and absolute inequality indices – or "rightist" and "leftist" indices to use Kolm's [1976] terminology. This distinction is analogous to the two different comparability conditions analyzed in the previous subsection.

[18] See Blackorby and Donaldson ([1980a], p. 117, fn. 9) and Atkinson [1970]. This is analogous to the result of Arrow ([1973], pp. 256-257), who presents it to show that the maximin criterion may appear as the limiting case of average utilitarianism. This argument is debated in Sen [1974a], where an analogous result is presented (under CC* though).

In our framework a *social welfare function* can be defined simply as real-valued function defined on the utility space E^N. A social welfare function can be used to define a SWO. The resulting SWO is then *represented* by the given social welfare function. We have done this several times in the preceding, using for instance the various utilitarian social welfare functions or, in the previous subsection, characterizing the social welfare functions defined as global (or generalized) means, on the one hand, and the Kolm-Pollak social welfare function, on the other.

As a first case take a continuous SWO R^* restricted to $\bar{E}_+^N$ and satisfying ratio-scale comparability. Then it can be represented by, and only by, a homothetic social welfare function W (see Theorem 3.5.5). Thus we may write, for any $u \in \bar{E}_+^N$,

$$W(u) = \varphi(\widetilde{W}(u))$$

where φ is an increasing transformation and $\widetilde{W}$ a social welfare function that is positively homogeneous of degree 1. The *equally distributed equivalent utility level* is a $w_u \in E$ such that

$$\begin{aligned} & W(w_u, w_u, \cdots, w_u) = W(u) \\ \text{or} \quad & \widetilde{W}(w_u, w_u, \cdots, w_u) = \widetilde{W}(u) \\ & w_u = \frac{\widetilde{W}(u)}{\widetilde{W}(\mathbf{1})} \text{with } \mathbf{1} \ = (1, 1, \cdots, 1) \in \bar{E}_+^N. \end{aligned}$$

Note that $\widetilde{W}$, and hence w, are both social welfare functions representing the same SWO as W. Now, letting $\bar{u} \stackrel{\text{def}}{=} (1/n) \sum u_i$, the social welfare function

$$I(u) \stackrel{\text{def}}{=} (\bar{u} - w_u)/\bar{u}$$

is called the *relative index of inequality corresponding* to W. It is a relative index since it is homogeneous of degree zero. Moreover, it is zero for any vector of equal utilities.[19]

Example 1. As an example take the Gini index of relative inequality

$$I_G(u) = \frac{1}{2n^2\bar{u}} \sum_{i=1}^{n} \sum_{j=1}^{n} |u_i - u_j|, \quad u \in \bar{E}_+^N$$

[19]Blackorby and Donaldson [1978] show how to generate a family of reasonable indices of relative inequality (continuous, homogeneous of degree zero, and S-concave) from any social welfare function that is continuous, increasing along rays and S-concave. Conversely, from any reasonable index of relative inequality one may generate at least one social welfare function with the foregoing properties.

or, as shown in Sen [1973], p. 31,

$$I_G(u) = 1 + \frac{1}{n} - \frac{2}{n^2\bar{u}}[n\,u_{i(1)} + (n-1)u_{i(2)} + \cdots + u_{i(n)}],\ u \in \bar{E}^N_+$$

where $i(\cdot)$ is defined in Section 3.1.

Then, as shown by Donaldson and Weymark [1980],

$$\widetilde{W}_G(u) = \sum_{k=1}^{n} [2(n-k)+1]u_{i(k)}, \quad u \in \bar{E}^N_+.$$

Since the coefficients in this function look arbitrary, they propose to consider a larger class of social welfare functions given by

$$W_\gamma(u) = \sum_{k=1}^{n} a_k\, u_{i(k)}, \quad u \in \bar{E}^N_+,$$

where $a_k > 0$, $k = 1, \cdots, n$, and $a_1 \geq a_2 \geq \cdots \geq a_n$. (This is axiomatized in Gevers [1979]).

Corresponding to this function, we get the following relative index

$$I_\gamma(u) = 1 - \frac{\sum_{k=1}^{n} a_k\, u_{i(k)}}{\bar{u}\sum_{k=1}^{n} a_k}, \quad u \in \overline{E}^N_+.$$

This is called the generalized Gini relative index.

