MULTI-PROFILE WELFARISM: A GENERALISATION

Charles Blackorby

Walter Bossert

And

David Donaldson

No 710

WARWICK ECONOMIC RESEARCH PAPERS

DEPARTMENT OF ECONOMICS

THE UNIVERSITY OF WARWICK

Multi-Profile Welfarism: A Generalization*

Charles Blackorby, Walter Bossert and David Donaldson

Final version, February 2004

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

David Donaldson: Department of Economics, University of British Columbia, dvdd@telus.net

Abstract. This paper characterizes welfarist social evaluation in a multi-profile setting where, in addition to multiple utility profiles, there may be more than one profile of non-welfare information. We prove a new version of the welfarism theorem in this alternative framework, and we demonstrate that adding a plausible and weak anonymity property to the welfarism axioms generates welfarist social-evaluation orderings that are anonymous. *Journal of Economic Literature* Classification Number: D63.

Keywords: Welfarism, Multiple-Profile Social Choice.

* This is a revised version of our earlier paper "Welfarism with multiple non-welfare profiles." We thank François Maniquet, Kotaro Suzumura, John Weymark, an editor and two referees for comments and discussions. The paper was presented at the International Symposium on Intergenerational Equity in Tokyo, December 2001 and at the Meeting of the Society for Social Choice and Welfare in Pasadena, July 2002. Financial support through a grant from the Social Sciences and Humanities Research Council of Canada is gratefully acknowledged.

1. Introduction

Welfarist principles for social evaluation rank social alternatives using information about individual well-being (welfare, utility) alone, ignoring non-welfare information. As a result, those principles regard things such as liberty, freedom of expression or a healthy environment as desirable because of their contribution to well-being. Welfarist principles can, however, account for the needs of the disadvantaged and give priority to worse-off people. Principles that are inequality-averse in utilities have this property and utilitarianism—which is insensitive to utility inequality—gives priority to those who benefit most in terms of well-being. Thus, if individual utility is a strictly concave function of consumption, the criterion used for the allocation of resources favours the disadvantaged.

Welfarism does not, by itself, have conservative implications for economic policy. As an example, there are many objections, consistent with welfarism, to markets for body parts (such as kidneys) for transplantation: sellers may favour short-term gain over their own long-term well-being; sellers may be poorly informed; buyers may have monopsony power over desperate sellers; inequality of well-being may be increased; and criminal acquisition of body parts may increase. The desirability of such markets therefore depends on factual information in addition to the principle for social evaluation, welfarist or not, that is employed.

Sen [1987] has criticized welfarism on the grounds that preferences or desires may not always be consistent with well-being, noting that individual preferences may be affected by incomplete information and that "the underdog comes to terms with social inequalities by bringing desires in line with feasibilities" (Sen [1987, p. 11]). Because of this, welfarist principles should be coupled with accounts of well-being, such as those of Broome [1991], Griffin [1986], Mongin and d'Aspremont [1998] and Sumner [1996], that take account of information problems, are based on individuals' subjective self-interest and are comprehensive enough to capture all aspects of the good life. Due to the complete-information and self-interest qualifications of these accounts, expressed preferences may not always be consistent with individual well-being. Without such accounts of well-being, the appeal of the welfarism axioms would be significantly diminished.

A principle for social evaluation is a social-evaluation functional which associates an ordering of the alternatives with each possible information profile. Such a functional is welfarist if and only if there is a single social-evaluation ordering of utility vectors such that, for all information profiles, the ranking of any two alternatives is given by the ranking of the corresponding utility vectors without regard to non-welfare information.

Conventional social-choice theory employs multiple profiles of welfare (utility) information only: non-welfare information is implicitly fixed. In that setting, welfarism is a consequence of the axioms unlimited domain, Pareto indifference and binary independence of irrelevant alternatives.[1] Because non-welfare information is fixed, it is impossible to discern the way in which a principle makes use of it. For that, multiple non-welfare profiles are needed.

In this paper, we present a characterization of welfarism in a framework in which both social and individual non-welfare information may vary across information profiles. Social non-welfare information may include information about the presence or absence of democratic institutions or freedom of the press. Individual non-welfare information may include length of life, whether the person has a propensity to work hard and whether he or she likes classical music.

Each information profile includes a vector of individual utility functions which represent welfare information and a vector of functions which describe social and individual non-welfare information.[2] In that setting, the independence axiom is formulated in terms of both welfare and non-welfare information and it, together with unlimited domain and Pareto indifference, is used to make a case in favour of welfarism.

Our approach permits a compelling justification of anonymous welfarism. The standard axiom requires the social ordering to be unaffected by a permutation of utility functions across individuals with non-welfare information unchanged. It is possible, however, that some individual may have non-welfare characteristics, such as being hardworking, that may be thought to justify special consideration and this lessens the ethical attractiveness of the axiom. By contrast, our anonymity axiom requires the social ordering to be unaffected if *both* utility functions and individual non-utility-information functions are permuted. Together with a restriction on the ranges of the individual non-welfare-information functions (which is needed to ensure that permuted profiles are well-defined) and our other conditions, it implies that the social-evaluation ordering must be anonymous: it ranks all permutations of any utility vector as equally good.

