Utilitarianism and the Theory of Justice*

by

Charles Blackorby, Walter Bossert and David Donaldson

August 1999

Prepared as Chapter 20 of the
Handbook of Social Choice and Welfare
K. Arrow, A. Sen and K. Suzumura, eds., Elsevier, Amsterdam

Charles Blackorby: University of British Columbia and GREQAM
Walter Bossert: University of Nottingham
David Donaldson: University of British Columbia

August 10, 1999

Abstract

This paper, which is to be published as a chapter in the Handbook of Social Choice and Welfare, provides a survey of Utilitarianism as a theory of justice. We review and discuss axiomatizations of Utilitarian and Generalized-Utilitarian social-evaluation functionals in a welfarist framework. In addition, we analyze extensions of Utilitarian principles to variable-population environments and to situations in which the alternatives resulting from choices among feasible actions are not known with certainty. *Journal of Economic Literature* Classification Numbers: D63, D71.

Keywords: Social Choice, Utilitarianism, Welfarism.

1. Introduction

In *A Theory of Justice,* Rawls (1971) describes justice as "the first virtue of social institutions" (p. 3) and identifies "the primary subject of justice" as "the basic structure of society, or more exactly, the way in which the major social institutions distribute fundamental rights and duties and determine the division of advantages from social cooperation" (p. 7).[1] The view of justice investigated in this chapter asserts that a just society is a good society: good for the individual people that comprise it. To implement such an approach to justice, the social good is identified and used to rank social alternatives. Of the alternatives that are feasible, given the constraints of human nature and history, the best is identified with justice. Even if the best alternative is not chosen, however, better ones are considered to be more just than worse ones. If societies are not perfectly just, therefore, social improvements can be recognized.

Social choices are not made in isolation, however. Decisions made in a particular society affect people in other parts of the world and people who are not yet born. In addition, both the number and identities of future people are influenced by choices made in the present. For that reason, principles that identify the social good are typically extended to rank complete histories of the world (or the universe if necessary) from remote past to distant future.[2] The principle that asserts that a just society is a good society must be qualified, therefore, with a ceteris paribus clause.

In this chapter, we investigate a particular conception of the social good, one that is based exclusively on individual good or well-being. Principles that reflect this view are called welfarist (Sen, 1979) and they treat values such as freedom and individual autonomy as 'instrumental'—valuable only 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) or Sumner (1996). We therefore focus on lifetime well-being and include enjoyment, pleasure and the absence of pain, good health, length of life, autonomy, liberty, understanding, accomplishment, and good human relationships as aspects of it.

Popularized by Bentham (1789, 1973), Utilitarianism is a welfarist principle that can be used to rank social alternatives according to their goodness.[3] Utility is an index of

---

[1] See Murphy (1998) for a discussion of the relationship between the principles of social justice and the principles that guide individual conduct.

[2] Because some non-human animals are sentient—capable of having experiences—their interests are often included. Sidgwick (1907, 1966, p. 414) argues that we should "extend our concern to all the beings capable of pleasure and pain whose feelings are affected by our conduct". Throughout the chapter, however, we assume that only human well-being counts in social evaluation, a simplification that makes our presentation simpler. Readers who are interested in the extension of welfarist principles to non-human sentient creatures are referred to Blackorby and Donaldson (1992).

[3] See also Mill (1861, 1969) and Sidgwick (1907, 1966) for other early formulations of Utilitarianism.

individual lifetime well-being and, for a fixed population, Utilitarianism declares alternative x to be better than alternative y if and only if total utility is greater in x than in y. If total utility if fixed, this principle is unconcerned with its distribution, but it does not follow that it is unconcerned with income inequality or social provision for special needs.

A family of principles whose value functions have the same additively separable mathematical structure as the Utilitarian value function is the Generalized-Utilitarian family of principles, which includes Utilitarianism as a special case. Each of these principles employs transformed utilities and some exhibit aversion to utility inequality. The Generalized-Utilitarian principles satisfy an important property: if a social change affects the utilities of a particular group of individuals only, the ranking of such changes is independent of the utility levels of others. This means that independent subprinciples exist for subgroups (including generations) and are consistent with the overall principle. Both the Utilitarian and Generalized-Utilitarian families can be extended to cover changes in population size and composition.

A social-evaluation functional assigns a ranking of all possible social alternatives to every admissible utility profile where a profile contains one utility function for each member of society. Arrow's (1951, 1963) seminal contribution to social-choice theory employs a domain that uses individual preference information only. His impossibility result can be avoided if the domain is changed to include profiles of utility functions together with conditions that ensure that utilities are, to some degree, interpersonally comparable.[4] Although the principles discussed in this chapter all require a certain amount of interpersonal comparability, numerically meaningful utility functions are not necessary for most of them. For example, Utilitarianism requires only cardinal measurability of individual utilities—with utility functions that are unique up to increasing affine transformations— and interpersonal comparability of utility gains and losses between pairs of alternatives (cardinal measurability and unit comparability). In a two-person society, if one person gains more in moving from y to x than the other loses, Utilitarianism declares x to be socially better than y.

In this chapter, we survey a set of results in social-choice theory that provides an axiomatic basis for Generalized Utilitarianism and, in some cases, Utilitarianism itself. Although we are concerned, for the most part, with comparisons of alternatives according to their goodness, we include a brief discussion of rankings of actions. If the elements of a set of feasible actions for an agent (a person or a government) lead with certainty to particular social alternatives, the actions can be ranked with a welfarist principle. If, however, the consequences of actions are uncertain, ranking them is more difficult. One way of doing it is to attach probabilities (possibly subjective) to a set of 'states of nature' and use them to rank prospects: lists of corresponding alternatives.

---

[4] See Sen (1970).

Section 2 introduces social-evaluation functionals. They make use of some or all available utility and non-utility information to rank alternatives according to their goodness. In addition, the section provides a formal account of information environments and information-invariance conditions. Information environments are described by partitions of profiles of utility functions into equivalence classes. They determine which statements involving both intrapersonal and interpersonal utility comparisons are meaningful. Information-invariance conditions specify the amount of utility information that a social-evaluation functional may make use of.[5]

Section 3 turns to welfarism and it is shown that, given an unlimited domain of utility profiles, welfarism is characterized by the axioms binary independence of irrelevant alternatives and Pareto indifference. Section 4 contains a set of results that characterize Generalized Utilitarianism. Section 5 focuses on Utilitarianism itself and presents characterization theorems for it. The axioms employed in Sections 4 and 5 are of two types. Some are information assumptions and they require the social-evaluation functional to make use of meaningful utility information only. Other axioms, such as strong Pareto and anonymity, are ethical in nature. Anonymity, for example, captures the idea of impartiality in social evaluation, an essential feature of many welfarist principles.

Population issues are the concern of Section 6 which discusses extensions of Utilitarianism and Generalized Utilitarianism to environments in which alternatives may differ in population size and composition. The axioms used in that section strengthen the case for the Utilitarian and Generalized-Utilitarian principles and, in addition, characterize families of principles that extend the fixed-population principles in an ethically attractive way. These families are known as Critical-Level Utilitarianism and Critical-Level Generalized Utilitarianism. They require the specification of a fixed 'critical level' of lifetime utility above which additions to a population are, ceteris paribus, valuable.