As a second case, take now a continuous SWO R^* satisfying difference comparability. Then it can be represented by, and only by, a translatable social welfare function W (see Theorem 3.5.6). In other words we may write, for an $u \in E^N$

$$W(u) = \varphi(\widetilde{\widetilde{W}}(u))$$

where φ is an increasing transformation and $\widetilde{\widetilde{W}}$ is a social welfare function that is unit-translatable; for any scalar b

$$\widetilde{\widetilde{W}}(u + \mathbf{1}\,b) = \widetilde{\widetilde{W}}(u) + b,\ \text{with } \mathbf{1} = (1, 1, \cdots, 1) \in E^N.$$

Again we have for w_u the equally distributed equivalent utility level

$$\widetilde{\widetilde{W}}(\mathbf{1}\,w_u) = \widetilde{\widetilde{W}}(u), \quad u \in E^N$$

and, since $\widetilde{\widetilde{\widetilde{W}}}$ is unit-translatable, we get

$$w_u = \widetilde{\widetilde{\widetilde{W}}}(u) - \widetilde{\widetilde{\widetilde{W}}}(\mathbf{0}), \text{ with } \mathbf{0} = (0, 0, \cdots, 0) \in E^N.$$

We may now define an *absolute index of inequality corresponding to* W:

$$A(u) = \bar{u} - w_u, \quad u \in E^N.$$

It is an absolute index since it is invariant to any translation.[20]

Example 2. We can define the Gini index of absolute inequality by

$$A_G(u) = \bar{u} - \frac{1}{n^2} \sum_{k=1}^{n} [2(n-k)+1] u_{i(k)}, \quad u \in E^N.$$

Similarly we may write the generalized Gini absolute index:

$$A_\gamma(u) = \bar{u} - \frac{\sum_{k=1}^{n} a_k \, u_{i(k)}}{\sum_{k=1}^{n} a_k}, \quad u \in E^N,$$

where $a_k > 0$, $k = 1, \cdots, n$, and $a_1 \geq a_2 \geq \cdots \geq a_n$.

The example of the Gini indices was good to take since every W_γ is both homothetic and translatable.

4 Conclusion: two fundamental justifications

In this concluding section I would like to come back to the main difficulty, which is linked to the extended sympathy approach I have used as the basis for this presentation. As Elisha Pazner justly noted, "it seems natural to presume that there is a limit to one's ability to put oneself into somebody else' shoes. The implication of such a limit is that interpersonal orderings become a subjective matter and will generally differ for different individuals" ([1979], p. 163). How, then, can one justify, on ethical grounds, the use, at any level of social choice, of a single social welfare ordering? The theories of Harsanyi and Rawls may be viewed as providing two such justifications, albeit rather distinct ones.

[20] In Blackorby and Donaldson [1980a] it is shown that any well-behaved social welfare function can generate a family of absolute indices of inequality. Conversely for each absolute index there exists a family of social welfare functions that imply this index and that represent the same SWO whenever they are translatable.

4.1 Expected utility and the fundamental preference

Both theories may be seen as theories of justice based on a notion of "fairness". This notion is inseparable from the construction of some sort of procedure. The fairness of the result can only be secured by the reasonable character of the rules defining the procedure. Applications of such a constructivist approach are numerous in normative economics and in game theory. A well-known example is the procedure of fair division. The simplest case of it is when two individuals have to divide a cake. The fair solution is presented as the one resulting from the procedure consisting in asking the first individual to divide the cake into two parts, and the second individual to choose the part that pleases him most (i.e., the biggest). However, the most important application of this idea, from the present viewpoint, is the one developed by Lerner [1944], Vickrey [1945, 1960], and Harsanyi [1953, 1955, 1977]. Because the most systematic development of this application is due to Harsanyi, we shall concentrate on his procedure.