If, in any two alternatives, each person is equally well off, the Pareto-indifference axiom requires the two alternatives to be ranked as equally good. Pareto indifference is implied by an axiom based on the view that, if one alternative is ranked as better than another, it must be better for at least one person (Goodin [1991]). Without this requirement, we run the risk of recommending social changes that are empty gestures, benefitting no one and, perhaps, harming some or all. We use this intuition, which is fundamental for welfarist social evaluation, to underline the ethical appeal of the Pareto-indifference requirement. Pareto indifference is employed in our theorems because the stronger axiom is not needed to prove the results. See Blackorby, Bossert and Donaldson

[1] See, for example, Blackorby, Bossert and Donaldson [2002], Bossert and Weymark [2003], d'Aspremont and Gevers [1977], Guha [1972], Hammond [1979], Sen [1977, 1979] and Weymark [1998].

[2] See also Kelsey [1987] and Roberts [1980] for approaches to social choice where non-welfare information is explicitly modelled.

[2002] for a discussion of the relationship between various Pareto axioms and Goodin's intuition.

In Section 2, we discuss accounts of individual well-being that we consider suitable in connection with welfarist social evaluation. Section 3 contains the definitions of our notation, the welfarism axioms and social-evaluation functionals. In Section 4, we prove that any social-evaluation functional with an unlimited domain that satisfies Pareto indifference and binary independence of irrelevant alternatives must be welfarist, disregarding all non-welfare information. In Section 5, we characterize welfarist social-evaluation orderings that are anonymous. Section 6 concludes.

2. Individual well-being

We begin with a discussion of the accounts of individual well-being that are, in our opinion, the most suitable for welfarist social evaluation. Because welfarist principles regard individual well-being as the only entity with intrinsic value, the account of well-being that is used in a welfarist principle should be comprehensive enough to capture all aspects of the good life. Such accounts are provided by Broome [1991], Griffin [1986], Mongin and d'Aspremont [1998] and Sumner [1996].

Bentham [1789, 1973] understands individual well-being in terms of pleasure and pain. Life is seen as a series of pleasurable or painful experiences, differing only in intensity and duration, and well-being or utility is seen as an aggregate that measures overall hedonic value. Although this view has been rejected as too narrow, it contains the important idea that well-being is mediated by experience (see Griffin [1986, Chapter 1]). If someone's experiences are identical in two alternatives, therefore, he or she must be equally well off in both. We find this view attractive, but it is not needed for welfarism. As an example, it has been suggested that events that occur after a person's death may influence his or her well-being. Although it is not our own, this view can be made consistent with an account of individual well-being. Note that events that take place after someone's death are very different from expectations regarding the events that may occur after life is over.

Bentham's theory implies that individual well-being is subjective: if one alternative is better than another for someone, it must be better for the person who is the subject of the life, not better by some external standard. A theory treats well-being as subjective if it makes it depend, at least in part, on some actual or hypothetical attitude on the part of the person (see Sumner [1996, Chapter 2]).

It is difficult to maintain that all the elements of the good life are reducible to pleasure and pain. Both pleasure and pain are complex, multi-faceted experiences. In addition, enjoyment, freedom from anxiety, good health, limbs and senses that work well, length of life (when it is worth living), autonomy, liberty, understanding, accomplishment, satisfying work and good human relationships also make significant contributions to well-being. Moreover, there is a moral dimension to well-being. Most people value being moral agents and they want to contribute to a better world.

A 'list' view of well-being (Griffin [1986]) uses an enumeration of basic elements of the good life such as the one above. A ceteris paribus increase in any element on the list increases well-being. But individual people may differ in the way that the items on the list contribute to their welfare. It may be best for a person not to be autonomous, for example, if he or she is plagued by anxiety. In addition, the importance of an ability depends on the skills a person has and intends to use. A musician might place a great value on the ability to move his or her fingers quickly.

Sumner [1996] presents an account of well-being that focuses on happiness. Happiness is equated with life satisfaction "which has both an affective component (experiencing the conditions of your life as fulfilling and rewarding) and a cognitive component (judging that your life is going well for you)" (p. 172). Like Griffin, Sumner allows for many determinants of well-being but sees their importance in their contribution to happiness. Self-evaluations are useful as long as the person is informed and autonomous.

Desire and preference accounts identify well-being with the satisfaction of self-interested individual wants.[3] The (hypothetical) person must be fully informed and, for that reason, the preferences thus identified do not, in general, coincide with actual preferences. Sen [1987, p. 11] criticizes such accounts of well-being on the grounds that "the battered slave, the broken unemployed, the hopeless destitute, the tamed housewife, may have the courage to desire little." This observation points to the need for full-information and, possibly, autonomy qualifications.

A theory of 'functionings and capabilities,' presented by Sen [1985], is similar to a list view of well-being with an added dimension. Functionings are the 'doings and beings' a person achieves. Refining the list of possible functionings to the list actually used is seen as a valuational exercise and aggregation of the items on the resulting list is influenced by individual differences. Unlike in Griffin's approach, Sen views capabilities as opportunities to achieve various functionings and they are seen as valuable in themselves.[4] The presence of capabilities on Sen's list gives him a way to value individual liberty.

It is possible to employ Sen's theory in a welfarist context, nevertheless. What is needed is an individual goodness relation which ranks all the possible combinations of functionings and capabilities. Although the resulting view of well-being would be more

[3] See Broome [1991], Griffin [1986], Mongin and d'Aspremont [1998] and Sumner [1996] for discussions.

