

INTERPERSONAL COMPARISONS OF WELL-BEING

Charles Blackorby

Walter Bossert

No 711

WARWICK ECONOMIC RESEARCH PAPERS

DEPARTMENT OF ECONOMICS

THE UNIVERSITY OF
WARWICK

Interpersonal comparisons of well-being*

Charles Blackorby and Walter Bossert

April 2004

Prepared as Chapter V.3 of the *Oxford Handbook of Political Economy*,
B. Weingast and D. Wittman, eds., Oxford University Press, Oxford

Charles Blackorby: Department of Economics, University of Warwick and GREQAM,
c.blackorby@warwick.ac.uk

Walter Bossert: Département de Sciences Economiques and CIREQ, Université de Montréal,
walter.bossert@umontreal.ca

Abstract. This paper, which is to be published as a chapter in the Oxford Handbook of Political Economy, provides an introduction to social-choice theory with interpersonal comparisons of well-being. We argue that the most promising route of escape from the negative conclusion of Arrow's theorem is to use a richer informational environment than ordinal measurability and the absence of interpersonal comparability of well-being. We discuss welfarist social evaluation (which requires that the levels of individual well-being in two alternatives are the only determinants of their social ranking) and present characterizations of some important social-evaluation orderings. *Journal of Economic Literature* Classification Number: D63.

Keywords: Arrow's theorem, social choice with interpersonal utility comparisons, welfarism.

* Financial support through a grant from the Social Sciences and Humanities Research Council of Canada is gratefully acknowledged.

April 19, 2004

1. Introduction

This chapter provides a brief survey of the use of interpersonal comparisons in social evaluation. We focus on principles for social evaluation that are welfarist (Sen, 1979): principles that use information about individual well-being to rank alternatives, disregarding all other information. Utility functions are indicators of individual well-being and we use the terms utility and well-being synonymously. Sentient non-human animals have experiences and it is possible to take account of their interests in social evaluation. We focus on human beings in this chapter and refer the interested reader to Blackorby and Donaldson (1992) for a discussion of the ethics of animal exploitation.

We begin with the idea that the society has a large number of options from which to choose; to do so it wants to be able rank them. We call these options alternatives; an alternative is a complete description of everything that matters to society. This social ranking is to be determined by a social evaluation functional. Of course, the individual members of this society can also rank these alternatives; in fact, we assume that each individual has a utility function that represents his or her preferences over all alternatives. A list of utility functions, one for each individual is called a utility profile. A social-evaluation functional associates a social ordering of the alternatives with every utility profile in its domain. Welfarism obtains if and only if there exists a social-evaluation ordering of vectors of individual utilities that can be used, together with the information about well-being in a profile, to rank the alternatives.

Welfarist principles regard values such as individual liberty and autonomy as instrumental: valuable because of their contribution to well-being. Because of this, it is important to employ a comprehensive notion of well-being such as that of Griffin (1986). Individuals who are autonomous and fully informed may have self-regarding preferences that accord with their well-being, but we do not assume that they do. If they do, the individual utility functions are representations of these preferences.

Welfarism rests mainly on the view that any two alternatives in which everyone is equally well off are equally good, a condition that is called Pareto indifference. The Pareto-indifference axiom is implied by a condition proposed by Goodin (1991). Goodin suggests that if one alternative is declared socially better than another, then, the former should be better than the latter for at least one member of society. This is a fundamental property of a principle for social evaluation and we consider it a strong argument in favor of welfarism. Without this requirement, we run the risk of recommending social changes that are empty gestures, benefiting no one and, perhaps, harming some or all.

Because Pareto indifference applies separately to each utility profile, we need a way to require a principle for social evaluation to behave consistently across different profiles. Such a condition is provided by the axiom binary independence of irrelevant alternatives. We say that two profiles coincide on a pair of alternatives, x and y , if every individual

who finds x to be at least as good as y in the first profile also ranks them in the same way in the second profile and if every individual who ranks y at least as good as x in the first profile also does so in the second profile. If two profiles coincide on a pair of alternatives, the independence axiom requires the social ranking of that pair to be the same for both profiles.

The most commonly employed domain of a social-evaluation functional consists of all logically possible utility profiles. On this unlimited domain, a social-evaluation functional satisfies Pareto indifference and binary independence of irrelevant alternatives if and only if it is welfarist.

Arrow's (1951, 1963) fundamental theorem states that there do not exist satisfactory welfarist principles if the only information that can be used in social evaluation is ordinally measurable and interpersonally non-comparable utility information. Sen (1970) shows that the conclusion of Arrow's theorem remains true if Arrow's ordinal interpretation of individual utility is replaced by a cardinal interpretation and no interpersonal comparisons of well-being are permitted. Taking these results as our starting point, we illustrate how the Arrow-Sen impossibility can be avoided if various forms of interpersonal utility comparisons are possible.

Information-invariance conditions require that the social-evaluation principle respect the informational environment regarding the measurability and interpersonal comparability of well-being. As is standard in the literature, we express information assumptions by specifying the transformations that can be applied to utility profiles without changing their informational contents. If two utility vectors u and v are subjected to a vector of admissible transformations under a given informational environment, information invariance with respect to that environment demands that the social ranking of the transformed vectors is the same as that of u and v .