Section 7 makes use of subjective probabilities to rank prospects. We present a multi-profile version of Harsanyi's (1955, 1977) social-aggregation theorem. Instead of using lotteries as social alternatives, we employ prospects and assume that probabilities are fixed and common to all individuals and the social evaluator. Both individual ex-ante utilities and social preferences are assumed to satisfy the expected-utility hypothesis (von Neumann and Morgenstern, 1944) and individual ex-ante utilities are equal to the expected value of von Neumann – Morgenstern utilities (this is called the Bernoulli hypothesis by Broome, 1991a). Given that, we show that any welfarist ex-ante social-evaluation functional satisfying anonymity and the weak Pareto principle is Utilitarian. Harsanyi (1953) presents another argument for Utilitarianism in his impartial-observer theorem.

_____

[5] For a more complete guide to these requirements, see Chapter 19 of this volume (d'Aspremont and Gevers, 1999).

Because it is discussed in detail in Chapter 21 of this volume (Mongin, 1999), it is omitted in this survey.

Welfarist social evaluation is an attractive option but it is not the only one. Many people who reject welfarism, however, do not believe that welfare considerations are completely irrelevant to social decision-making. Accordingly, the results of this chapter should be of interest to most people who are concerned with social evaluation. Other theories, such as Sen's (1985) treatment of functionings and capabilities, can be modified to fit our framework. Functionings are the things that people can do, and the idea can be used to provide an account of (some aspects of) well-being. Capabilities are opportunities and include freedoms. If the two are aggregated into a single index of 'advantage' for each person, welfarist social-evaluation functionals can be used to rank alternatives. On the other hand, it is possible to use welfarist principles to aggregate functionings into an index of social functioning, which would be only one factor in overall social evaluation.

A major challenge to welfarism has appeared in recent years. It replaces concern for well-being with concern for *opportunities* for well-being on the grounds that individual people are responsible for their choices (in certain circumstances they may be thought to be responsible for their preferences as well).[6] In practice, welfarists often agree that the provision of opportunities is socially warranted, but their concern is with actual well-being. If autonomy is a significant aspect of well-being, people must be free to make important choices for themselves, and this provides a constraint which restricts the feasible set of social possibilities. By way of analogy, parents typically provide opportunities to their children, but that does not mean that opportunities are what they care about.

2. Social-Evaluation Functionals

Social-evaluation functionals use information about the members of a set of social alternatives to rank them according to their social goodness. Let X be a set of alternatives that contains at least three members (some slightly stronger requirements on the minimal number of alternatives are employed in Sections 6 and 7). No other restrictions are imposed on X: it may be finite, countably infinite, or uncountable.

The set of individuals in an n-person society is $\{1, \ldots, n\}$ where $n \in \mathcal{Z}_{++}$.[7] Except for our discussion in Section 6, we consider only comparisons of alternatives with the same population. For any $i \in \{1, \ldots, n\}$, $U_i : X \longrightarrow \mathcal{R}$ is i's utility function[8] and $u_i = U_i(x)$ is the utility level of individual $i \in \{1, \ldots, n\}$ in alternative $x \in X$. Utilities are interpreted

---

[6] See, for example, Arneson (1989), Roemer (1996) and, for an unsympathetic critique, Anderson (1999).

[7] $\mathcal{Z}_{++}$ is the set of positive integers and $\mathcal{Z}_+$ is the set of nonnegative integers.

[8] $\mathcal{R}$, $\mathcal{R}_+$ and $\mathcal{R}_{++}$ are the sets of all real numbers, nonnegative real numbers and positive real numbers respectively. In addition, $\mathbf{1}_n = (1, \ldots, 1) \in \mathcal{R}^n$.

as indicators of lifetime well-being and measure how good a person's life is from his or her own point of view. This does not mean that the utility function U_i is a representation of person i's actual preferences. Preferences and utility functions may be inconsistent because of individual non-rationalities, altruism or insufficient information.[9]

A profile of utility functions is an n-tuple $U = (U_1, \ldots, U_n)$ with one utility function for each individual in society. The set of all possible profiles is denoted by $\mathcal{U}$. For $U \in \mathcal{U}$ and $x \in X$, we write $U(x) = (U_1(x), \ldots, U_n(x))$. This vector represents the welfare information for alternative x given the profile U.

In addition to welfare information, non-welfare information may be available, and we assume that each x in the set X contains a full description of all non-welfare aspects of alternatives that may be considered relevant for social evaluation.

A social-evaluation functional is a mapping $F \colon \mathcal{D} \longrightarrow \mathcal{O}$, where $\emptyset \neq \mathcal{D} \subseteq \mathcal{U}$ and $\mathcal{O}$ is the set of all orderings on X.[10] $\mathcal{D}$ is the domain of admissible utility profiles and it may consist of a single profile or many profiles. In the latter case, the social-evaluation functional can cope with different profiles of utility functions and inter-profile consistency conditions such as binary independence of irrelevant alternatives or various information-invariance conditions (see below) may be imposed on it. The social-evaluation functional may make use of non-welfare information in addition to welfare information contained in the profile U. For simplicity of notation, we write $R_U = F(U)$; I_U and P_U are the symmetric and asymmetric components of R_U. For any $x, y \in X$, $x R_U y$ means that x is socially at least as good as y, $x I_U y$ means that x and y are equally good, and $x P_U y$ means that x is socially better than y.

General social-evaluation functionals may make use of both welfare and non-welfare information, but welfarist functionals ignore non-welfare information and completely non-welfarist functionals ignore welfare information, making use of non-welfare information only. In the latter case, a single ordering of the alternatives in X is produced because only a single set of non-welfare information is available.

In a multi-profile environment, it is possible to restrict the welfare information that the social-evaluation functional F may make use of. This is done by partitioning $\mathcal{D}$ into subsets of informationally equivalent profiles called information sets. Usable information in a profile in $\mathcal{D}$ is that which all informationally equivalent profiles in the corresponding information set have in common. If utilities are ordinally measurable and interpersonally noncomparable, for example, two profiles $U, V \in \mathcal{D}$ are informationally equivalent if and only if there exist increasing functions $\phi_1, \ldots, \phi_n$ with $\phi_i \colon \mathcal{R} \longrightarrow \mathcal{R}$ for all $i \in \{1, \ldots, n\}$

---

[9] See Broome (1991a) and Mongin and d'Aspremont (1998) for discussions of individual well-being and its relationship to preferences, information and self-interest. Hammond (1999) offers a very different account, interpreting "individual welfare as a purely ethical concept".

[10] An ordering is a reflexive, transitive, and complete binary relation. Social-evaluation functionals are also referred to as social-welfare functionals and were introduced by Sen (1970).

such that $(V_1(x), \ldots, V_n(x)) = (\phi_1(U_1(x)), \ldots, \phi_n(U_n(x)))$ for all $x \in X$. The utility comparison $U_1(x) > U_1(y)$ is meaningful in such an environment because it is true in all informationally equivalent profiles or false in all of them. On the other hand, if $\mathcal{D} = \mathcal{U}$, the interpersonal comparison $U_1(x) > U_2(x)$ is not meaningful because it is true in some informationally equivalent profiles and false in others.

Restrictions on social-evaluation functionals imposed by the available information regarding the measurability and interpersonal comparability of individual utilities can be summarized using information-invariance conditions. In order to represent the informational environment, the set of admissible profiles is partitioned into information sets and an information-invariance condition requires F to be constant on each of them. That is, if two profiles U and V are informationally equivalent, R_U and R_V are identical.