The objective of Harsanyi is to build up a "general theory of rational behavior" divided in two main parts. The first part is the theory of rational behavior of the individual, respectively under certainty, risk, and uncertainty. The second part concerns rational behavior in a social setting and is developed at two levels. At the first level, the game-theoretic one, each individual pursues his own self-interest and is moved according to personal preferences. At the second level, where the ethical norms constraining the first level have to be chosen, each individual pursues the interests of society as a whole and is moved according to social (or moral) preferences. The fairness procedure is introduced to determine these moral preferences. In this procedure each individual, who has to judge different possible situations for society, is to adopt an impartial view, and this he may achieve by acting *as if* "he simply did not know in advance what his own social position would be in each social situation" ([1977], p. 49). For that purpose, the moral preference of every individual should be based on complete information not only about the objective social situation of every other individual but also about the subjective attitudes characterizing their personal preferences. To understand this distinction, let us return to Harsanyi's example. Suppose a society consists of two individuals, and consider two possible social situations, one in which fish is the main item of everyone's diet and the other in which everyone's diet consists mainly of meat. Suppose in addition that the first individual has a mild personal preference for fish and the second a strong distaste for fish. The requirement is that if some individual, say the first, wants to order socially the two situations, he must take into account not only the objective diet characterizing them (meat or fish) but also his own personal taste and the other indivudual's taste. It seems then that, socially, he would prefer the meat diet. Thus, in Harsanyi's theory, the social preference is defined on the set of "extended" social alternatives that are alternatives of the kind: "being in social alternative x with the objective position and the subjective attitude of individual i". The fairness or impartiality of the moral preference of an individual comes from the fact that it is determined in a hypothetical situation where, for each social alternative, the individual supposes that he has an *an equal chance* of being in the objective position and of adopting the subjective attitude of any of the individuals. Then assuming that, in such an "original position", the individual's moral preference would satisfy the conditions imposed in the theory of rational behavior in the face of risk, and assuming that, whenever he adopts some other individual's subjective attitudes, he thereby adopts the personal preference of this other individual, Harsanyi is led to infer that the individual's moral preference can be represented by a von Neumann-Morgenstern utility function. This is simply, here, the average of the utilities representing the personal preferences of all individuals.

More precisely, consider a society of n individuals and a set X of social alternatives $a, b, c, \cdots$ etc. Then the hypothetical decision problem for the individual in the original position may be presented in the following table.

	1	2	3	$\cdots$	n
a	C_{1a}	C_{2a}	C_{3a}	$\cdots$	C_{na}
b	C_{1b}	C_{2b}	C_{3b}	$\cdots$	C_{nb}
c	C_{1c}	C_{2c}	C_{3c}	$\cdots$	C_{nc}
etc.					

This table summarizes a decision problem in the face of risk, where the decisions to be chosen are the social alternatives $a, b, c, \cdots$, and the "states of the world" are the individuals $1, 2, \cdots, n$. In addition each consequence C_{iz}, resulting from some decision z in state i, is a complete description of the objective situation and of the subjective attitude of individual i in the social situation z, say $C_{iz} = (x_i, p_i)$, where x_i and p_i are the respective vectors of "objective" and "subjective" characteristics of individual i. Finally the probability attached to every state of the world is simply $1/n$. Then using the axioms of decision theory under risk one gets that the chosen alternative should maximize on X:

$$\sum_{i=1}^{n} \frac{1}{n} u(C_{ix}) = \sum_{i=1}^{n} \frac{1}{n} u(x_i, p_i)$$

where u is some (cardinal utility) function defined on the set of consequences. By the assumption that to adopt some individual subjective attitude is to adopt his personal preferences (as they are embodied at the game-theoretic level of behavior) we may identify

$$u(x_i, p_i) = U_i(x)$$

where $U(\cdot)$ is individual i's personal utility. Therefore the objective of an individual in the original position becomes

$$\sum_{i=1}^{n} \frac{1}{n} U_i(x)$$

which, in the context of a fixed population, is simply pure utilitarianism.

Of course the crucial step, from our viewpoint, is the identification of $u(C_{ix})$ to $U_i(x)$ for every i, since this is where the interpersonal judgment is introduced. On this basis any individual who would put himself in the original position would end up with the same social preference. This step has been often criticized. First, each individual solving the problem in the original position may have what is called in decision theory a different "risk attitude" (see the detailed discussion in Pattanaik [1968]). Second, the utilities introduced are representations that are not unique; their unit and origins may be arbitrarily changed. As we have seen in the preceding text, some invariance condition must be used and justified. To such questions Harsanyi provides a general answer that has been called the "theory of fundamental preference."[21] In such a theory the subjective attitudes characterizing each individual (including his risk attitude) may be reduced

[21] For an analysis of different possible formal presentations of Harsanyi's theory, see Blackorby, Donaldson, and Weymark [1980b].

to the same list of parameters the value of which may vary from individual to individual. This is based on the presumption that

> the different individuals choice behavior and preferences are at least governed by the *same basic psychological laws.* For in this case each individual's preferences will be determined by the same general causal variables. Thus the differences we can observe between different people's preferences can be predicted, at least in principle, from differences in these causal variables, such as differences in their biological inheritance, in their past life histories, and in their current environmental conditions. (Harsanyi [1977], p. 58).