We review some of the most important characterization results for welfarist social-evaluation principles. Due to space limitations, we cannot provide an exhaustive survey but we attempt to mention the most relevant references for further reading. For the same reason, we do not provide any proofs but refer the interested reader to the original contributions or more extensive surveys.

Section 2 introduces our basic notation along with a formal definition of social-evaluation functionals. In addition, we present the welfarism theorem which shows that welfarism is a consequence of three fundamental axioms. Because welfarism permits us to work with a single ordering of utility vectors (called a social-evaluation ordering), this ordering is employed instead of the social-evaluation functional in the remainder of the chapter. In Section 3, we formulate some basic axioms for social-evaluation orderings and define the orderings that are of particular importance in this chapter. They are the strongly dictatorial orderings, the strong positional dictatorships, the utilitarian principle,

the global means of order r , the Kolm-Pollak class and the leximin ordering. Information-invariance properties are introduced in Section 4, and Section 5 contains an overview of some important results. Section 6 concludes the chapter with a discussion of possible extensions! and applications of our model in choice problems.

2. Welfarist social evaluation

We use \mathcal{R} (\mathcal{R}_{++}) to denote the set of all (positive) real numbers, and \mathcal{Z}_{++} is the set of all positive integers. Consider a set of alternatives X with at least three elements and a society $N = \{1, \dots, n\}$ of $n \in \mathcal{Z}_{++}$ individuals where n is finite and greater than one. A binary relation R on X is (i) reflexive if and only if xRx for all $x \in X$; (ii) transitive if and only if xRy and yRz together imply xRz for all $x, y, z \in X$; (iii) complete if and only if xRy or yRx for all distinct $x, y \in X$. A utility profile is an n -tuple $U = (U_1, \dots, U_n)$ of individual utility functions $U_i: X \rightarrow \mathcal{R}$, one for each individual $i \in N$.

A social-evaluation functional assigns a social ranking of alternatives to each profile of utility functions in its domain. Let \mathcal{U} denote the set of all logically possible utility profiles. The domain of a social-evaluation functional is a non-empty subset \mathcal{D} of \mathcal{U} . Denoting the set of all logically possible orderings on X by \mathcal{O} , a social-evaluation functional is a mapping $F: \mathcal{D} \rightarrow \mathcal{O}$. The social ranking of the alternatives obtained for a profile $U \in \mathcal{D}$ is $R_U = F(U)$. The subscript U on the induced ordering indicates that the social ordering of alternatives may depend upon the preferences of the individual members of society. P_U and I_U are the asymmetric and symmetric parts of R_U , that is, for all $x, y \in X$ and for all $U \in \mathcal{D}$, (i) xP_Uy if and only if xR_Uy and not yR_Ux ; (ii) xI_Uy if and only if xR_Uy and yR_Ux .

The social-evaluation functional F is welfarist if and only if there exists a social-evaluation ordering R (referred to as a social welfare ordering by Gevers, 1979) on the set \mathcal{R}^n of all n -dimensional utility vectors such that, for any utility profile $U \in \mathcal{D}$ and for any two alternatives $x, y \in X$, x is socially at least as good as y for the profile U if and only if the utility vector $u = (u_1, \dots, u_n) = (U_1(x), \dots, U_n(x)) = U(x)$ is at least as good as the utility vector $v = (v_1, \dots, v_n) = (U_1(y), \dots, U_n(y)) = U(y)$ according to R , that is, if and only if uRv . We use P and I to denote the asymmetric and symmetric parts of R . Welfarism is a consequence of three axioms: unlimited domain, Pareto indifference and binary independence of irrelevant alternatives.

Unlimited domain requires the social-evaluation functional to produce a social ordering for every logically possible utility profile.

Unlimited domain: $\mathcal{D} = \mathcal{U}$.

Pareto indifference demands that if, according to a utility profile U in the domain of the social-evaluation functional, the individual utilities for two alternatives x and y are the same, then x and y must be equally good according to the social ranking generated by U . This axiom is implied by a condition introduced by Goodin (1991). Goodin suggests that if an alternative x is socially better than an alternative y according to the ranking obtained for a profile U , then it must be the case that, in the given profile U , x is better than y for at least one individual—otherwise we run the risk of making empty gestures that benefit no-one and may make some worse off.

Pareto indifference: For all $x, y \in X$ and for all $U \in \mathcal{D}$, if $U(x) = U(y)$, then $xI_U y$.

Binary independence of irrelevant alternatives is a consistency condition that imposes restrictions across different profiles. If the utilities for two alternatives x and y are the same in two profiles U and V , then the social rankings of x and y resulting from the two profiles should be the same.

Binary independence of irrelevant alternatives: For all $x, y \in X$ and for all $U, V \in \mathcal{D}$, if $U(x) = V(x)$ and $U(y) = V(y)$, then

$$xR_U y \Leftrightarrow xR_V y.$$

For a social-evaluation functional that satisfies unlimited domain, Pareto indifference and binary independence of irrelevant alternatives together are equivalent to welfarism. This result, which is implicit in d'Aspremont and Gevers (1977) and explicit in Hammond (1979), is referred to as the welfarism theorem. It requires the assumption that there are at least three alternatives in X .