A partition of $\mathcal{D}$ into information sets can be defined using an equivalence relation $\sim$ on $\mathcal{D}$ with an information set given by an equivalence class of $\sim$.[11] That is, for $U, V \in \mathcal{D}$, $U \sim V$ if and only if U and V are informationally equivalent. A social-evaluation functional F satisfies information invariance with respect to the information environment described by $\sim$ if and only if it assigns the same social ordering to all profiles in an equivalence class of $\sim$.

Information Invariance with Respect to $\sim$: *For all $U, V \in \mathcal{D}$, if $U \sim V$, then $R_U = R_V$.*

The most commonly used approach to formalizing various types of informational assumptions identifies the equivalence relation $\sim$ by specifying a set of admissible transformations of utility profiles that lead to informationally equivalent profiles.[12] An invariance transformation is a vector $\phi = (\phi_1, \ldots, \phi_n)$ of functions $\phi_i \colon \mathcal{R} \longrightarrow \mathcal{R}$ for all $i \in \{1, \ldots, n\}$ whose application to a profile U results in an informationally equivalent profile. Let Φ denote the set of invariance transformations used to generate the equivalence relation $\sim$. That is, for all $U, V \in \mathcal{D}$, $U \sim V$ if and only if there exists $\phi \in \Phi$ such that $V = \phi \circ U$, where $\circ$ denotes component-by-component function composition.[13]

Various information assumptions that can be expressed in terms of admissible transformations have been considered in contributions by Blackorby and Donaldson (1982), Blackorby, Donaldson and Weymark (1984), d'Aspremont and Gevers (1977), De Meyer and Plott (1971), Dixit (1980), Gevers (1979), Roberts (1980b), and Sen (1970, 1974,

---

[11] An equivalence relation is a reflexive, transitive and symmetric binary relation.

[12] This approach was developed in contributions such as d'Aspremont and Gevers (1977), Roberts (1980a,b) and Sen (1974). See Basu (1983), Bossert (1991, 1999), Bossert and Stehling (1992, 1994), Falmagne (1981), Fishburn, Marcus-Roberts and Roberts (1988), Fishburn and Roberts (1989) and Krantz, Luce, Suppes and Tversky (1971) for discussions of information-invariance assumptions in terms of meaningful statements and their relations to uniqueness properties of measurement scales.

[13] We only consider sets Φ of invariance transformations such that the resulting relation $\sim$ is an equivalence relation. See Bossert and Weymark (1999) for a discussion and for conditions guaranteeing this.

1977a, 1986) among others. We restrict attention to the information assumptions that are relevant for the purposes of this chapter and refer the interested reader to Bossert and Weymark (1999) or d'Aspremont and Gevers (1999) for more detailed treatments. Each of these assumptions is defined by specifying the set of invariance transformations Φ which induces the equivalence relation $\sim$ that partitions $\mathcal{D}$ into sets of informationally equivalent utility profiles. For each information assumption listed below, we implicitly assume that the domain $\mathcal{D}$ contains all profiles that are informationally equivalent to U for each $U \in \mathcal{D}$. This is true, in particular, for the unrestricted domain $\mathcal{D} = \mathcal{U}$ on which we focus for most of the chapter. Additional information assumptions are introduced in Section 6 in a variable-population context.

If utilities are cardinally measurable, individual utility functions are unique up to increasing affine transformations, thereby allowing for intrapersonal comparisons of utility differences. If, in addition, some comparisons of utility are meaningful interpersonally, these transformations must be restricted across individuals. An example is cardinal unit comparability. In that information environment, admissible transformations are increasing affine functions and, in addition, the scaling factor must be the same for all individuals. This information assumption allows for interpersonal comparisons of utility differences but utility levels cannot be compared interpersonally because the intercepts of the the affine transformations may differ across individuals.

Cardinal Unit Comparability (CUC): $\phi \in \Phi$ *if and only if there exist* $a_1, \ldots, a_n \in \mathcal{R}$ *and* $b \in \mathcal{R}_{++}$ *such that* $\phi_i(\tau) = a_i + b\tau$ *for all* $\tau \in \mathcal{R}$ *and for all* $i \in \{1, \ldots, n\}$.

An information environment that provides more information than CUC is one in which the unit in which utilities are measured is numerically significant. In this case, we say that utilities are translation-scale measurable. Utility differences are interpersonally comparable and, in addition, their numerical values are meaningful. Because the functions $\phi_1, \ldots, \phi_n$ may be different for each person, utility levels are, again, not interpersonally comparable.

Translation-Scale Measurability (TSM): $\phi \in \Phi$ *if and only if there exist* $a_1, \ldots, a_n \in \mathcal{R}$ *such that* $\phi_i(\tau) = a_i + \tau$ *for all* $\tau \in \mathcal{R}$ *and for all* $i \in \{1, \ldots, n\}$.

If utilities are cardinally measurable and fully interpersonally comparable, both utility levels and differences can be compared interpersonally. In this case, utility functions are unique up to increasing affine transformations which are identical across individuals.

Cardinal Full Comparability (CFC): $\phi \in \Phi$ *if and only if there exist* $a \in \mathcal{R}$ *and* $b \in \mathcal{R}_{++}$ *such that* $\phi_i(\tau) = a + b\tau$ *for all* $\tau \in \mathcal{R}$ *and for all* $i \in \{1, \ldots, n\}$.

CFC defines a finer partition of the set of admissible utility profiles than CUC and, therefore, places weaker invariance requirements on the social-evaluation functional.

If all the information in a profile is meaningful, we say that utilities are numerically measurable and fully interpersonally comparable. In this case, each information set consists of a singleton.

Numerical Full Comparability (NFC): $\phi \in \Phi$ *if and only if* $\phi_i(\tau) = \tau$ *for all* $\tau \in \mathcal{R}$ *and for all* $i \in \{1, \ldots, n\}$.

In general, increases in available information reduce the restrictions on F implied by the information-invariance condition. For example, information invariance with respect to TSM is a weaker restriction than information invariance with respect to CUC, and invariance with respect to NFC provides no restriction at all.

3. Welfarism

The orderings on X generated by welfarist social-evaluation functionals compare any two alternatives $x, y \in X$ solely on the basis of the individual utilities experienced in x and in y. All non-welfare information is ignored when establishing the social ranking.

Welfarism is a consequence of three axioms, the first of which is an unrestricted-domain assumption. This axiom requires the social-evaluation functional F to be defined on the set of all possible utility profiles.

Unrestricted Domain (UD): $\mathcal{D} = \mathcal{U}$.

The next axiom is an independence condition which links the orderings associated with different profiles. It requires the social ranking of any two alternatives to be independent of the utility levels associated with other alternatives.

Binary Independence of Irrelevant Alternatives (BI): For all $x, y \in X$, for all $U, V \in \mathcal{D}$, if $U(x) = V(x)$ and $U(y) = V(y)$, then xR_Uy if and only if xR_Vy.

The above independence axiom for social-evaluation functionals is weaker than the corresponding independence axiom for social-welfare functions (see Arrow, 1951, 1963, and Sen, 1970). Arrow's independence axiom requires the social ordering of a pair of alternatives to depend only on the individual rankings of the two alternatives. BI is equivalent to Arrow's binary-independence axiom if the social-evaluation functional satisfies information invariance with respect to ordinally measurable, interpersonally noncomparable utilities. As formulated above, binary independence is compatible with any assumption concerning the measurability and interpersonal comparability of individual utilities.

The final axiom used to generate welfarism is Pareto indifference. If all individuals are equally well off in two alternatives, it requires the social-evaluation functional to rank them as equally good.