[4] Nussbaum [2000a,b] focuses almost exclusively on capabilities. For a discussion of Sen's approach, see Sumner [1996, pp. 60–68].

objective than the ones considered above, there would be no difficulty in using it with welfarist principles.

To summarize, our view is that theories of well-being such as the ones of Griffin and Sumner capture the complexities of individual well-being best. Both are subjective and provide the comprehensive accounts needed if welfarist social rankings are to assign value to things such as liberty, freedom and good human relationships.

3. Basic definitions

The set of all positive integers is denoted by $\mathcal{Z}_{++}$ and the set of real numbers by $\mathcal{R}$. For $n \in \mathcal{Z}_{++}$, let $\mathcal{R}^n$ be the n-fold Cartesian product of $\mathcal{R}$. Our notation for vector inequalities is $\geq$, $>$ and $\gg$.

The (fixed) set of individuals is $N = \{1, \ldots, n\}$ with $n \in \mathcal{Z}_{++}$. The set of alternatives is X, and we assume that it contains at least three elements.

A utility (welfare) profile is an n-tuple $U = (U_1, \ldots, U_n)$, where $U_i: X \to \mathcal{R}$ is the utility function of individual $i \in N$. Utility is an index of individual well-being. The set of all possible utility profiles is $\mathcal{U}$, and we write $U(x) = (U_1(x), \ldots, U_n(x))$ for all $x \in X$ and for all $U \in \mathcal{U}$.

Non-welfare information is described by a profile $K = (K_0, K_1, \ldots, K_n)$, where $K_0: X \to \mathcal{S}_0$ is a function that associates social non-welfare information with each alternative in X and, for all $i \in N$, $K_i: X \to \mathcal{S}_i$ associates individual non-welfare information for individual i with each alternative in X. The set $\mathcal{S}_0 \neq \emptyset$ is the set of possible values of social non-welfare information and, for all $i \in N$, $\mathcal{S}_i \neq \emptyset$ is the set of possible values for individual i's non-welfare information. The set of all possible profiles of non-welfare information is $\mathcal{K}$ and, for all $x \in X$ and for all $K \in \mathcal{K}$, we define $K(x) = (K_0(x), K_1(x), \ldots, K_n(x))$.

The set of all orderings on X is denoted by $\mathcal{O}$. A social-evaluation functional is a mapping $F: \mathcal{D} \to \mathcal{O}$, where $\mathcal{D} \subseteq \mathcal{U} \times \mathcal{K}$ and $\mathcal{D} \neq \emptyset$. We use the notation $\Upsilon = (U, K)$ and, for convenience, we define $R_\Upsilon = F(\Upsilon)$ for all $\Upsilon \in \mathcal{D}$. The asymmetric and symmetric factors of R_Υ are denoted by P_Υ and I_Υ. Furthermore, we write $\Upsilon(x) = (U(x), K(x))$ for all $x \in X$ and for all $\Upsilon \in \mathcal{D}$.

The first axiom we introduce is a generalization of the standard unlimited-domain assumption. We assume that the social-evaluation functional is capable of producing a social ordering for all logically possible profiles of welfare and non-welfare information.

Unlimited Domain: $\mathcal{D} = \mathcal{U} \times \mathcal{K}$.

Although unlimited domain requires all possible profiles of non-welfare information to be in the domain of the social-evaluation functional, the sets of possibilities for non-welfare information can be different for different individuals. As an example, consider a society of two individuals and four alternatives in which, for simplicity, there is no social non-welfare information and individual non-welfare information consists in specifying whether the person is fat (f) or thin (t). Person 1 can be fat or thin so $\mathcal{S}_1 = \{f, t\}$, but person 2 is thin in all profiles so $\mathcal{S}_2 = \{t\}$. The possibilities for the two individuals are, therefore, (f, t) and (t, t) and, by suitable choice of a non-welfare profile in $\mathcal{K}$, they can be assigned independently to the four alternatives. Consequently, there are sixteen profiles in $\mathcal{K}$. If $\mathcal{S}_i$ is a singleton for all $i \in N$, a single profile is produced but, in it, non-welfare information is not necessarily the same in all alternatives. Thus, the standard fixed non-welfare-information profile is not a special case.

The assumption that the domain is a Cartesian product is important in the proofs of our theorems. It might be argued that certain non-welfare characteristics, such as extreme disabilities, limit the possibilities for well-being. Because profiles in the domain are used to investigate the properties of social-evaluation functionals, however, the multi-profile approach uses all possible profiles, including 'unlikely' ones. Alternatively, it is possible to classify individual preferences, which are normally correlated with well-being, as a kind of non-welfare information. Preferences can and do reveal information about well-being, at least when people are fully informed, rational and autonomous adults. For that reason, preferences should not be regarded as independent components of non-welfare information.

It is possible to restrict the domain somewhat by making the sets $\mathcal{S}_0, \mathcal{S}_1, \ldots, \mathcal{S}_n$ conditional on the utility vector achieved in an alternative. Although this is a technical possibility, we do not believe that it increases the power of the theorems significantly and we therefore use the simpler domain.

Pareto indifference requires any two alternatives to be ranked as equally good whenever each individual is equally well off in both.