Theorem 1: *Suppose F satisfies unlimited domain. F satisfies Pareto indifference and binary independence of irrelevant alternatives if and only if there exists a social-evaluation ordering R on \mathcal{R}^n such that, for all $x, y \in X$ and for all $U \in \mathcal{D}$,*

$$xR_U y \Leftrightarrow U(x)RU(y).$$

Blackorby, Bossert and Donaldson (2004a) prove a generalized version of the welfarism theorem where multiple profiles of individual and social non-welfare information are permitted.

3. Axioms and examples

We now formulate some basic axioms regarding the social-evaluation ordering R . The first of these is anonymity. It ensures that the ordering R treats individuals impartially, paying

no attention to their identities. Thus, any permutation of a utility vector u must be as good as u itself. This is a strengthening of Arrow's (1951, 1963) condition that prevents the existence of a dictator.

Anonymity: For all $u \in \mathcal{R}^n$ and for all bijections $\rho: N \rightarrow N$,

$$uI(u_{\rho(1)}, \dots, u_{\rho(n)}).$$

Pareto principles impose monotonicity properties on the ordering R . They require that the social-evaluation ordering respond positively to increases in utility. We use the following notation for vector inequalities. For all $u, v \in \mathcal{R}^n$, (i) $u \geq v$ if and only if $u_i \geq v_i$ for all $i \in N$; (ii) $u > v$ if and only if $u \geq v$ and $u \neq v$; (iii) $u \gg v$ if and only if $u_i > v_i$ for all $i \in N$.

The weak Pareto principle requires an increase in everyone's utility to be regarded as a social improvement.

Weak Pareto: For all $u, v \in \mathcal{R}^n$, if $u \gg v$, then uPv .

The strong Pareto requirement extends weak Pareto to cases in which no-one's utility decreases and at least one individual's well-being increases.

Strong Pareto: For all $u, v \in \mathcal{R}^n$, if $u > v$, then uPv .

Continuity is a regularity condition. It ensures that 'small' changes in individual utilities do not lead to 'large' changes in the social ranking.

Continuity: For all $u \in \mathcal{R}^n$, the sets $\{v \in \mathcal{R}^n \mid vRu\}$ and $\{v \in \mathcal{R}^n \mid uRv\}$ are closed in \mathcal{R}^n .

The next axiom is an equity requirement; see d'Aspremont and Gevers (1977) and Deschamps and Gevers (1978). It is called minimal equity and it prevents the social ordering from exhibiting a strong version of preference for inequality.

Minimal equity: There exist $u, v \in \mathcal{R}^n$ and $i, j \in N$ such that $u_k = v_k$ for all $k \in N \setminus \{i, j\}$, $v_j > u_j > u_i > v_i$ and uRv .

Finally, we introduce a separability property. This independence condition limits the influence of the well-being of unconcerned individuals on the social ordering. Suppose that a social change affects only the utilities of the members of a population subgroup. Independence of the utilities of unconcerned individuals requires the social assessment of the change to be independent of the utility levels of people outside the subgroup.

Independence of the utilities of unconcerned individuals: For all non-empty $M \subset N$ and for all $u, v, \bar{u}, \bar{v} \in \mathcal{R}^n$, if $[u_i = v_i \text{ and } \bar{u}_i = \bar{v}_i]$ for all $i \in M$ and $[u_j = \bar{u}_j \text{ and } v_j = \bar{v}_j]$ for all $j \in N \setminus M$, then

$$uRv \Leftrightarrow \bar{u}R\bar{v}.$$

In this definition, the individuals in M are the unconcerned—they have the same utilities in u and v and in \bar{u} and \bar{v} . Independence of the utilities of unconcerned individuals requires the ranking of u and v to depend on the utilities of the concerned individuals, those in $N \setminus M$, only. The corresponding separability axiom for social-evaluation functionals can be found in d’Aspremont and Gevers (1977) where it is called separability with respect to unconcerned individuals. d’Aspremont and Gevers’ separability axiom is called elimination of (the influence of) indifferent individuals in Maskin (1978) and in Roberts (1980b). In the case of two individuals, the independence axiom is implied by the strong Pareto principle. Therefore it is usually applied to societies with at least three individuals.

We conclude this section with some examples of social-evaluation orderings, restricting attention to those characterized in this chapter.

The strongly dictatorial social-evaluation orderings pay attention to the utility of a single individual only. That is, R is strongly dictatorial if and only if there exists an individual $k \in N$ such that, for all $u, v \in \mathcal{R}^n$,

$$uRv \Leftrightarrow u_k \geq v_k.$$

Strong dictatorships satisfy weak Pareto, continuity, minimal equity and independence of the utilities of unconcerned individuals. Anonymity and strong Pareto are violated.

A strong positional dictatorship assigns dictatorial power to a position in the society rather than to a named individual. For $u \in \mathcal{R}^n$, let $(u_{(1)}, \dots, u_{(n)})$ be a permutation of u such that $u_{(i)} \geq u_{(i+1)}$ for all $i \in N \setminus \{n\}$. R is a strong positional dictatorship if and only if there exists a position $k \in N$ such that, for all $u, v \in \mathcal{R}^n$,

$$uRv \Leftrightarrow u_{(k)} \geq v_{(k)}.$$

An important special case is the maximin ordering which is obtained for $k = n$, that is, the social ranking is determined by the utility of the worst-off. If $k = 1$, the maximax ordering, which pays attention to the best-off only, results. Strong positional dictatorships satisfy anonymity, weak Pareto and continuity. They violate strong Pareto and independence of the utilities of unconcerned individuals. All strong positional dictatorships except maximax satisfy minimal equity.