In the notation of this chapter this means that for a sufficient a priori specification of the subjective attitudes of the individuals, both the variables x_i and the variables p_i appearing in the function u are in some sense "objective". Hence, by taking as part of the decision problem, all those causal variables that explain individual differences, we may eventually obtain one single fundamental social preference that would be, in the words of Kolm [1972], a formal expression for the notion of human nature.

4.2 Contractualism and social unity

Although Rawls's theory of justice also relies on a notion of fairness resulting from the original position procedure, he insists that "the unity of society and the allegiance of its citizens to their common institutions rest not on their espousing one rational conception of the good but on an agreement as to what is just for free and equal persons with different and opposing conceptions of the good" ([1982], p. 160). Despite the universal character of the principles of justice, they should not reduce the incommensurability of the different individual conceptions of the good. In Kant's view this is the compatibility of the universal character of the categorical imperative ("Act always on such a maxim as thou canst at the same time will to be a universal law") with the principle of the autonomy of the will. The moral person cannot be governed by maxims, even if they can be taken as universal laws, since he would then be reduced to a simple mean.

He himself, as a reasonable being, should be the legislator. For Rawls, this universality and this autonomy should both be realized through the procedure of the original position. Universality is warranted by such a procedure since every individual should be represented in this hypothetical situation and because each representative individual should be "behind a veil of ignorance"; that is, he should not know his place in society, his class position, his endowment in natural assets and abilities, his psychological propensities, and even his own conception of the good. Autonomy is implied because the parties in the original position are supposed to be moved by "their highest-order interests" to promote the understanding and the realization of a conception of justice and of a conception of the good (whatever it is). However as representatives, they should only possess the "minimum adequate powers of moral personality" ([1980], p. 529). This should only be enough to make possible the deliberation procedure through which an agreement on the principles of justice may be reached.

Rawls, in defining the original position, not only specifies the (artificial) personality of the representative individuals, but also the subject of their deliberation, that is, the social alternatives. These are limited to be the various ways of providing and distributing the "primary goods" that is the goods that are under the control of the major social, political and economic institutions (as opposed to natural goods). More precisely Rawls gives the following list:

(a) First, the basic liberties as given by a list, for example: freedom of thought and liberty of conscience; freedom of association; and the freedom defined by the liberty and integrity of the person, as well as by the rule of law; and finally the political liberties;

(b) Second, freedom of movement and choice of occupation against a background of diverse opportunities;

(c) Third, powers and prerogatives of offices and positions of responsibility, particularly those in the main political and economic institutions;

(d) Fourth, income and wealth; and

(e) Finally, the social bases of self-respect. (Rawls [1982], p. 162)

To arrive at an agreement, the parties in their original deliberation are supposed to have a preference ordering on the set of alternatives which can be derived from this list of primary goods. This preference is represented by a utility function or more precisely by "an index of primary goods" and this index is supposed to be the same for everyone. However here this identical index is not to be founded on some understanding of the "basic psychological laws" but is part of the agreement in the original position. The identity of the index does not result from some conception of a "human nature" but is agreed upon by representatives of free and equal persons (having incommensurable conceptions of the goods). If this identity may be viewed as a form of fundamentalism, it is a contractual fundamentalism (as opposed to a natural fundamentalism). Here I should quote Rawls again:

> To clarify this contrast, we can write the function which represents interpersonal comparisons in questions of justice made by citizens in the well-ordered society of justice as fairness: $g = f(x_i, \bar{p})$. Here g is the index of primary goods (a real number), f is the function that determines the value of g for individual i, and x_i is the vector of primary goods held or enjoyed by individual i. The vector y, which in $w = u(x, y)$ includes entries for all features of the person which may affect satisfaction, is here replaced by a constant vector $\bar{p}$ which has entries only for the characteristics of free and equal moral persons presumed to be fully cooperating members of society over a complete life. This vector is constant since all citizens are taken to possess these features to the minimum sufficient degree. Thus the same function holds for all citizens and interpersonal comparisons are made accordingly. The difference between the functions f and u expresses the fact that in justice as fairness individuals' different final ends and desires, and their greater or less capacities for satisfaction, play no role in determining the justice of the basic structure. They do not enter into $\bar{p}$. (Rawls [1982], p. 178, fn. 21)

In the terminology of the preceding subsection, $\bar{p}$ would be a subvector of p, common to every individual i, which characterizes the basic common subjective attitude making social unity possible.