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

Pareto indifference is an attractive axiom if utility functions measure everything that is of value to individuals. For that reason, we endorse comprehensive accounts of lifetime utility such as the ones provided by Griffin (1986) and Sumner (1996).[14] Griffin includes enjoyment, pleasure and the absence of pain, good health, autonomy, liberty, understanding, accomplishment, and good human relationships as aspects of well-being. He argues, in addition, that there is a moral dimension to well-being. Sumner also discusses the role that individual attitudes can play.

In the presence of unrestricted domain, BI and PI together imply that non-welfare information about the alternatives must be ignored by the social-evaluation functional. If, in one profile, utility numbers for a pair of alternatives are equal to the utility numbers for another pair in a possibly different profile, the rankings of the two pairs must be the same. This property is called strong neutrality.

Strong Neutrality (SN): For all $x, y, z, w \in X$, for all $U, V \in \mathcal{D}$, if $U(x) = V(z)$ and $U(y) = V(w)$, then $x R_U y$ if and only if $z R_V w$.

We obtain the following theorem (see, for example, Blau, 1976, d'Aspremont and Gevers, 1977, Guha, 1972, and Sen, 1977a, for this and related results).

Theorem 1: *Suppose that a social-evaluation functional F satisfies UD. F satisfies BI and PI if and only if F satisfies SN.*

Proof. First, suppose that F satisfies UD and SN. That BI is satisfied follows immediately by setting $x = z$ and $y = w$ in the definition of SN. Setting $U = V$ and $y = z = w$, SN implies that $x R_U y$ if and only if $y R_U y$ when $U(x) = U(y)$. Because R_U is reflexive, this implies $x I_U y$, which demonstrates that PI is satisfied.

Now suppose that F satisfies UD, BI and PI. Suppose that $U(x) = V(z) = u$ and $U(y) = V(w) = v$. By UD, there exists an alternative $\bar{x} \in X$ and profiles $\bar{U}, \hat{U}, \tilde{U} \in \mathcal{D}$ such that $\bar{U}(x) = \bar{U}(\bar{x}) = u$ and $\bar{U}(y) = v$, $\hat{U}(z) = \hat{U}(\bar{x}) = u$ and $\hat{U}(w) = v$, and $\tilde{U}(\bar{x}) = u$ and $\tilde{U}(y) = \tilde{U}(w) = v$. By BI, $x R_U y$ if and only if $x R_{\bar{U}} y$. By PI and the transitivity of $R_{\bar{U}}$, it follows that $x R_{\bar{U}} y$ if and only if $\bar{x} R_{\bar{U}} y$. A similar argument implies that $\bar{x} R_{\bar{U}} y$ if and only if $\bar{x} R_{\tilde{U}} y$ if and only if $\bar{x} R_{\tilde{U}} w$. Applying the same argument once again, we have $\bar{x} R_{\tilde{U}} w$ if and only if $\bar{x} R_{\hat{U}} w$ if and only if $z R_{\hat{U}} w$. By BI, $z R_{\hat{U}} w$ if and only if $z R_V w$. Therefore, $x R_U y$ if and only if $z R_V w$ which proves that F satisfies SN. ∎

Given unrestricted domain, SN is equivalent to the existence of an ordering $\overset{*}{R}$ on $\mathcal{R}^n$ which can be used to rank the alternatives in X for any utility profile U. The social betterness (strict preference) relation and the equal-goodness (indifference) relation corresponding to $\overset{*}{R}$ are denoted by $\overset{*}{P}$ and $\overset{*}{I}$ respectively. We refer to $\overset{*}{R}$ as a social-evaluation

---

[14] See also Broome (1991a) and Mongin and d'Aspremont (1998) for accounts based on self-interested preferences under conditions of full information.

ordering.[15] Combined with Theorem 1, this observation yields the following welfarism theorem (see d'Aspremont and Gevers, 1977, and Hammond, 1979).[16]

Theorem 2: *Suppose that a social-evaluation functional F satisfies UD. F satisfies BI and PI if and only if there exists a social-evaluation ordering $\overset{*}{R}$ on $\mathcal{R}^n$ such that, for all $x, y \in X$ and for all $U \in \mathcal{D}$,*

$$x\, R_U\, y \iff U(x)\, \overset{*}{R}\, U(y). \tag{3.1}$$

Proof. If there exists a social-evaluation ordering $\overset{*}{R}$ such that, for all $x, y \in X$ and all $U \in \mathcal{D}$, (3.1) is satisfied, BI and PI are satisfied.

Now suppose that F satisfies BI and PI. By Theorem 1, F satisfies SN. Define the relation $\overset{*}{R}$ as follows. For all $u, v \in \mathcal{R}$, $u\overset{*}{R}v$ if and only if there exist a profile $U \in \mathcal{D}$ and two alternatives $x, y \in X$ such that $U(x) = u$, $U(y) = v$, and xR_Uy. By SN, the relative ranking of any two utility vectors u and v does not depend on the profile U or on the alternatives x and y used to generate u and v and, therefore, $\overset{*}{R}$ is well-defined. That $\overset{*}{R}$ is reflexive and complete follows immediately because R_U is reflexive and complete for all $U \in \mathcal{D}$. It remains to be shown that $\overset{*}{R}$ is transitive. Suppose $u, v, w \in \mathcal{R}^n$ are such that $u\overset{*}{R}v$ and $v\overset{*}{R}w$. By UD, there exists a profile $U \in \mathcal{D}$ and three alternatives $x, y, z \in X$ such that $U(x) = u$, $U(y) = v$, and $U(z) = w$. Because $U(x)\overset{*}{R}U(y)$ and $U(y)\overset{*}{R}U(z)$, it follows that xR_Uy and yR_Uz by definition of $\overset{*}{R}$. Transitivity of R_U then implies that xR_Uz. Hence, $U(x)\overset{*}{R}U(z)$ or, equivalently, $u\overset{*}{R}w$ which shows that $\overset{*}{R}$ is transitive. ■

Note that the social-evaluation ordering $\overset{*}{R}$ in the statement of Theorem 2 is profile-independent. Pairs of alternatives whose utility vectors are the same are ranked in the same way, regardless of the utility profile. If $\mathcal{D}$ consists of a single profile, the result is true but BI is not needed (Blackorby, Donaldson and Weymark, 1990).

For notational convenience, we concentrate on the social-evaluation ordering $\overset{*}{R}$ in most of the remainder of the chapter. All axioms and results regarding this ordering can be reformulated in terms of the social-evaluation functional F by defining the properties analogous to those defined for $\overset{*}{R}$ and adding the welfarism axioms.

Given a set of invariance transformations Φ (and, hence, an equivalence relation $\sim$) and the welfarism axioms UD, BI, and PI, the information-invariance axiom for the social-evaluation functional F is equivalent to an analogous condition formulated in terms of the corresponding social-evaluation ordering $\overset{*}{R}$.

---

[15] Gevers (1979) calls $\overset{*}{R}$ a social-welfare ordering.

[16] Bordes, Hammond and Le Breton (1997) and Weymark (1998) prove variants of this theorem with specific domain restrictions, that is, with weaker domain assumptions than UD.

Information Invariance with Respect to Φ**:** For all u, v, u', $v' \in \mathcal{R}^n$, if there exists $\phi \in \Phi$ such that $u' = \phi(u)$ and $v' = \phi(v)$, then $u \overset{*}{R} v$ if and only if $u' \overset{*}{R} v'$.