Utilitarianism ranks any two utility vectors by comparing their sums of utilities. Thus, according to utilitarianism, for all $u, v \in \mathcal{R}^n$,

$$uRv \Leftrightarrow \sum_{i=1}^n u_i \geq \sum_{i=1}^n v_i.$$

Utilitarianism satisfies all of the axioms introduced earlier in this section. R is a weakly utilitarian ordering if and only if it respects the asymmetric part of utilitarianism, that is, if and only if, for all $u, v \in \mathcal{R}^n$,

$$\sum_{i=1}^n u_i > \sum_{i=1}^n v_i \Rightarrow uPv.$$

The class of Kolm-Pollak orderings (see Kolm, 1969, and Pollak, 1971) is a subclass of the class of generalized-utilitarian principles. Generalized utilitarianism uses the sum of transformed utilities as the criterion for social evaluation and the Kolm-Pollak orderings are obtained for specific transformations. R is a Kolm-Pollak ordering if and only if R is utilitarian or there exists $\gamma \in \mathcal{R}_{++}$ such that, for all $u, v \in \mathcal{R}^n$,

$$uRv \Leftrightarrow -\sum_{i=1}^n e^{-\gamma u_i} \geq -\sum_{i=1}^n e^{-\gamma v_i}.$$

Utilitarianism is the limiting case when γ approaches zero and, therefore, the utilitarian ordering can be defined as the Kolm-Pollak ordering with a parameter value of $\gamma = 0$. As γ approaches infinity, maximin is approximated. All our axioms are satisfied by these orderings.

The class of global means of order r is another subclass of the generalized-utilitarian class. R is a global mean of order r if and only if there exist $r, \beta \in \mathcal{R}_{++}$ such that, for all $u, v \in \mathcal{R}^n$,

$$uRv \Leftrightarrow \sum_{i \in N: u_i \geq 0} (u_i)^r - \beta \sum_{i \in N: u_i < 0} (-u_i)^r \geq \sum_{i \in N: v_i \geq 0} (v_i)^r - \beta \sum_{i \in N: v_i < 0} (-v_i)^r.$$

In this case, the transformation assigns the value τ^r to all non-negative utility levels τ and $-\beta(-\tau)^r$ to all negative τ . Utilitarianism is obtained for the parameter values $r = \beta = 1$. The orderings that result for $r = 1$ and $\beta > 1$ are modifications of utilitarianism such that negative utilities get a higher weight than positive utilities. The global means of order r satisfy all of the axioms of this section.

Leximin is a modified version of maximin in which utility vector u is better than utility vector v if the worst-off individual in u is better off than the worst-off individual in v . If those individuals are equally well off, the utilities of the next-worse-off individuals are used to determine the social ranking, and the procedure continues until either there is a strict ranking or the two utility vectors are permutations of each other, in which case they are declared equally good. Thus, the leximin ordering is defined by letting, for all $u, v \in \mathcal{R}^n$,

$$uRv \Leftrightarrow u \text{ is a permutation of } v \text{ or there exists } j \in N \text{ such that} \\ u_{(j)} = v_{(j)} \text{ for all } i > j \text{ and } u_{(j)} > v_{(j)}.$$

Leximin satisfies all of our axioms except continuity.

4. Information invariance

Information-invariance conditions restrict the information regarding the measurability and interpersonal comparability of individual utilities that can be used in social evaluation. The most common way to represent informational environments is to define the set of invariance transformations that can be applied to utility vectors without changing their informational contents. Information invariance with respect to the information assumption represented by the set of admissible transformations then requires that the ranking of any two utility vectors is the same as the ranking of the transformed vectors. This approach was developed in contributions such as d'Aspremont and Gevers (1977), Roberts (1980a,b) and Sen (1974) and we follow it in this chapter. We present the information-invariance assumptions that are used in the remainder of the chapter and refer the reader to Blackorby and Donaldson (1982), Blackorby, Donaldson and Weymark (1984),! Bossert and Weymark (2004), d'Aspremont (1985), d'Aspremont and Gevers (1977, 2002), DeMeyer and Plott (1971), Dixit (1980), Gevers (1979), Roberts (1980b) and Sen (1970, 1974, 1977, 1986), for example, for more extensive discussions.

If the only information that can be used is ordinal utility information without interpersonal comparability, we obtain Arrow's (1951, 1963) informational environment that requires information invariance with respect to ordinal non-comparability. The set of admissible transformations consists of all n -tuples of independent increasing transformations. Ordinal non-comparability implies that intrapersonal comparisons of utility levels are possible. This is the case because an inequality such as $u_i \geq v_i$ is preserved whenever an increasing transformation is applied to all utility values of individual $i \in N$.

Information invariance with respect to ordinal non-comparability: For all $u, v \in \mathcal{R}^n$ and for all increasing functions $\Phi_1, \dots, \Phi_n: \mathcal{R} \rightarrow \mathcal{R}$,

$$uRv \Leftrightarrow (\Phi_1(u_1), \dots, \Phi_n(u_n)) R (\Phi_1(v_1), \dots, \Phi_n(v_n)).$$

Of the social-evaluation orderings defined in the previous section, only strong dictatorships satisfy information invariance with respect to ordinal non-comparability.