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

Our Pareto-indifference assumption is different from the usual one in welfare economics which is applied to preferences rather than well-being. The intent of the standard assumption is, however, to use preferences as a proxy for well-being. If preferences and well-being generate the same ranking of alternatives, the two conditions coincide.

Binary independence of irrelevant alternatives is a condition that ensures consistency across profiles. It requires the social ranking of any two alternatives to depend on the utility and non-welfare information associated with those two alternatives only. An important property of this axiom is that it does not prevent non-welfare information from being taken into consideration.

Binary Independence of Irrelevant Alternatives: For all $x, y \in X$ and for all $\Upsilon, \bar{\Upsilon} \in \mathcal{D}$, if $\Upsilon(x) = \bar{\Upsilon}(x)$ and $\Upsilon(y) = \bar{\Upsilon}(y)$, then

$$xR_{\Upsilon}y \Leftrightarrow xR_{\bar{\Upsilon}}y.$$

We conclude this section with a formulation of strong neutrality. If the utility vectors for alternatives x and y in one profile are the same as the utility vectors for two (possibly different) alternatives z and w in another, strong neutrality requires the ranking of x and y by the social ordering associated with the first profile to be the same as the ranking of z and w by the social ordering associated with the second.

Strong Neutrality: For all $x, y, z, w \in X$ and for all $\Upsilon, \bar{\Upsilon} \in \mathcal{D}$, if $U(x) = \bar{U}(z)$ and $U(y) = \bar{U}(w)$, then

$$xR_{\Upsilon}y \Leftrightarrow zR_{\bar{\Upsilon}}w.$$

4. Welfarism

Our first step toward proving a welfarism theorem with multiple non-welfare profiles consists of showing that unlimited domain, Pareto indifference and binary independence of irrelevant alternatives together imply that the social ordering cannot depend on non-welfare information. It is easy to see why this is the case if there are four or more alternatives.

	Welfare Information				**Non-Welfare Information**			
	x	y	z	w	x	y	z	w
Profile Υ	u	v			k	ℓ		
Profile Υ^1	u	v	u	v	k	ℓ	$\bar{k}$	$\bar{\ell}$
Profile Υ^2	u	v	u	v	$\bar{k}$	$\bar{\ell}$	$\bar{k}$	$\bar{\ell}$
Profile $\bar{\Upsilon}$	u	v			$\bar{k}$	$\bar{\ell}$		

Table 1

In Table 1, x, y, z and w are distinct alternatives, entries under the welfare-information heading are utility vectors and entries under the non-welfare-information heading are non-welfare-information vectors. In profile Υ, utility vectors for x and y are $u \in \mathcal{R}^n$ and $v \in \mathcal{R}^n$ and non-welfare information vectors for x and y are $k \in \mathcal{S}_0 \times \mathcal{S}_1 \times \ldots \times \mathcal{S}_n$ and

$\ell \in \mathcal{S}_0 \times \mathcal{S}_1 \times \ldots \times \mathcal{S}_n$. In profile $\bar{\Upsilon}$, utility vectors for x and y are the same, but the non-welfare-information vectors may be different and are denoted by $\bar{k}$ and $\bar{\ell}$. Information for all other alternatives is unspecified and can be anything in the domain.

We show that the ranking of x and y by R_Υ, the ordering corresponding to profile Υ, is the same as the ranking of x and y by $R_{\bar{\Upsilon}}$, the ordering corresponding to profile $\bar{\Upsilon}$. To do so, we construct two other profiles which are feasible by unlimited domain. Profile Υ^1 coincides with profile Υ on x and y but is specified for z and w. By binary independence of irrelevant alternatives, the rankings of x and y by R_Υ and R_{Υ^1} are the same. Because the pairs (x, z) and (y, w) have the same utility vectors, Pareto indifference requires R_{Υ^1} to declare x and z to be equally good and y and w to be equally good. Consequently, the two pairs are ranked in the same way by R_{Υ^1}. Profiles Υ^1 and Υ^2 coincide on z and w and, by binary independence, the rankings of z and w by R_{Υ^1} and R_{Υ^2} are identical . In addition, Pareto indifference requires R_{Υ^2} to rank the pairs (x, y) and (z, w) in the same way. Because profiles Υ^2 and $\bar{\Upsilon}$ coincide on x and y, binary independence requires the rankings of x and y by R_{Υ^2} and $R_{\bar{\Upsilon}}$ to be the same. Together, these observations prove the result.

The above discussion provides only a partial demonstration. The additional complexity in the following proof is a consequence of the possibility that X may contain three distinct elements only.

Theorem 1: *If F satisfies unlimited domain, Pareto indifference and binary independence of irrelevant alternatives, then, for all $x, y \in X$ and for all $\Upsilon, \bar{\Upsilon} \in \mathcal{D}$ such that $U(x) = \bar{U}(x)$ and $U(y) = \bar{U}(y)$,*