Next, we introduce some axioms that are commonly required in welfarist social evaluation.

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

Continuity (C): For all $u \in \mathcal{R}^n$, the sets $\{v \in \mathcal{R}^n \mid v \overset{*}{R} u\}$ and $\{v \in \mathcal{R}^n \mid u \overset{*}{R} v\}$ are closed in $\mathcal{R}^n$.

Anonymity ensures that the ordering $\overset{*}{R}$ treats individuals impartially, paying no attention to their identities. That is, any permutation of a given utility vector must be indifferent to the utility vector itself. Note that this is a strengthening of Arrow's (1951, 1963) condition that prevents the existence of a dictator.

Anonymity (A): For all $u \in \mathcal{R}^n$, for all bijective mappings $\pi \colon \{1, \ldots, n\} \longrightarrow \{1, \ldots, n\}$,

$$u \overset{*}{I} \left(u_{\pi(1)}, \ldots, u_{\pi(n)} \right). \tag{3.2}$$

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

Weak Pareto (WP): For all $u, v \in \mathcal{R}^n$, if $u_i > v_i$ for all $i \in \{1, \ldots, n\}$, then $u \overset{*}{P} v$.

A strengthening of both weak Pareto and Pareto indifference is the strong Pareto principle. In addition to Pareto indifference, it requires that if no one's utility has decreased and at least one person's utility has increased, the change is a social improvement.

Strong Pareto (SP): For all $u, v \in \mathcal{R}^n$, (i) if $u_i = v_i$ for all $i \in \{1, \ldots, n\}$, then $u \overset{*}{I} v$; and (ii) if $u_i \geq v_i$ for all $i \in \{1, \ldots, n\}$ with at least one strict inequality, then $u \overset{*}{P} v$.

Note that (i) in the definition of SP is redundant in a welfarist framework—this restriction is implied by the reflexivity of $\overset{*}{R}$. We have chosen to include it in the definition of strong Pareto order to follow the conventional terminology.

Finally, we introduce an axiom which prevents the social ordering from exhibiting a strong version of inequality preference.

Minimal Equity (ME): There exist $i, j \in \{1, \ldots, n\}$ and $u, v \in \mathcal{R}^n$ such that $u_k = v_k$ for all $k \in \{1, \ldots, n\} \setminus \{i, j\}$, $v_j > u_j > u_i > v_i$, and $u \overset{*}{R} v$.

See d'Aspremont (1985), d'Aspremont and Gevers (1977), and Deschamps and Gevers (1978) for this axiom and Hammond (1976) for a related condition.

Continuity and the weak Pareto principle ensure the existence of a continuous representation of the social-evaluation ordering $\overset{*}{R}$. We obtain

Theorem 3: *If a social-evaluation ordering $\overset{*}{R}$ satisfies C and WP, then, for each $u \in \mathcal{R}^n$, there exists a unique $\xi = \Xi(u) \in [\min\{u_1, \ldots, u_n\}, \max\{u_1, \ldots, u_n\}]$ such that $u \overset{*}{I} \xi \mathbf{1}_n$.*

Proof. If $n = 1$, WP and reflexivity imply that, for all $u, v \in \mathcal{R}$, $u \overset{*}{R} v$ if and only if $u \geq v$. Consequently, the result follows from letting $\xi = \Xi(u) = u$ (which is equal to both maximum and minimum utility) for all $u \in \mathcal{R}$.

Suppose $n \geq 2$, let $u \in \mathcal{R}^n$ be arbitrary and suppose, by way of contradiction, that $\min\{u_1, \ldots, u_n\}\mathbf{1}_n \overset{*}{P} u$. By C, there exists a neighborhood of u such that $\min\{u_1, \ldots, u_n\}\mathbf{1}_n$ is preferred to all points in that neighborhood according to $\overset{*}{R}$. Because this neighborhood contains points that strictly dominate u and, thus, $\min\{u_1, \ldots, u_n\}\mathbf{1}_n$, this contradicts WP. Therefore, $u \overset{*}{R} \min\{u_1, \ldots, u_n\}\mathbf{1}_n$. Analogously, it follows that $\max\{u_1, \ldots, u_n\}\mathbf{1}_n \overset{*}{R} u$. By C, it follows that there exists $\xi \in [\min\{u_1, \ldots, u_n\}, \max\{u_1, \ldots, u_n\}]$ such that $u \overset{*}{I} \xi \mathbf{1}_n$. WP implies that ξ must be unique for each u and thus can be written as a function $\Xi: \mathcal{R}^n \longrightarrow \mathcal{R}$. ∎

The representative utility ξ is analogous to the equally-distributed-equivalent income used in ethical approaches to income-inequality measurement.[17] The function Ξ is the representative-utility function corresponding to $\overset{*}{R}$, and it is easy to see that it is a representation of $\overset{*}{R}$—that is, for all $u, v \in \mathcal{R}^n$,

$$u \overset{*}{R} v \iff \Xi(u) \geq \Xi(v). \tag{3.3}$$

Furthermore, Ξ is continuous because $\overset{*}{R}$ is, and WP implies that Ξ is weakly increasing.

We conclude this section with some examples of welfarist social-evaluation orderings. The Utilitarian social-evaluation ordering uses the sum of the individual utilities to make social comparisons. According to Utilitarianism, for all $u, v \in \mathcal{R}^n$,

$$u \overset{*}{R} v \iff \sum_{i=1}^{n} u_i \geq \sum_{i=1}^{n} v_i. \tag{3.4}$$

The class of social-evaluation orderings that respect all strict rankings of utility vectors according to Utilitarianism is the class of Weakly Utilitarian orderings (see Deschamps and Gevers, 1978). $\overset{*}{R}$ is Weakly Utilitarian if and only if

$$\sum_{i=1}^{n} u_i > \sum_{i=1}^{n} v_i \implies u \overset{*}{P} v \tag{3.5}$$

for all $u, v \in \mathcal{R}^n$.

---

[17] See, for example, Atkinson (1970), Dalton (1920), Kolm (1969), and Sen (1973).

The members of the class of Generalized-Utilitarian orderings perform social comparisons by adding the transformed utilities of the members of society. If the transformation applied to individual utilities is (strictly) concave, the resulting ordering represents (strict) inequality aversion (to utility inequality). Formally, a social-evaluation ordering $\overset{*}{R}$ is Generalized Utilitarian if and only if there exists a continuous and increasing function $g \colon \mathcal{R} \longrightarrow \mathcal{R}$ such that, for all $u, v \in \mathcal{R}^n$,

$$u \overset{*}{R} v \iff \sum_{i=1}^{n} g(u_i) \geq \sum_{i=1}^{n} g(v_i). \tag{3.6}$$

All Generalized-Utilitarian social-evaluation orderings satisfy C, A, and SP (and thus WP). If g is concave, ME is satisfied as well.