If utilities are cardinally measurable, individual utilities are unique up to increasing affine transformations, thereby allowing for intrapersonal comparisons of utility differences: inequalities such as $u_i - v_i \geq w_i - t_i$ are preserved if an increasing affine transformation is applied to all utility values of individual $i \in N$. If no requirements regarding the interpersonal comparison of utilities are imposed, the transformations may differ across individuals and we obtain the information assumption cardinal non-comparability.

Information invariance with respect to cardinal non-comparability: For all $u, v \in \mathcal{R}^n$, for all $a_1, \dots, a_n \in \mathcal{R}_{++}$ and for all $b_1, \dots, b_n \in \mathcal{R}$,

$$uRv \Leftrightarrow (a_1u_1 + b_1, \dots, a_nu_n + b_n) R (a_1v_1 + b_1, \dots, a_nv_n + b_n).$$

Formulated in terms of a social-evaluation ordering R defined on \mathcal{R}^n , information invariance with respect to ordinal non-comparability and information invariance with respect to cardinal non-comparability are equivalent; see Sen (1970) and, for a diagrammatic illustration, Blackorby, Donaldson and Weymark (1984). Clearly, information invariance with respect to ordinal non-comparability implies information invariance with respect to cardinal non-comparability. To prove the converse implication, suppose R satisfies information invariance with respect to cardinal non-comparability. Let $\Phi_1, \dots, \Phi_n: \mathcal{R} \rightarrow \mathcal{R}$ be increasing and consider any $i \in N$. If $u_i = v_i$, let $a_i = 1$ and $b_i = \Phi_i(u_i) - u_i$. If $u_i \neq v_i$, let

$$a_i = \frac{\Phi_i(v_i) - \Phi_i(u_i)}{v_i - u_i}$$

(which is positive because Φ_i is increasing) and

$$b_i = \frac{v_i\Phi_i(u_i) - u_i\Phi_i(v_i)}{v_i - u_i}.$$

In either case, it follows that $\Phi_i(u_i) = a_iu_i + b_i$ and $\Phi_i(v_i) = a_iv_i + b_i$. By information invariance with respect to cardinal non-comparability,

$$uRv \Leftrightarrow (\Phi_1(u_1), \dots, \Phi_n(u_n)) R (\Phi_1(v_1), \dots, \Phi_n(v_n))$$

which establishes information invariance with respect to ordinal non-comparability.

As an immediate consequence of this equivalence result, it follows that, among the social-evaluation orderings of the previous section, strong dictatorships are the only ones satisfying information invariance with respect to cardinal non-comparability.

If utility levels are comparable both intrapersonally and interpersonally but no further information is available, we obtain ordinal full comparability. In that case, only common increasing transformations can be applied to the utilities without changing the information relevant for social evaluation. This implies that utility levels can be compared interpersonally: the inequality $u_i \geq v_j$ is preserved even for different individuals $i, j \in N$ if a common transformation is applied to all utility values.

Information invariance with respect to ordinal full comparability: For all $u, v \in \mathcal{R}^n$ and for all increasing functions $\Phi_0: \mathcal{R} \rightarrow \mathcal{R}$,

$$uRv \Leftrightarrow (\Phi_0(u_1), \dots, \Phi_0(u_n)) R (\Phi_0(v_1), \dots, \Phi_0(v_n)).$$

Information invariance with respect to ordinal full comparability is satisfied by the strongly dictatorial social-evaluation orderings, the strong positional dictatorships and leximin. The Kolm-Pollak orderings and the global means of order r (including utilitarianism) do not satisfy this invariance property.

If utilities are cardinally measurable and fully interpersonally comparable, both utility levels and utility differences can be compared interpersonally. In this case, the only admissible transformations are increasing affine transformations which are identical across individuals. Utility levels and utility difference can be compared because inequalities both of the form $u_i \geq u_j$ and of the form $u_i - v_i \geq w_j - t_j$ are preserved for all $i, j \in N$ if a common increasing affine transformation is applied to all utilities.

Information invariance with respect to cardinal full comparability: For all $u, v \in \mathcal{R}^n$, for all $a_0 \in \mathcal{R}_{++}$ and for all $b_0 \in \mathcal{R}$,

$$uRv \Leftrightarrow (a_0u_1 + b_0, \dots, a_0u_n + b_0) R (a_0v_1 + b_0, \dots, a_0v_n + b_0).$$

This information-invariance axiom is satisfied by the strong dictatorships, the strong positional dictatorships, utilitarianism and leximin but not by the Kolm-Pollak orderings and the global means of order r that are not utilitarian.

Translation-scale full comparability provides an informational environment in which the numerical values of utility differences are meaningful (and, therefore, interpersonally comparable) and, in addition, utility levels can be compared interpersonally. Admissible transformations are affine with a common scaling factor equal to one and a common additive constant. Clearly, utility differences such as $u_i - v_j$ are unchanged for all $i, j \in N$ if a common constant is added to all individual utilities.