Once it is assumed that the parties in the original position will agree on a common index of primary goods, Rawls is led to argue in favor of an explicit rule, which is the negotiated solution to this particular form of the social choice problem and which has to be defined for every such index. This rule is described by a list of three principles – the principles of justice – which should be taken in hierarchical order. The first is a *principle of equal liberty*: the basic liberties should be distributed equally and the (equal) share of every individual should be the largest possible. The second principle in this hierarchy is a *principle of equal opportunity:* the offices and positions considered should be fairly and equally open to every individual. These first two principles take care, respectively, of the distribution of the two first categories of primary goods (see (a) and (b) above). The third and last principles in this hierarchy is the already mentioned *difference principle*, which is to be applied to all the remaining primary goods: the allocation of all these should be to the greatest benefit of the least advantaged.

This will conclude our sketchy presentation of Rawls's fundamental justification for social unity. In contrast to Harsanyi's natural fundamentalism, which leads to some form of pure utilitarianism, this contractual fundamentalism leads to some form of pure utilitarianism, this contractual fundamentalism leads to some rule of the maximin type. However Rawls's (non-axiomatic approach), by introducing a hierarchy of principles, takes explicitly into account the nature of the social alternatives which are defined in terms of primary goods. From the viewpoint of social choice theory, this may be taken as a strong indication of the need to develop axiomatics for welfare models that would essentially diverge from a universal application of welfarism.

References

Armstrong, W.E. Utility and the theory of welfare. *Oxford Economic Papers*, 3, 257–271, 1951.

Arrow, K.J. *Social Choice and Individual Values, 2nd ed.*. New Haven, Connecticut: Yale University Press, 1963 (1st ed. 1951).

Arrow, K.J. Some ordinalist utilitarian notes on Rawls's theory of justice. *Journal of Philosophy*, 70(9), 1973a.

Arrow, K.J. Rawls's principle of just saving. *The Swedish Journal of Economics*, 75, 323–335, 1973b.

Arrow, K.J. Extended sympathy and the possibility of social choice. *American Economic Review*, 67(1), 1977.

Atkinson, A.B. On the measurement of inequality. *Journal of Economic Theory*, 2, 244-263, 1970.

Barone, E. The ministry of production in the collectivist state. In F.A. von Hayek (ed.), *Collective Economic Planning*. London: Routledge, 1935.

Baumol, W.J. Community indifference. *Review of Economic Studies*, 14:44–48, 1946.

Bentham, J. *An Introduction to the Principle of Morals and Legislation*. Payne, 1789, Oxford: Clarendon Press, 1907.

Bergson, A. A reformulation of certain aspects of welfare economics. *Quarterly Journal of Economics*, 52, 310-334, 1938.

Blackorby, C. and D. Donaldson. Measures of relative equality and their meaning in terms of social welfare. *Journal of Economic Theory*, 18: 59–80, 1978.

Blackorby, C. and D. Donaldson. A theoretical treatment of indices of absolute inequality. *International Economic Review*, 21(1), 107–136, 1980.

Blackorby, C. and D. Donaldson. Ratio-scale and translation-scale full interpersonal comparability without domain restrictions: admissible socia-evaluation functions. *International Economic Review*, 23(2), 249–268, 1982.

Blackorby, C. , Donaldson, D. and J.A. Weymark. On John Harsanyi's defences of utilitarianism. CORE Discussion Paper 8013, Université catholique de Louvain, 1980.

Blackorby, C., Donaldson, D. and J. Weymark. Social choice with interpersonal utility comparisons: a diagrammatic introduction. *International Economic Review*, 25(2): 327–356, 1984.

Blackwell, D. and M.A. Girshick. *Theory of Games and Statistical Decisions*. New York: Wiley, 1954.

Blau, J.H. Neutrality, monotonicity and the right of veto: A comment. *Econometrica*, 44, 603, 1976.

Border, K.C. Social welfare functions for economic environments with and without the Pareto principle. *Journal of Economic Theory*, 29, 205-216, 1983.

Bordes, G. Individualisme, ordinalisme et bien-être social. Discussion Paper of the Laboratoire d'analyse et de recherche économiques, Faculté des Sciences économiques, Université de Bordeaux I, 1980.

Dahl, R.A. *A Preface to Democratic Theory*. Chicago: University of Chicago Press, 1956.

Dalton, H. The measurement of the inequality of incomes. *Economic Journal*, 30, 1920.

Dasgupta, P. On some problems arising from Professor Rawls' conception of distributive justice. *Theory and Decision*, 4, 225–344, 1974.

d'Aspremont, C. and L. Gevers. Equity and the informational basis of collective choice. *Review of Economic Studies*, 44, 199–209, 1977.