Utilitarianism is a special case of Generalized Utilitarianism in which the transformation g is affine. An example of a class of Generalized-Utilitarian orderings is the class of Symmetric Global Means of Order r. $\overset{*}{R}$ is a Symmetric Global Mean of Order r if and only if there exist $\beta, r \in \mathcal{R}_{++}$ such that, for all $u, v \in \mathcal{R}^n$,

$$u \overset{*}{R} v \iff \sum_{\substack{i \in \{1,\ldots,n\}: \\ u_i \geq 0}} u_i^r - \beta \sum_{\substack{i \in \{1,\ldots,n\}: \\ u_i < 0}} |u_i|^r \geq \sum_{\substack{i \in \{1,\ldots,n\}: \\ v_i \geq 0}} v_i^r - \beta \sum_{\substack{i \in \{1,\ldots,n\}: \\ v_i < 0}} |v_i|^r. \tag{3.7}$$

In this case, $g(\tau) = \tau^r$ for all $\tau \geq 0$ and $g(\tau) = -\beta|\tau|^r$ for all $\tau < 0$. The special cases of (3.7) with $r = 1$ and $\beta > 1$ are of particular interest. They represent the only cases that exhibit (weak) inequality aversion in all of $\mathcal{R}^n$ (Blackorby and Donaldson, 1982). The resulting ordering is a modification of Utilitarianism such that negative utilities get a higher weight than positive utilities. This principle requires that the zero level of utility has some meaning. In the variable-population principles of Section 6, it is used to represent the value of a 'neutral' life. Above neutrality, an individual life, taken as a whole, is worth living from the viewpoint of the individual; below neutrality, it is not (see Section 6).

The Utilitarian social-evaluation ordering pays no attention to utility inequality and, in contrast, the Maximin ordering exhibits complete inequality aversion. It pays attention to the utility of the worst-off individual only. For $u \in \mathcal{R}^n$, let $(u_{(1)}, \ldots, u_{(n)})$ be a permutation of u such that $u_{(i)} \geq u_{(i+1)}$ for all $i \in \{1, \ldots, n-1\}$. For all $u, v \in \mathcal{R}^n$, the Maximin ordering requires

$$u \overset{*}{R} v \iff u_{(n)} \geq v_{(n)}. \tag{3.8}$$

Leximin is a variant of Maximin in which the utility vector u is socially preferred to the utility vector v if the worst-off individual in u is better off than the worst-off individual in v. In case of a tie, however, the two are not necessarily equally good. Instead, the utilities of the next-to-worst-off individuals are used to determine the social preference, and the procedure continues until either there is a strict preference or the two utility vectors are

permutations of each other, in which case they are declared equally good. Formally, the Leximin ordering is given by letting, for all $u, v \in \mathcal{R}^n$,

$$u \overset{*}{R} v \iff u \text{ is a permutation of } v \text{ or there exists a } j \in \{1, \ldots, n\} \text{ such that}$$

$$u_{(i)} = v_{(i)} \text{ for all } i > j \text{ and } u_{(j)} > v_{(j)}. \tag{3.9}$$

Maximin is continuous and violates the strong Pareto principle (but satisfies weak Pareto) and Leximin satisfies strong Pareto but not continuity. Both orderings satisfy A and ME.

The extremely equality-averse counterpart of Leximin is the Leximax ordering, which is defined by letting, for all $u, v \in \mathcal{R}^n$,

$$u \overset{*}{R} v \iff u \text{ is a permutation of } v \text{ or there exists a } j \in \{1, \ldots, n\} \text{ such that}$$

$$u_{(i)} = v_{(i)} \text{ for all } i < j \text{ and } u_{(j)} > v_{(j)}. \tag{3.10}$$

Leximax satisfies A and SP (and therefore WP) but violates C and ME.

The class of Single-Parameter Gini social-evaluation orderings provides another possibility for a generalization of Utilitarianism—their level sets are linear in rank-ordered subspaces of $\mathcal{R}^n$. A social-evaluation ordering $\overset{*}{R}$ is a Single-Parameter Gini ordering if and only if there exists a real number $\delta \geq 1$ such that, for all $u, v \in \mathcal{R}^n$,

$$u \overset{*}{R} v \iff \sum_{i=1}^{n} \left[i^\delta - (i-1)^\delta \right] u_{(i)} \geq \sum_{i=1}^{n} \left[i^\delta - (i-1)^\delta \right] v_{(i)}. \tag{3.11}$$

The Single-Parameter Ginis are special cases of the Generalized Ginis introduced by Weymark (1981) and discussed in Bossert (1990a) and Donaldson and Weymark (1980) in the context of ethical inequality measurement. For $\delta = 1$, we obtain Utilitarianism, and as δ approaches infinity, Maximin is obtained in the limit. The case $\delta = 2$ yields the social-evaluation ordering corresponding to the Gini index of inequality (Blackorby and Donaldson, 1978). All Single-Parameter Ginis satisfy C, A, SP, and ME.

4. Generalized Utilitarianism

A distinguishing feature of Generalized-Utilitarian social-evaluation orderings is that they possess an additively separable structure. This separability property is closely related to several plausible independence conditions regarding the influence of unconcerned individuals in establishing a social ordering, and those conditions can, together with some of our earlier axioms, be used to provide characterizations of Generalized Utilitarianism.

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 (IUUI): For all $M \subseteq \{1, \ldots, n\}$, for all $u, v, u', v' \in \mathcal{R}^n$, if $u_i = v_i$ and $u'_i = v'_i$ for all $i \in M$ and $u_j = u'_j$ and $v_j = v'_j$ for all $j \in \{1, \ldots, n\} \setminus M$, then

$$u \overset{*}{R} v \iff u' \overset{*}{R} v'. \tag{4.1}$$

In this definition, the individuals in M are the unconcerned individuals—they are equally well off in u and v and in u' and v'. IUUI requires the ranking of u and v to depend on the utilities of the concerned individuals—those not in M—only. In terms of a real-valued representation, this axiom is referred to as complete strict separability in Blackorby, Primont and Russell (1978). 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 Roberts (1980b).

In the case of two individuals, this axiom is implied by strong Pareto. Therefore, its use is typically restricted to societies with at least three individuals. In that case, we obtain the following characterization of Generalized Utilitarianism.

Theorem 4: *Suppose that $n \geq 3$. A social-evaluation ordering $\overset{*}{R}$ satisfies C, A, SP, and IUUI if and only if $\overset{*}{R}$ is a Generalized-Utilitarian social-evaluation ordering.*

Proof. Applying Debreu's (1954) representation theorem, continuity implies that there exists a continuous function $f \colon \mathcal{R}^n \longrightarrow \mathcal{R}$ such that, for all $u, v \in \mathcal{R}^n$,

$$u \overset{*}{R} v \iff f(u) \geq f(v). \tag{4.2}$$

By SP, f is increasing in all arguments, and A implies that f is symmetric.

IUUI requires that $\{1, \ldots, n\} \setminus M$ is separable from its complement M for any choice of $M \subseteq \{1, \ldots, n\}$. Gorman's (1968) theorem on overlapping separable sets of variables (see also Aczél, 1966, p. 312, and Blackorby, Primont, and Russell, 1978, p. 127) implies that f is additively separable. Therefore, there exist continuous and increasing functions $H \colon \mathcal{R} \longrightarrow \mathcal{R}$ and $g_i \colon \mathcal{R} \longrightarrow \mathcal{R}$ for all $i \in \{1, \ldots, n\}$ such that

$$f(u) = H\left(\sum_{i=1}^{n} g_i(u_i) \right) \tag{4.3}$$

for all $u \in \mathcal{R}^n$. Because f is symmetric, each g_i can be chosen to be independent of i, and we define $g = g_i$ for all $i \in \{1, \ldots, n\}$. Therefore, because f is a representation of $\overset{*}{R}$,

$$u \overset{*}{R} v \iff H\Big(\sum_{i=1}^{n} g(u_i)\Big) \geq H\Big(\sum_{i=1}^{n} g(v_i)\Big) \tag{4.4}$$
$$\iff \sum_{i=1}^{n} g(u_i) \geq \sum_{i=1}^{n} g(v_i)$$

for all $u, v \in \mathcal{R}^n$. ∎

See also Debreu (1960) and Fleming (1952) for variants of this theorem. Due to the presence of A, IUUI could be weakened by suitably restricting the possible sets of unconcerned individuals.