Information invariance with respect to translation-scale full comparability: For all $u, v \in \mathcal{R}^n$ and for all $b_0 \in \mathcal{R}$,

$$uRv \Leftrightarrow (u_1 + b_0, \dots, u_n + b_0) R (v_1 + b_0, \dots, v_n + b_0).$$

Strong dictatorships, strong positional dictatorships, the Kolm-Pollak orderings (including utilitarianism) and leximin satisfy information invariance with respect to translation-scale full comparability; the global means of order r that are not utilitarian do not.

If, instead of a common translation scale, a common ratio scale is employed, we obtain information invariance with respect to ratio-scale full comparability. Utility ratios such as u_i/v_j with $v_j \neq 0$ are preserved if all utility values are multiplied by the same positive constant for all $i, j \in N$.

Information invariance with respect to ratio-scale full comparability: For all $u, v \in \mathcal{R}^n$ and for all $a_0 \in \mathcal{R}_{++}$,

$$uRv \Leftrightarrow (a_0u_1, \dots, a_0u_n) R (a_0v_1, \dots, a_0v_n).$$

Strong dictatorships, strong positional dictatorships, the global means of order r (including utilitarianism) and leximin satisfy information invariance with respect to ratio-scale full comparability; the Kolm-Pollak orderings that are not utilitarian do not.

5. Impossibilities and characterizations

In the absence of interpersonal comparisons of well-being, there do not exist satisfactory social-evaluation principles. A variant of Arrow's (1951, 1963) theorem (see also Blau, 1957) formulated for social-evaluation orderings states that only strong dictatorships satisfy weak Pareto, continuity and information invariance with respect to ordinal non-comparability.

Theorem 2: *R satisfies weak Pareto, continuity and information invariance with respect to ordinal non-comparability if and only if R is a strong dictatorship.*

Sen (1970) proves that replacing information invariance with respect to ordinal non-comparability by information invariance with respect to cardinal non-comparability does not provide an escape from the negative conclusion of Arrow's theorem. This observation is an immediate consequence of the equivalence of the two information-invariance conditions established in the previous section. Thus, we obtain the following result.

Theorem 3: *R satisfies weak Pareto, continuity and information invariance with respect to cardinal non-comparability if and only if R is a strong dictatorship.*

Clearly, strongly dictatorial principles are not anonymous and, therefore, if this fundamental impartiality requirement is added to the list of axioms in Theorem 2 or in Theorem 3, an impossibility results. Because continuity is not required in this impossibility, we omit it.

Theorem 4: *There exists no social-evaluation ordering satisfying anonymity, weak Pareto and information invariance with respect to ordinal or cardinal non-comparability.*

The impossibility result of Theorem 4 remains true if anonymity is weakened to the requirement that rules out the existence of a dictator; see Arrow (1951, 1963).

The conclusion we draw from Theorem 4 is that interpersonal comparisons of utility must be permitted to obtain reasonable principles for social evaluation. One possibility is to replace information invariance with respect to ordinal or cardinal non-comparability by information invariance with respect to ordinal full comparability. If anonymity, weak Pareto and continuity are added, we obtain a characterization of the class of strong positional dictatorships.

Theorem 5: *R satisfies anonymity, weak Pareto, continuity and information invariance with respect to ordinal full comparability if and only if R is a strong positional dictatorship.*

We now consider informational environments where more information than ordinal measurability is available. Deschamps and Gevers (1978) examine the class of social-evaluation orderings satisfying information invariance with respect to cardinal full comparability. If this axioms is added to anonymity, strong Pareto, minimal equity and independence of the utilities of unconcerned individuals, only weakly utilitarian orderings and leximin remain as possibilities. This theorem, which is quite remarkable, is valid for societies with at least three members because, for two-person societies, the independence condition is implied by strong Pareto and fails to be of sufficient strength to obtain the result.

Theorem 6: *Suppose $n \geq 3$. If R satisfies anonymity, strong Pareto, minimal equity, independence of the utilities of unconcerned individuals and information invariance with respect to cardinal full comparability, then R is weakly utilitarian or leximin.*

Theorem 6 is not an if-and-only-if result because not all weakly utilitarian orderings satisfy all axioms: the axioms also place restrictions on how a weakly utilitarian ordering ranks utility vectors which have the same sum. However, it shows that there do not exist many orderings other than utilitarianism and leximin satisfying the axioms of the theorem statement. The following two theorems illustrate how utilitarianism or leximin can be obtained by modifying one or another of the axioms.

We begin with utilitarianism, which has received a considerable amount of attention in the literature on social choice. The following theorem is due to Maskin (1978); see also Deschamps and Gevers (1978). It is obtained by replacing minimal equity with continuity in Theorem 6.

Theorem 7: *Suppose $n \geq 3$. R satisfies anonymity, strong Pareto, continuity, independence of the utilities of unconcerned individuals and information invariance with respect to cardinal full comparability if and only if R is utilitarian.*

d'Aspremont and Gevers (1977) provide an alternative characterization of utilitarianism that does not require the independence axiom; see also Blackwell and Girshick (1954), Milnor (1954) and Roberts (1980b).

A characterization of the leximin ordering due to d'Aspremont and Gevers (1977) replaces information invariance with respect to cardinal full comparability by information invariance with respect to ordinal full comparability in Theorem 6.