$$xR_\Upsilon y \Leftrightarrow xR_{\bar{\Upsilon}} y. \tag{1}$$

Proof. Let $x, y \in X$ and $\Upsilon, \bar{\Upsilon} \in \mathcal{D}$ be such that $U(x) = \bar{U}(x)$ and $U(y) = \bar{U}(y)$. Let $u = U(x) = \bar{U}(x)$, $v = U(y) = \bar{U}(y)$, $k = K(x)$, $\ell = K(y)$, $\bar{k} = \bar{K}(x)$ and $\bar{\ell} = \bar{K}(y)$. Because X contains at least three alternatives, there exists $z \in X \setminus \{x, y\}$. By unlimited domain, we can define the profiles Υ^1, Υ^2, Υ^3 and Υ^4 as follows. Let $\Upsilon^1(x) = (u, k)$, $\Upsilon^1(y) = (v, \ell)$, $\Upsilon^1(z) = (v, \bar{\ell})$, $\Upsilon^2(x) = (u, k)$, $\Upsilon^2(y) = (v, \bar{\ell})$, $\Upsilon^2(z) = (v, \bar{\ell})$, $\Upsilon^3(x) = (u, k)$, $\Upsilon^3(y) = (v, \bar{\ell})$, $\Upsilon^3(z) = (u, \bar{k})$, $\Upsilon^4(x) = (u, \bar{k})$ $\Upsilon^4(y) = (v, \bar{\ell})$ and $\Upsilon^4(z) = (u, \bar{k})$.

By binary independence of irrelevant alternatives, we have

$$xR_\Upsilon y \Leftrightarrow xR_{\Upsilon^1} y.$$

By Pareto indifference, $yI_{\Upsilon^1} z$ and it follows that

$$xR_{\Upsilon^1} y \Leftrightarrow xR_{\Upsilon^1} z.$$

Using binary independence again, we obtain

$$xR_{\Upsilon^1} z \Leftrightarrow xR_{\Upsilon^2} z.$$

By Pareto indifference, $zI_{\Upsilon^2}y$ and, therefore,

$$xR_{\Upsilon^2}z \Leftrightarrow xR_{\Upsilon^2}y.$$

Now binary independence implies

$$xR_{\Upsilon^2}y \Leftrightarrow xR_{\Upsilon^3}y.$$

By Pareto indifference, $xI_{\Upsilon^3}z$ and it follows that

$$xR_{\Upsilon^3}y \Leftrightarrow zR_{\Upsilon^3}y.$$

Using binary independence again, we obtain

$$zR_{\Upsilon^3}y \Leftrightarrow zR_{\Upsilon^4}y.$$

By Pareto indifference, $zI_{\Upsilon^4}x$ and it follows that

$$zR_{\Upsilon^4}y \Leftrightarrow xR_{\Upsilon^4}y.$$

Using binary independence once more, we obtain

$$xR_{\Upsilon^4}y \Leftrightarrow xR_{\bar{\Upsilon}}y.$$

Combining the above equivalences, (1) results. ■

If two profiles have the same welfare profiles, Theorem 1 demonstrates that the corresponding social orderings must be identical. Analogously to the standard single-non-welfare-profile environment (see, for example, Blau [1976], Bossert and Weymark [2003], d'Aspremont and Gevers [1977], Guha [1972] and Sen [1977]), it is straightforward to show that Pareto indifference and binary independence of irrelevant alternatives together are equivalent to strong neutrality if F satisfies unlimited domain.

Theorem 2: *Suppose F satisfies unlimited domain. F satisfies Pareto indifference and binary independence of irrelevant alternatives if and only if F satisfies strong neutrality.*

Proof. First, suppose that F satisfies strong neutrality. That binary independence of irrelevant alternatives is satisfied follows from setting $x = z$ and $y = w$ in the definition of strong neutrality. To show that Pareto indifference is implied, let $U = \bar{U}$ and $y = z = w$. Strong neutrality implies that $xR_{\Upsilon}y$ if and only if $yR_{\Upsilon}y$ whenever $U(x) = U(y)$. Because R_{Υ} is reflexive, this implies $xI_{\Upsilon}y$.

Now suppose that F satisfies unlimited domain, Pareto indifference and binary independence of irrelevant alternatives. By Theorem 1, we know that non-welfare information is irrelevant. Consider two profiles $\Upsilon, \bar{\Upsilon} \in \mathcal{D}$ and four (not necessarily distinct) alternatives $x, y, z, w \in X$ such that $U(x) = \bar{U}(z) = u$ and $U(y) = \bar{U}(w) = v$.

By unlimited domain, there exist profiles $\Upsilon^1, \Upsilon^2, \Upsilon^3, \Upsilon^4 \in \mathcal{D}$ such that $U^1(x) = u$, $U^1(y) = v$, $U^1(w) = v$, $U^2(x) = u$, $U^2(y) = v$, $U^2(w) = v$, $U^3(x) = u$, $U^3(y) = v$, $U^3(z) = u$, $U^4(y) = v$, $U^4(z) = u$ and $U^4(w) = v$.