An alternative to independence of the utilities of unconcerned individuals is the population substitution principle. It considers a family of social-evaluation orderings $\{\overset{*}{R}^n\}_{n \in \mathcal{Z}_{++}}$, one for each population size in $\mathcal{Z}_{++}$. Given, for each $n \in \mathcal{Z}_{++}$, anonymity and the axioms guaranteeing the existence of the representative-utility functions $\{\Xi^n\}_{n \in \mathcal{Z}_{++}}$ (see Theorem 3), the population substitution principle requires that replacing the utilities of a subgroup of the population with the representative utility of that subgroup is a matter of indifference (see Blackorby and Donaldson, 1984) .

Population Substitution Principle (PSP): For all $n \geq 3$, for all $u \in \mathcal{R}^n$, for all $M \subseteq \{1, \ldots, n\}$,

$$u \overset{*}{I} \left(\Xi^{|M|}\big((u_i)_{i \in M}\big) \mathbf{1}_{|M|}, (u_j)_{j \in \{1,\ldots,n\} \setminus M} \right). \tag{4.5}$$

As is the case for IUUI, PSP implies that for all $n \geq 3$, the representation Ξ^n of $\overset{*}{R}^n$ must be additively separable and, therefore, it provides an alternative way of characterizing Generalized Utilitarianism in the presence of continuity, anonymity, and strong Pareto. The axiom also implies that additive separability also applies to the cases $n = 1, 2$ and that the transform in the representation of $\overset{*}{R}^n$ can be chosen to be the same for all $n \in \mathcal{Z}_{++}$. In the theorem statement, we say that $\{\overset{*}{R}^n\}_{n \in \mathcal{Z}_{++}}$ satisfies a fixed population axiom if and only if the axiom is satisfied for all $n \in \mathcal{Z}_{++}$. The proof of this result is analogous to the one of Theorem 4 and is thus omitted.

Theorem 5: *A family of social-evaluation orderings* $\{\overset{*}{R}{}^n\}_{n\in\mathcal{Z}_{++}}$ *satisfies C, A, SP, and PSP if and only if there exists a continuous and increasing function* $g\colon \mathcal{R} \longrightarrow \mathcal{R}$ *such that, for all* $n \in \mathcal{Z}_{++}$, $\overset{*}{R}{}^n$ *is a Generalized-Utilitarian social-evaluation ordering with*

$$u \, \overset{*}{R}{}^n \, v \iff \sum_{i=1}^n g(u_i) \geq \sum_{i=1}^n g(v_i) \tag{4.6}$$

for all $u, v \in \mathcal{R}^n$.

See Blackorby and Donaldson (1984) for a related result in a variable-population framework.

Generalized Utilitarianism can also be characterized in an intertemporal framework. In that case, a very weak and natural separability axiom can be employed. This condition, the variable-population version of which was introduced in Blackorby, Bossert and Donaldson (1995), requires social evaluations to be independent of the utilities of individuals whose lives are over in both of any two alternatives and who had the same birth dates, lengths of life, and lifetime utilities in both.

Consider a model where each alternative $x \in X$ contains (among other features that may be considered relevant for social evaluation) information about individual birth dates and lengths of life. We assume that no one can live longer than $L \in \mathcal{Z}_{++}$ periods (L may be arbitrarily large). For $i \in \{1, \ldots, n\}$ and $x \in X$, let $s_i = S_i(x) \in \mathcal{Z}_+$ be the period before individual i is born in alternative x, and let $l_i = L_i(x) \in \{1, \ldots, L\}$ be i's lifetime in x (in periods). Thus, individual i is alive in periods $s_i + 1$ to $s_i + l_i$, and $u_i = U_i(x)$ is i's lifetime utility in alternative x. Let $s = (s_1, \ldots, s_n)$, $l = (l_1, \ldots, l_n)$ and, as before, $u = (u_1, \ldots, u_n)$.

In order to extend our model to this intertemporal framework, instead of a social-evaluation ordering $\overset{*}{R}$, we employ an ordering $\overset{\circ}{R}$ on $\mathcal{A} = \mathcal{Z}_+^n \times \{1, \ldots, L\}^n \times \mathcal{R}^n$ and the objects to be ranked are vectors (s, l, u) of the birth dates, lifetimes, and lifetime utilities of everyone in society. It is straightforward to reformulate intertemporal versions of the axioms continuity, anonymity, and strong Pareto in this framework. We use the following definitions, each of which is stated for an arbitrary $n \in \mathcal{Z}_{++}$.

Intertemporal Continuity (IC): For all $(s, l, u) \in \mathcal{A}$, the sets $\{v \in \mathcal{R}^n \mid (s, l, v)\overset{\circ}{R}(s, l, u)\}$ and $\{v \in \mathcal{R}^n \mid (s, l, u)\overset{\circ}{R}(s, l, v)\}$ are closed in $\mathcal{R}^n$.

Intertemporal Anonymity (IA): For all $(s, l, u) \in \mathcal{A}$, for all bijective mappings $\pi\colon \{1, \ldots, n\} \longrightarrow \{1, \ldots, n\}$,

$$(s, l, u) \, \overset{\circ}{I} \left((s_{\pi(1)}, \ldots, s_{\pi(n)}), (l_{\pi(1)}, \ldots, l_{\pi(n)}), (u_{\pi(1)}, \ldots, u_{\pi(n)}) \right). \tag{4.7}$$

Intertemporal Strong Pareto (ISP): For all (s, l, u), $(r, k, v) \in \mathcal{A}$, (i) if $u_i = v_i$ for all $i \in \{1, \ldots, n\}$, then $(s, l, u) \overset{\circ}{I} (r, k, v)$; and (ii) if $u_i \geq v_i$ for all $i \in \{1, \ldots, n\}$ with at least one strict inequality, then $(s, l, u) \overset{\circ}{P} (r, k, v)$.

The axiom independence of the utilities of the dead requires that, in any period $t \in \mathcal{Z}_{++}$, the relative ranking of any two alternatives is independent of the utilities of those individuals whose lives are over in t and who had the same birth dates, lifetimes, and lifetime utilities in both alternatives. To define this axiom formally, we need more notation. Let, for all $(s, l, u) \in \mathcal{A}$ and all $t \in \mathcal{Z}_{++}$, $D_t(s, l, u) = \{i \in \{1, \ldots, n\} \mid s_i + l_i < t\}$ and $B_t(s, l, u) = \{i \in \{1, \ldots, n\} \mid s_i + 1 < t\}$. The individuals in $D_t(s, l, u)$ are those individuals whose lives are over before period t, and $B_t(s, l, u)$ contains the individuals who are born before t. We can now define our intertemporal independence condition.