Theorem 8: *Suppose $n \geq 3$. R satisfies anonymity, strong Pareto, minimal equity, independence of the utilities of unconcerned individuals and information invariance with respect to ordinal full comparability if and only if R is leximin.*

Hammond (1976) provides an alternative characterization of leximin that replaces minimal equity, independence of the utilities of unconcerned individuals and information invariance with respect to cardinal full comparability with a stronger equity axiom. Because it does not employ an information-invariance condition, we do not state it here.

We conclude this section with characterizations of the Kolm-Pollak orderings and of the class of global means of order r . These results are obtained by replacing the information-invariance assumption of Theorem 7 with information invariance with respect to translation-scale full comparability in the case of the Kolm-Pollak orderings and with respect to ratio-scale full comparability in the case of the global means of order r .

Theorem 9: *Suppose $n \geq 3$. R satisfies anonymity, strong Pareto, continuity, independence of the utilities of unconcerned individuals and information invariance with respect to translation-scale full comparability if and only if R is a Kolm-Pollak ordering.*

Theorem 10: *Suppose $n \geq 3$. R satisfies anonymity, strong Pareto, continuity, independence of the utilities of unconcerned individuals and information invariance with respect to ratio-scale full comparability if and only if R is a global mean of order r .*

6. Concluding remarks

This chapter provides a brief introduction to welfarist social-choice theory. It is argued that the most promising route of escape from the negative conclusion of Arrow's theorem is to consider informational environments that allow for interpersonal comparisons of well-being.

We focus on establishing a social ranking of alternatives in this chapter. Because actual decision problems typically are choice problems, it is natural to ask what to do about constraints facing a society, such as those resulting from resource limitations. The

approach to constrained social choice implicitly followed in this chapter is that of constrained optimization. We propose to select, for each choice problem identified by a feasible set of alternatives, a best element according to the social objective represented by the social ordering. If the suitability of a social objective on ethical grounds is independent of the constraints (a position we advocate), our approach which does not model constraints explicitly does not involve any loss of generality.

The model discussed in this chapter can be generalized in various ways. Two of the most important extensions involve considerations of uncertainty and the possibility of population change. Due to space constraints, we cannot present them in detail and provide brief summaries instead, accompanied by some suggestions for further reading.

Welfarist social evaluation under uncertainty is discussed, for example, in Blackorby, Bossert and Donaldson (2002, 2003), Blackorby, Donaldson and Weymark (1999, 2004), Harsanyi (1955, 1977), Mongin (1994) and Weymark (1991, 1993, 1994, 1995). While most contributions (such as those of Harsanyi) focus on the ranking of probability distributions, an attractive alternative is to fix the probabilities of the possible states that may occur and represent the social-choice situation under uncertainty by considering prospects—vectors of alternatives, one for each possible state that may materialize. This model is a natural generalization of the one considered here. Instead of actual (ex-post) utility functions, ex-ante utilities are employed and a welfarism theorem that is formulated in terms of ex-ante utility functions is obtained. See Blackorby, Bossert and Donaldson (2003) for details.

Variable-population social choice is discussed, for example, in Blackorby and Donaldson (1984) and in Blackorby, Bossert and Donaldson (2004b). As is the case for uncertainty, natural generalizations of welfarism can be characterized if the population is allowed to vary from one alternative to another. Moreover, the model can be extended to cover both uncertainty and variable-population issues at the same time.

REFERENCES

- Arrow, K., 1951 (second ed. 1963), *Social Choice and Individual Values*, Wiley, New York.
- Blackorby, C., W. Bossert and D. Donaldson, 2002, Utilitarianism and the theory of justice, in *Handbook of Social Choice and Welfare, vol. 1*, K. Arrow, A. Sen and K. Suzumura, eds., Elsevier, Amsterdam, 543–596.
- Blackorby, C., W. Bossert and D. Donaldson, 2003, Harsanyi's theorem: a multi-profile approach and variable-population extensions, Discussion Paper 03-2003, Université de Montréal, CIREQ.
- Blackorby, C., W. Bossert and D. Donaldson, 2004a, Multi-profile welfarism: a generalization, *Social Choice and Welfare*, forthcoming.
- Blackorby, C., W. Bossert and D. Donaldson, 2004b, Population issues in social-choice theory, welfare economics and ethics, Cambridge University Press, Cambridge, forthcoming.