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

Debreu, G. Topological methods in cardinal utility theory. In K.J. Arrow, S. Karlin and P. Suppes (eds.), *Mathematical Methods in the Social Sciences.* Stanford, CA: Stanford University Press.

Deschamps, R. and L. Gevers. Separability, risk-bearing and social welfare judgments. *European Economic Review*, 10, 77–94, 1977.

Deschamps, R. and L. Gevers. Leximin and utilitarian rules: a joint characterization. *Journal of Economic Theory*, 17, 143–163, 1978.

Donaldson, D. and J.A. Weymark. A single-parameter generalization of the Gini indices of inequality. *Journal of Economic Theory*, 22(1), 67–86, 1980.

Edgeworth, F.H. *Mathematical Psychics.* London: Kegan Paul: 1881, New York: A.M. Kelly, 1967.

Eichhorn, W. *Functional Equations in Economics.* Reading (Mass.): Addison-Wesley, 1978.

Farrell, M.J. Liberalism in the theory of social choice. *Review of Economic Studies*, 43, 3–10, 1976.

Fisher, I. A statistical method for measuring marginal utility and testing the justice of a progressive income tax. *Economic Essays in Honor of J.B. Clark.* New York: Macmillan, 1927.

Fleming, M. A cardinal concept of welfare. *The Quarterly Journal of Economics*, 66, 366–384, 1952.

Frisch, R. *New Methods of Measuring Marginal Utility.* Tubingen: J.C.B. Morh, 1932.

Gevers, L. On interpersonal comparability and social welfare orderings. *Econometrica*, 47, 75–89, 1979.

Gibbard, A. A Pareto-consistent Libertarian claim. *Journal of Economic Theory*, 7, 388–410, 1974.

Goodman, L.A. and H. Markovitz. Social welfare function based on individual rankings. *American Journal of Sociology*, 58, 257–262, 1952.

Graaff, J. de V. *Theoretical Welfare Economics.* Cambridge: Cambridge University Press, 1957.

Guha, A.S. Neutrality, monotonicity and the right of veto. *Econometrica*, 40, 821–826, 1972.

Guilbaud, G.-Th. Les théories de l'intérêt général et le problème logique de l'agrégation. *Economie Appliquée*, 5, 1952.

Hammond, P.J. Equity, Arrow's conditions and Rawls' difference principle. *Econometrica*, 44, 793–804, 1976a.

Hammond, P.J. Why ethical measures of inequality need interpersonal comparisons. *Theory and Decision*, 7, 263–274, 1976b.

Hammond, P.J. Equity in two person situations: some consequences. *Econometrica*, 47, 1127–1135, 1979.

Harsanyi, J.C. Cardinal utility in welfare economics and in the theory of risk-taking. *Journal of Political Economy*, 61, 434–435, 1953.

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

Harsanyi, J.C. *Essays on Ethics, Social Behavior and Scientific Explanation.* Dordrecht, Holland: D. Reidel, 1976.

Harsanyi, J.C. *Rational Behavior and Bargaining Equilibrium in Games and Social Situations.* Cambridge: Cambridge University Press, 1977.

Hicks, J.R. Foundations of welfare economics. *Economic Journal*, 49, 696–712, 1939.

Hotelling, H. The general welfare in relation to problems of taxation and of railway and utility rates. *Econometrica*, 6, 242–269, 1938.

Hutcheson, F. *An Inquity Concerning Moral Good and Evil*, 1725.

Kahneman, D. and A. Tversky. Prospect theory: an analysis of decision under risk. *Econometrica*, 47(2), 1979.

Kalai, E., Muller, E. and M. Satterthwaite. Social welfare functions when preferences are convex, strictly monotonic and continuous. *Public Choice*, 34, 87–97, 1979.

Kaldor, N. Welfare proposition in Economics. *Economic Journal*, 49, 1939.

Kelly, J.S. Rights exercising and a Pareto-consistent Libertarian claim. *Journal of Economic Theory*, 13, 138–153, 1976.

Kemp, M.C. and Y.K Ng. On the existence of social welfare functions, social orderings, and social decision functions. *Economica*, 43, 59–66, 1976.

Kemp, M.C. and Y.K Ng. More on social welfare functions: the incompatibility of individualism with ordinalism. *Economica*, 44, 89–90, 1977.

Kolm, S.C. The optimum production of social justice. In J. Margolis and H. Guitton (eds.), *Public Economics.* London: Macmillan, 1969.

Kolm, S.C. *Justice et équité.* Paris: CNRS, 1972.