By binary independence of irrelevant alternatives,

$$xR_{\Upsilon}y \Leftrightarrow xR_{\Upsilon^1}y.$$

By Pareto indifference, $yI_{\Upsilon^1}w$ and, therefore,

$$xR_{\Upsilon^1}y \Leftrightarrow xR_{\Upsilon^1}w.$$

Using binary independence of irrelevant alternatives again, we obtain

$$xR_{\Upsilon^1}w \Leftrightarrow xR_{\Upsilon^2}w.$$

By Pareto indifference, $wI_{\Upsilon^2}y$ and, therefore,

$$xR_{\Upsilon^2}w \Leftrightarrow xR_{\Upsilon^2}y.$$

By binary independence of irrelevant alternatives,

$$xR_{\Upsilon^2}y \Leftrightarrow xR_{\Upsilon^3}y.$$

By Pareto indifference, $xI_{\Upsilon^3}z$ and, therefore,

$$xR_{\Upsilon^3}y \Leftrightarrow zR_{\Upsilon^3}y.$$

By binary independence of irrelevant alternatives,

$$zR_{\Upsilon^3}y \Leftrightarrow zR_{\Upsilon^4}y.$$

By Pareto indifference, $yI_{\Upsilon^4}w$ and, therefore,

$$zR_{\Upsilon^4}y \Leftrightarrow zR_{\Upsilon^4}w.$$

Using binary independence of irrelevant alternatives once more, we obtain

$$zR_{\Upsilon^4}w \Leftrightarrow zR_{\bar{\Upsilon}}w.$$

Combining the above equivalences, we obtain

$$xR_{\Upsilon}y \Leftrightarrow zR_{\bar{\Upsilon}}w,$$

and strong neutrality is satisfied. ■

Given unlimited domain and our assumption that X contains at least three elements, strong neutrality is equivalent to the existence of a social-evaluation ordering R on $\mathcal{R}^n$ which can be used to rank the alternatives in X for any profile $\Upsilon \in \mathcal{D}$.[5] The asymmetric and symmetric factors of R are P and I. Combined with Theorem 2, this observation yields the following welfarism theorem.[6]

[5] Gevers [1979] uses the term social-welfare ordering for R.

[6] See d'Aspremont and Gevers [1977] and Hammond [1979] for a version with a single non-welfare profile. Bordes, Hammond and Le Breton [1997] and Weymark [1998] prove variants of this theorem with specific domain restrictions, again in the single-non-welfare-profile case.

Theorem 3: *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 $\Upsilon \in \mathcal{D}$,*

$$xR_\Upsilon y \Leftrightarrow U(x)RU(y). \tag{2}$$

Proof. Clearly, if there exists a social-evaluation ordering R such that (2) is satisfied for all $x, y \in X$ and for all $\Upsilon \in \mathcal{D}$, then F satisfies Pareto indifference and binary independence of irrelevant alternatives.

Now suppose F satisfies unlimited domain, Pareto indifference and binary independence of irrelevant alternatives. By Theorem 2, F satisfies strong neutrality. We complete the proof by constructing the social-evaluation ordering R. For all $u, v \in \mathcal{R}^n$, let uRv if and only if there exist a profile $\Upsilon \in \mathcal{D}$ and two alternatives $x, y \in X$ such that $U(x) = u$, $U(y) = v$ and $xR_\Upsilon y$. Strong neutrality implies that non-welfare information is irrelevant and that the relative ranking of any two utility vectors u and v does not depend on the profile Υ or on the alternatives x and y used to generate u and v. Therefore, R is well-defined. That R is reflexive and complete follows immediately because R_Υ is reflexive and complete for all $\Upsilon \in \mathcal{D}$. It remains to show that R is transitive. Suppose $u, v, q \in \mathcal{R}^n$ are such that uRv and vRq. By unlimited domain and the assumption that X contains at least three alternatives, there exist a profile $\Upsilon \in \mathcal{D}$ and three alternatives $x, y, z \in X$ such that $U(x) = u$, $U(y) = v$ and $U(z) = q$. Because $U(x)RU(y)$ and $U(y)RU(z)$, it follows that $xR_\Upsilon y$ and $yR_\Upsilon z$ by definition of R. Because R_Υ is transitive, we have $xR_\Upsilon z$. Hence, $U(x)RU(z)$ or, equivalently, uRq. ■

Theorem 3 implies that, for all $x, y \in X$ and for all $\Upsilon \in \mathcal{D}$, $xP_\Upsilon y$ if and only if $U(x)PU(y)$ and $xI_\Upsilon y$ if and only if $U(x)IU(y)$.

The unlimited-domain axiom is crucial for this result. There are some non-welfare-information domains that permit non-welfarist social evaluation. This occurs because, on those domains, the constructions used in the proofs are not possible. Consider again the fat-thin example of Section 3 with four alternatives, two individuals, $\mathcal{S}_1 = \{f, t\}$ and $\mathcal{S}_2 = \{t\}$. Suppose that, instead of the unlimited domain, the domain of the social-evaluation functional is $\mathcal{U} \times \{\bar{K}, \hat{K}\}$, where $\bar{K}$ assigns (f, t, t, f) to the four alternatives for person 1 and (t, t, t, t) for person 2 and $\hat{K}$ assigns (t, f, f, t) to the four alternatives for person 1 and, again, (t, t, t, t) for person 2. This means that non-welfare information for the four alternatives is (f, t), (t, t), (t, t) and (f, t) in $\bar{K}$ and (t, t), (f, t), (f, t) and (t, t) in $\hat{K}$. Note that there is no pair of alternatives with the same non-welfare information. Consequently, binary independence does not apply.