- Blackorby, C. and D. Donaldson, 1982, Ratio-scale and translation-scale full interpersonal comparability without domain restrictions: admissible social evaluation functions, *International Economic Review* **23**, 249–268.
- Blackorby, C. and D. Donaldson, 1984, Social criteria for evaluating population change, *Journal of Public Economics* **25**, 13–33.
- Blackorby, C. and D. Donaldson, 1992, Pigs and guinea pigs: a note on the ethics of animal exploitation, *Economic Journal* **102**, 1345–1369.
- Blackorby, C., D. Donaldson and J. Weymark, 1984, Social choice with interpersonal utility comparisons: a diagrammatic introduction, *International Economic Review* **25**, 327–356.
- Blackorby, C., D. Donaldson and J. Weymark, 1999, Harsanyi’s social aggregation theorem for state-contingent alternatives, *Journal of Mathematical Economics* **32**, 365–387.
- Blackorby, C., D. Donaldson and J. Weymark, 2004, Social aggregation and the expected utility hypothesis, in *Justice, Political Liberalism, and Utilitarianism: Themes from Harsanyi and Rawls*, M. Salles and J. Weymark, eds., Cambridge University Press, Cambridge, forthcoming.
- Blackwell, D. and M. Girshick, 1954, *Theory of Games and Statistical Decisions*, Wiley, New York.
- Blau, J., 1957, The existence of social welfare functions, *Econometrica* **25**, 302–313.
- Bossert, W. and J. Weymark, 2004, Utility in social choice, in *Handbook of Utility Theory, vol. 2: Extensions*, S. Barberà, P. Hammond and C. Seidl, eds., Kluwer, Dordrecht, 1099–1177.
- d’Aspremont, C., 1985, Axioms for social welfare orderings, in *Social Goals and Social Organizations: Essays in Memory of Elisha Pazner*, L. Hurwicz, D. Schmeidler and H. Sonnenschein, eds., Cambridge University Press, Cambridge, 19–76.
- d’Aspremont, C. and L. Gevers, 1977, Equity and the informational basis of collective choice, *Review of Economic Studies* **44**, 199–209.
- d’Aspremont, C. and L. Gevers, 2002, Social welfare functionals and interpersonal comparability, in *Handbook of Social Choice and Welfare, vol. 1*, K. Arrow, A. Sen and K. Suzumura, eds., Elsevier, Amsterdam, 459–541.
- DeMeyer, F. and C. Plott, 1971, A welfare function using ‘relative intensity’ of preference, *Quarterly Journal of Economics* **85**, 179–186.
- Deschamps, R. and L. Gevers, 1978, Leximin and utilitarian rules: a joint characterization, *Journal of Economic Theory* **17**, 143–163.
- Dixit, A., 1980, Interpersonal comparisons and social welfare functions, unpublished manuscript, University of Warwick, Department of Economics.
- Gevers, L., 1979, On interpersonal comparability and social welfare orderings, *Econometrica* **47**, 75–89.
- Goodin, R., 1991, Utility and the good, in *A Companion to Ethics*, P. Singer, ed., Basil Blackwell, Oxford, 241–248.
- Griffin, J., 1986, *Well-Being: Its Meaning, Measurement, and Moral Importance*, Clarendon, Oxford.

- Hammond, P., 1976, Equity, Arrow's conditions, and Rawls' difference principle, *Econometrica* **44**, 793–804.
- Hammond, P., 1979, Equity in two person situations: some consequences, *Econometrica* **47**, 1127–1135.
- Harsanyi, J., 1955, Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility, *Journal of Political Economy* **63**, 309–321.
- Harsanyi, J., 1977, *Rational Behavior and Bargaining Equilibrium in Games and Social Situations*, Cambridge University Press, Cambridge.
- Kolm, S.-C., 1969, The optimal production of social justice, in *Public Economics*, J. Margolis and S. Guitton, eds., Macmillan, London, 145–200.
- Maskin, E., 1978, A theorem on utilitarianism, *Review of Economic Studies* **45**, 93–96.
- Milnor, J., 1954, Games against nature, in *Decision Processes*, R. Thrall, C. Coombs and R. Davis, eds., Wiley, New York, 49–59.
- Mongin, P., 1994, Harsanyi's aggregation theorem: multi-profile version and unsettled questions, *Social Choice and Welfare* **11**, 331–354.
- Pollak, R., 1971, Additive utility functions and linear Engel curves, *Review of Economic Studies* **38**, 401–414.
- Roberts, K., 1980a, Possibility theorems with interpersonally comparable welfare levels, *Review of Economic Studies* **47**, 409–420.
- Roberts, K., 1980b, Interpersonal comparability and social choice theory, *Review of Economic Studies* **47**, 421–439.
- Sen, A., 1970, *Collective Choice and Social Welfare*, Holden-Day, San Francisco.
- Sen, A., 1974, Informational bases of alternative welfare approaches: aggregation and income distribution, *Journal of Public Economics* **3**, 387–403.
- Sen, A., 1977, On weights and measures: informational constraints in social welfare analysis, *Econometrica* **45**, 1539–1572.
- Sen, A., 1979, Personal utilities and public judgements: or what's wrong with welfare economics?, *Economic Journal* **89**, 537–558.
- Sen, A., 1986, Social choice theory, in *Handbook of Mathematical Economics, vol. III*, K. Arrow and M. Intriligator, eds., North-Holland, Amsterdam, 1073–1181.
- Weymark, J., 1991, A reconsideration of the Harsanyi-Sen debate on utilitarianism, in *Interpersonal Comparisons of Well-Being*, J. Elster and J. Roemer, eds., Cambridge University Press, Cambridge, 255–320.
- Weymark, J., 1993, Harsanyi's social aggregation theorem and the weak Pareto principle, *Social Choice and Welfare* **10**, 209–222.
- Weymark, J., 1994, Harsanyi's social aggregation theorem with alternative Pareto principles, in *Models and Measurement of Welfare and Inequality*, W. Eichhorn, ed., Springer, Berlin, 869–887.
- Weymark, J., 1995, Further remarks on Harsanyi's social aggregation theorem and the weak Pareto principle, *Social Choice and Welfare* **12**, 87–92.