Kolm, S.C. Unequal inequalities 'I' and 'II'. *Journal of Economic Theory*, 12, 416–442 and 13, 82–111, 1976a, b.

Lerner, A.P. The concept of monopoly and the measure of monopoly power. *Review of Economic Studies*, 1, 157–175, 1934a.

Lerner, A.P. Economic theory and socialist economy. *Review of Economic Studies*, 2, 51–61, 1934b.

Lerner, A.P. *The Economics of Control*. New York: Macmillan, 1944.

Little, I.M.D. Social choice and individual values. *Journal of Political Economy*, 60, 422–432.

Little, I.M.D. *A Critique of Welfare Economics*. Oxford: Clarendon Press, 1957 (1st ed. 1950).

Luce, R.D. and H. Raiffa. *Games and Decisions*. New York: Wiley, 1957.

Maskin, E. A theorem on utilitarianism. *Review of Economic Studies*, 45, 93–96, 1978.

May, K.O. A set of independent necessary and sufficient conditions for simple majority decision. *Econometrica*, 20, 680–684, 1952.

Milnor, J. Games against nature. In R.M. Thrall, C.H. Coombs and R.L. Davis (eds.), *Decision Processes*. New York: John Wiley, 1954.

Mirrlees, J.A. The economic uses of utilitarianism. In A. Sen and B. Williams (eds.), *Utilitarianism and Beyond*. Cambridge: Cambridge University Press; Paris: La Maison des Sciences de l'Homme, 1982.

Moulin, H. Choix social cardinal: résultats récents. Discussion Paper, Laboratoire d'Econométrie de l'Ecole Polytechnique, Paris, 1982.

Myerson, R.B. Linearity, concavity and scale invariance in social choice functions. Discussion Paper no. 321, Graduate School of Management, Northwestern University, Evanston (Ill), 1978.

Nash, J.F. The bargaining problem. *Econometricia*, 18, 155-162, 1950.

Ng, Y.K. Bentham or Bergson? Finite sensibility, utility functions and social welfare functions. *Review of Economic Studies*, 42(4), 545–569, 1975.

Nozick, R. *Anarchy, State and Utopia.* Oxford: Blackwell, 1974.

Pareto, V. *Manuel d'Economie Politique*, Paris: Giard, 1909 (1st ed. 1897).

Parks, R.P. An impossibility theorem for fixed preferences: a dictatorial Bergson-Samuelson welfare function. *Review of Economic Studies*, 43, 447–450, 1976.

Pattanaik, P.K. Risk, impersonality, and the social welfare function. *Journal of Political Economy*, 76: 1152–1169, 1968.

Pazner, E.A. Equity, nonfeasible alternatives and social choice: a reconsideration of the concept of social welfare. In J.J. Laffont (ed.), *Aggregation and Revelation of Preferences.* Amsterdam: North Holland, 1979.

Pazner, E.A. and D. Schmeidler. Social contract theory and ordinal distributive equity. *Journal of Public Economics*, 5, 261–268, 1976.

Pigou, A.C. *The Economics of Welfare.* London: Macmillan, 1920 (last reprint, 1962).

Pollak, R.A. Bergson-Samuelson social welfare functions and the theory of social choice. *The Quarterly Journal of Economics*, 73–90, 1979.

Rawls, J. *A Theory of Justice.* Oxford: Clarendon Press, 1972.

Rawls, J. Kantian constructivism in moral theory. Rational and full autonomy. *The Journal of Philosophy*, 77(9): 515–572, 1980.

Rawls, J. Social unity and primary goods. In A. Sen and B. Williams (eds.), *Utilitarianism and Beyond*, Cambridge: Cambridge University Press, 1982. Paris: La Maison des Sciences de l'Homme.

Robbins, L. *An Essay on the Nature and Significance of Economic Science*. London: Macmillan, 1932.

Roberts, K.W.S. Welfare theoretic social choice. Unpublished Ph.D. Thesis, Oxford University, 1977.

Roberts, K.W.S. Possibility theorems with interpersonally comparable welfare levels. *Review of Economic Studies*, 47, 409–420, 1980a.

Roberts, K.W.S. Interpersonal comparability and social choice theory. *Review of Economic Studies*, 47, 421–439, 1980b.

Roberts, K.W.S. Social choice theory: the single-profile and multi-profile approaches. *Review of Economic Studies*, 47, 441–450, 1980c.

Rothenberg, J. *The Measurement of Social Welfare*. Englewood Cliffs (NJ): Prentice-Hall, 1961.