Now consider the following social-evaluation functional. For all $x, y \in X$ and for all $U \in \mathcal{U}$,

$$xR_{(U,\bar{K})}y \Leftrightarrow U_1(x) + U_2(x) \geq U_1(y) + U_2(y)$$

and

$$xR_{(U,\hat{K})}y \Leftrightarrow U_1^r(x) + 3U_2^r(x) \geq U_1^r(y) + 3U_2^r(y)$$

where $U^r(x)$ and $U^r(y)$ are rank-ordered permutations of $U(x)$ and $U(y)$ such that $U_1^r(x) \geq U_2^r(x)$ and $U_1^r(y) \geq U_2^r(y)$. Thus, for all utility profiles, alternatives are ranked with utilitarianism when the non-welfare profile is $\bar{K}$ and with the Gini social-evaluation ordering when the non-welfare profile is $\hat{K}$. All of our axioms (except for unlimited domain) are satisfied but the principle is not welfarist because there is no profile-independent (single) ordering of utility vectors that can be used to rank the alternatives.

5. Anonymity

A principle for social evaluation may be welfarist and, at the same time, fail to be impartial. That would be the case, for example, if a weighted sum of utilities were used to rank alternatives with a weight of 2 for the utility of person 1 and a weight of 1 for all other utilities. If there is a single non-welfare profile, such a principle might be justified by the fact that person 1 is hardworking in every alternative.

In the single-non-welfare-profile environment, the anonymity axiom that is commonly used requires the social ordering to be unchanged if utility functions are permuted across individuals (see Sen [1970]). Although this produces the desired result, the permutation of utility functions does not change non-welfare information and, as a consequence, the case for anonymous welfarism is not convincing.

We employ a more compelling anonymity axiom. It requires the social ordering to be unchanged if *both* utility functions and individual non-welfare-information functions are permuted across individuals.

Anonymity: For all $\Upsilon, \bar{\Upsilon} \in \mathcal{D}$, if $K_0 = \bar{K}_0$ and there exists a bijection $\rho\colon N \to N$ such that $U_i = \bar{U}_{\rho(i)}$ and $K_i = \bar{K}_{\rho(i)}$ for all $i \in N$, then $R_\Upsilon = R_{\bar{\Upsilon}}$.

Anonymity is easily defended because it allows non-welfare information to matter. All that is ruled out is the claim that an individual's identity justifies special treatment, no matter what non-welfare information obtains.

An ordering R on $\mathcal{R}^n$ is anonymous if and only if, for all $u \in \mathcal{R}^n$ and for all bijections $\rho\colon N \to N$,

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

Together with unlimited domain, Pareto indifference and binary independence of irrelevant alternatives, anonymity is sufficient to ensure that the social-evaluation functional is welfarist and anonymous. To use anonymity, the permuted profiles of the axiom statement must be in the domain of the social-evaluation functional. Thus, the result of this section requires the additional assumption that the sets $\mathcal{S}_i$ are identical for all $i \in N$.

Theorem 4: *Suppose* $\mathcal{S}_i = \mathcal{S}_j$ *for all* $i, j \in N$ *and* F *satisfies unlimited domain.* F *satisfies Pareto indifference, binary independence of irrelevant alternatives and anonymity if and only if there exists an anonymous social-evaluation ordering* R *on* $\mathcal{R}^n$ *such that, for all* $x, y \in X$ *and for all* $\Upsilon \in \mathcal{D}$,

$$xR_\Upsilon y \Leftrightarrow U(x)RU(y). \tag{3}$$

Proof. Clearly, the existence of an anonymous social-evaluation ordering R such that (3) is satisfied for all $x, y \in X$ and for all $\Upsilon \in \mathcal{D}$ implies that F satisfies the required axioms.

Conversely, suppose $\mathcal{S}_i = \mathcal{S}_j$ for all $i, j \in N$ and F satisfies unlimited domain, Pareto indifference, binary independence of irrelevant alternatives and anonymity. By Theorem 3, there exists a social-evaluation ordering R on $\mathcal{R}^n$ such that (3) is satisfied for all $x, y \in X$ and for all $\Upsilon \in \mathcal{D}$. It remains to show that R must be anonymous.

For $j, k \in N$ with $j \neq k$, define the transposition bijection $\bar{\rho}_{jk}: N \to N$ by $\bar{\rho}_{jk}(j) = k$, $\bar{\rho}_{jk}(k) = j$ and $\bar{\rho}_{jk}(i) = i$ for all $i \in N \setminus \{j, k\}$. For $u \in \mathcal{R}^n$ and $j, k \in N$ with $j \neq k$, let $\bar{u}^{jk} = (u_{\bar{\rho}_{jk}(1)}, \ldots, u_{\bar{\rho}_{jk}(n)})$. By unlimited domain, there exist $\Upsilon \in \mathcal{D}$ and $x, y \in X$ such that $U(x) = u$ and $U(y) = \bar{u}^{jk}$. Let $\bar{\Upsilon}^{jk} = ((U_{\bar{\rho}_{jk}(1)}, \ldots, U_{\bar{\rho}_{jk}(n)}), (K_0, K_{\bar{\rho}_{jk}(1)}, \ldots, K_{\bar{\rho}_{jk}(n)}))$. By anonymity, $R_\Upsilon = R_{\bar{\Upsilon}^{jk}}$.

Because $U(x) = \bar{U}^{jk}(y) = u$ and $U(y) = \bar{U}^{jk}(x) = \bar{u}^{jk}$, we have

$$uR\bar{u}^{jk} \iff xR_\Upsilon y \iff yR_{\bar{\Upsilon}^{jk}}x \tag{4}$$

and

$$\bar{u}^{jk}Ru \iff yR_\Upsilon x \iff xR_{\bar{\Upsilon}^{jk}}y. \tag{5}$$

Because $R_\Upsilon = R_{\bar{\Upsilon}^{jk}}$, (4) and (5) together imply

$$uR\bar{u}^{jk} \iff \bar{u}^{jk}Ru$$

and, because R is complete, both $uR\bar{u}^{jk}$ and $\bar{u}^{jk}Ru$ are true, so $uI\bar{u}^{jk}$.

Now let $v = (u_{\rho(1)}, \ldots, u_{\rho(n)})$ for any bijection $\rho: N \to N$. Then there exist a finite number of transposition bijections such that ρ is the composition of those bijections. By repeated application of the above argument, uIv. ■

The anonymity axiom used in this section is not the only possible one. An alternate axiom applies to each profile separately. If the associated utility and individual non-welfare-information vectors for any one alternative are the same permutation of the corresponding vectors for a second, the axiom requires the two alternatives to be ranked as equally good. Neither it nor anonymity requires non-welfare information to be ignored and, in the presence of our other axioms, both imply anonymous welfarism.

6. Conclusion

Variants of the welfarism theorem can be proved on several different domains. If the domain consists of a single profile, the theorem requires Pareto indifference only: unlimited domain and binary independence of irrelevant alternatives are not needed because there is only one profile (Blackorby, Donaldson and Weymark [1990]). And, as is well known, the theorem is true with multiple welfare profiles and a single non-welfare-information profile.

In both these cases, it would be wrong to conclude that non-welfare information is irrelevant. In the single-profile case, if all of the utility vectors are distinct, Pareto indifference imposes no restriction and it might be true that the principle uses only non-welfare information in ranking alternatives. This is consistent with the formal definition of welfarism; a single ordering of utility vectors exists and it can be used to order the elements of X. A similar observation can be made in the single-non-welfare-profile case.

When the domain of the social-evaluation functional consists of multiple profiles of welfare and non-welfare information, no such ambiguity exists. As Theorem 1 indicates, any principle with an unlimited domain that satisfies Pareto indifference and binary independence of irrelevant alternatives must ignore non-welfare information. Our version of the welfarism theorem is, therefore, more powerful in this sense.

On a multi-profile domain, the welfarism theorem implies that any principle for social evaluation with an unlimited domain that uses non-welfare information must fail to satisfy Pareto indifference or binary independence of irrelevant alternatives. If it does not satisfy independence, it must be inconsistent across profiles. Because independence applies only to pairs of profiles for which welfare *and* non-welfare information coincide on a pair of alternatives, such inconsistency is not easily defended. On the other hand, if it does not satisfy Pareto indifference, it must also fail to satisfy the basic requirement that a social improvement should be an individual improvement for at least one person. Such principles can have little ethical appeal as long as the account of well-being that is employed is a comprehensive one.

An anonymity axiom that is weaker than the standard one requires the social ordering to be unaffected if both individual utility and non-utility-information functions are permuted across individuals. To ensure that permuted profiles are well-defined, the ranges of the individual non-welfare-information functions must be identical and, in that case, anonymity and our other conditions imply that the social-evaluation ordering must be anonymous: it ranks all permutations of any utility vector as equally good.

REFERENCES

Bentham, J., 1973 (originally published in 1789), Principles of morals and legislation, in *The Utilitarians,* Anchor, Garden City, 5–398.

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., D. Donaldson and J. Weymark, 1990, A welfarist proof of Arrow's theorem, *Recherches Economique de Louvain* **56**, 259–286.

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

Bordes, G., P. Hammond and M. Le Breton, 1997, Social welfare functionals on restricted domains and in economic environments, unpublished manuscript, Stanford University, Department of Economics.

Bossert, W. and J. Weymark, 2003, Utility in social choice, in *Handbook of Utility Theory, Vol. 2: Extensions,* S. Barberà, P. Hammond and C. Seidl, eds., Kluwer, Dordrecht, forthcoming.

Broome, J., 1991, *Weighing Goods,* Basil Blackwell, Oxford.

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

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., Blackwell, Oxford, 241–248.

Griffin, J., 1986, *Well-Being: Its Meaning, Measurement, and Moral Importance,* Clarendon, Oxford.

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

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

Kelsey, D., 1987, The role of information in social welfare judgements, *Oxford Economic Papers* **39**, 301–317.

Mongin, P. and C. d'Aspremont, 1998, Utility theory in ethics, in *Handbook of Utility Theory, Vol. 1: Principles,* S. Barberà, P. Hammond and C. Seidl, eds., Kluwer, Dordrecht, 371–481.

Nussbaum, M., 2000a, *Women and Human Development: The Capabilities Approach,* Cambridge University Press, Cambridge.

Nussbaum, M., 2000b, The central human capabilities, *Ethics* **111**, 6–7.

Roberts, K., 1980, 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., 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., 1985, *Commodities and Capabilities,* Elsevier, Amsterdam.

Sen, A., 1987, *The Standard of Living,* Cambridge University Press, Cambridge.

Sumner, L., 1996, *Welfare, Happiness, and Ethics,* Clarendon, Oxford.

Weymark, J., 1998, Welfarism on economic domains, *Mathematical Social Sciences* **36**, 251–268.

◇ ◇ ◇