Rubinstein, A. The single profile analogues to multi profile theorems: Mathematical logic's approach. Discussion Paper of the International Centre for Economics and Related Disciplines. London School of Economics, 1979.

Samuelson, P.A. *Foundations of Economic Analysis*. Cambridge (MA): Harvard University Press, 1947.

Samuelson, P.A. Arrow's mathematical politics. In S. Hook, *Human Values and Economic Policy: A symposium*, New York: New York University Press, 1967.

Samuelson, P.A. Reaffirming the existence of 'reasonable' Bergson-Samuelson social welfare functions. *Economica*, 44, 81–88, 1977.

Scitovsky, T. A note on welfare propositions in economics. *Review of Economic Studies*, 9, 77–88, 1941.

Scitovsky, T. A reconsideration of the theory of tariffs. *Review of Economic Studies*, 9, 89–110, 1942.

Seidl, C. On liberal values. *Zeitschrift für Nationalökonomie*, 35, 257–292, 1975.

Sen, A.K. *Collective Choice and Social Welfare.* San Francisco: Holden-Day, 1970.

Sen, A.K. *On Economic Inequality.* Oxford: Clarendon Press; and New York: Norton, 1973.

Sen, A.K. Informational bases of alternative welfare approaches: aggregation and income distribution. *Journal of Public Economics*, 3, 387–403, 1974a.

Sen, A.K. Rawls versus Bentham: an axiomatic examination of the pure distribution problem. *Theory and Decision*, 300–309, 1974b.

Sen, A.K. Liberty, unanimity and rights. *Economica*, 43, 217–245, 1976a.

Sen, A.K. Welfare inequalities and Rawlsian axiomatics. *Theory and Decision*, 7, 243–262, 1976b.

Sen, A.K. On weights and measures: informational constraints in social welfare analysis. *Econometrica*, 45, 1539–1572, 1977.

Sen, A.K. Utilitarianism and welfarism. *The Journal of Philosophy*, 76(9), 463–489, 1979a.

Sen, A.K. Social chice theory. In K.J. Arrow and M. Intriligator (eds.), *Handbook of Mathematical Economics*, Amsterdam: North-Holland, 1979b.

Sen, A.K. Personal utilities and public judgments: or what's wrong with welfare economics? *The Economic Journal*, 89-537–558, 1979c.

Sen, A.K. and B. Williams (eds.) *Utilitarianism and Beyond.* Cambridge: Cambridge University Press. Paris: La Maison des Sciences de l'Homme, 1982.

Sidgwick, H. *The Methods of Ethics.* London: Macmillan, 1907.

Simon, J.L. Interpersonal welfare comparisons can be made – and used for redistribution decisions. *Kyklos*, 27, 63–99, 1974.

Solow, R.M. Intergenerational equity and exhaustible resources. *Review of Economic Studies*, Symposium, 1974.

Stigler, G.J. A note on the new welfare economics. *American Economic Review*, 30, 1943.

Strasnick, S.L. Preference priority and the maximization of social welfare. Unpublished Ph.D. Dissertation. Harvard University, Cambridge, Mass., 1975.

Strasnick, S.L. Social choice and the derivation of Rawls's difference principle. *The Journal of Philosophy*, 73(4), 85–99, 1976a.

Strasnick, S.L. The problem of social choice: arrow to Rawls. *Philosophy and Public Affairs*, 5, 241–273, 1976b.

Suzumura, K. Liberal paradox and the voluntary exchange of rights exercising. *Journal of Economic Theory*, 22, 407–422, 1980.

Suppes, P. Some formal models of grading principles. *Synthese*, 6: 284–306, 1966.

Vickrey, W. Measuring marginal utility by reactions to risk. *Econometrica*, 13, 319–333, 1945.

Vickrey, W. Utility, strategy, and social decision rules. *The Quarterly Journal of Economics*, 74(4), 507–535, 1960.

von Neumann, J. and O. Morgenstern. *Theory of Games and Economic Behavior.* Princeton, N.J.: Princeton University Press, 1947 (1st ed. 1944).

Wilson, R. Social choice theory without the Pareto principle. *Journal of Economic Theory*, 5, 478–486, 1972.

Yaari, M.E. Rawls, Edgeworth, Shapley, Nash: theories of distributive justice re-examined. *Journal of Economic Theory*, 24, 1–39, 1981.

Young, H.P. An axiomatization of Borda's rule. *Journal of Economic Theory*, 9, 52–53, 1974.