Independence of the Utilities of the Dead (IUD): For all $(s, l, u), (r, k, v), (s', l', u')$, $(r', k', v') \in \mathcal{A}$, for all $t \in \mathcal{Z}_{++}$, if

$$
\begin{aligned}
B_t(s, l, u) = B_t(s', l', u') = B_t(r, k, v) = B_t(r', k', v') = \\
D_t(s, l, u) = D_t(s', l', u') = D_t(r, k, v) = D_t(r', k', v') = M_t,
\end{aligned}
\tag{4.8}
$$

$(s_i, l_i, u_i) = (r_i, k_i, v_i)$ and $(s'_i, l'_i, u'_i) = (r'_i, k'_i, v'_i)$ for all $i \in M_t$, and $(s_j, l_j, u_j) = (s'_j, l'_j, u'_j)$ and $(r_j, k_j, v_j) = (r'_j, k'_j, v'_j)$ for all $j \in \{1, \ldots, n\} \setminus M_t$, then

$$
(s, l, u) \overset{\circ}{R} (r, k, v) \iff (s', l', u') \overset{\circ}{R} (r', k', v').
\tag{4.9}
$$

IUD is a very weak separability condition because it applies to individuals whose lives are over only and not to all unconcerned individuals. However, when combined with the intertemporal version of the strong Pareto principle, this axiom has important consequences. In particular, independence of the utilities of the dead and intertemporal strong Pareto together imply an intertemporal version of independence of the utilities of unconcerned individuals, which is defined as follows.

Intertemporal Independence of the Utilities of Unconcerned Individuals (IIUUI): For all $M \subseteq \{1, \ldots, n\}$, for all $(s, l, u), (r, k, v), (s', l', u'), (r', k', v') \in \mathcal{A}$, if $u_i = v_i$ and $u'_i = v'_i$ for all $i \in M$ and $u_j = u'_j$ and $v_j = v'_j$ for all $j \in \{1, \ldots, n\} \setminus M$, then

$$
(s, l, u) \overset{\circ}{R} (r, k, v) \iff (s', l', u') \overset{\circ}{R} (r', k', v').
\tag{4.10}
$$

We obtain

Theorem 6: *Suppose that $n \geq 3$ and an intertemporal social-evaluation ordering $\overset{\circ}{R}$ satisfies ISP. $\overset{\circ}{R}$ satisfies IUD if and only if $\overset{\circ}{R}$ satisfies IIUUI.*

Proof. Clearly, IIUUI implies IUD. Now suppose $\overset{\circ}{R}$ satisfies ISP and IUD. Let $M \subseteq \{1, \ldots, n\}$, and suppose $(s, l, u), (r, k, v), (s', l', u'), (r', k', v') \in \mathcal{A}$ are such that $u_i = v_i$ and $u_i' = v_i'$ for all $i \in M$ and $u_j = u_j'$ and $v_j = v_j'$ for all $j \in \{1, \ldots, n\} \setminus M$. Let $s_i'' = r_i'' = 0$ and $l_i'' = k_i'' = 1$ for all $i \in M$, and $s_j'' = r_j'' = 1$ and $l_j'' = k_j'' = 1$ for all $j \in \{1, \ldots, n\} \setminus M$. By ISP, $(s'', l'', u)\overset{\circ}{I}(s, l, u)$, $(s'', l'', u')\overset{\circ}{I}(s', l', u')$, $(r'', k'', v)\overset{\circ}{I}(r, k, v)$, and $(r'', k'', v')\overset{\circ}{I}(r', k', v')$. Therefore,

$$(s, l, u) \overset{\circ}{R} (r, k, v) \iff (s'', l'', u)\overset{\circ}{R}(r'', k'', v) \tag{4.11}$$

and

$$(s', l', u') \overset{\circ}{R} (r', k', v') \iff (s'', l'', u')\overset{\circ}{R}(r'', k'', v'). \tag{4.12}$$

Furthermore, by definition,

$$B_2(s'', l'', u) = B_2(s'', l'', u') = B_2(r'', k'', v) = B_2(r'', k'', v') =$$
$$D_2(s'', l'', u) = D_2(s'', l'', u') = D_2(r'', k'', v) = D_2(r'', k'', v') = M_2 = M, \tag{4.13}$$

and IUD implies

$$(s'', l'', u) \overset{\circ}{R} (r'', k'', v) \iff (s'', l'', u') \overset{\circ}{R} (r'', k'', v'). \tag{4.14}$$

Together with (4.11) and (4.12), this implies that IIUUI is satisfied. ∎

The conclusion of Theorem 6 remains true if ISP is weakened by requiring part (i)—an intertemporal version of Pareto indifference—only; note that part (ii) of ISP is not used in the proof.

As an immediate consequence of Theorem 6, a result analogous to Theorem 4 can be obtained in this intertemporal setting. Thus, the characterization result for Generalized-Utilitarian principles is remarkably robust. $\overset{\circ}{R}$ is an Intertemporal Generalized-Utilitarian social-evaluation ordering if and only if there exists a continuous and increasing function $g: \mathcal{R} \longrightarrow \mathcal{R}$ such that, for all $(s, l, u), (r, k, v) \in \mathcal{A}$,

$$(s, l, u) \overset{\circ}{R} (r, k, v) \iff \sum_{i=1}^{n} g(u_i) \geq \sum_{i=1}^{n} g(v_i). \tag{4.15}$$

The proof of the following theorem is an immediate consequence of Theorem 6 and a change in notation that allows us to adapt Theorem 4 to the intertemporal model, and its proof is therefore omitted.

Theorem 7: *Suppose that $n \geq 3$. An intertemporal social-evaluation ordering $\overset{\circ}{R}$ satisfies IC, IA, ISP, and IUD if and only if $\overset{\circ}{R}$ is an Intertemporal Generalized-Utilitarian social-evaluation ordering.*

19

In addition to providing further support for Generalized Utilitarianism, the results of the intertemporal model discussed above illustrate an alternative way of obtaining fully welfarist social-evaluation functionals. Instead of imposing binary independence of irrelevant alternatives in an atemporal model, a limited version of welfarism that includes, in addition to lifetime utilities, birth dates and lengths of life as the available data, can be used to obtain welfarism by means of the strong Pareto principle alone. Weakenings of the strong Pareto principle that allow for birth dates or lifetimes to matter in intertemporal social evaluation are discussed in Blackorby, Bossert and Donaldson (1997a, 1999b).

5. Utilitarianism

The ethical appeal of Generalized Utilitarianism rests, in part, on its separability properties. Utilitarianism is but one possibility within that class of social-evaluation orderings, and it is appropriate to ask whether it should have special status. The arguments for Utilitarianism that we present in this section are based, for the most part, on information-invariance properties.

It is easy to verify that all of the information-invariance assumptions introduced formally in Section 2 are compatible with Utilitarianism. In an informational environment that allows for cardinal unit comparability at least, the Utilitarian social-evaluation ordering can be employed. This is not the case for Generalized Utilitarianism—many Generalized-Utilitarian orderings do not satisfy information invariance with respect to cardinally measurable and fully comparable utilities (CFC) or with respect to translation-scale measurable (TSM) utilities.[18] The application of Generalized Utilitarianism with the function g is restricted to informational environments that allow (at least) for the comparability properties described by the set of admissible transformations in the following theorem.

Theorem 8: *Suppose that $n \geq 2$ and Φ contains n-tuples of continuous and increasing functions only. Generalized Utilitarianism with a continuous and increasing function g satisfies information invariance with respect to Φ if and only if, for each $\phi \in \Phi$, there exist $a_1, \ldots, a_n \in \mathcal{R}$ and $b \in \mathcal{R}_{++}$ such that*

$$\phi_i(\tau) = g^{-1}\big(a_i + bg(\tau)\big) \tag{5.1}$$

for all $\tau \in \mathcal{R}$ and for all $i \in \{1, \ldots, n\}$.

---

[18] In addition, some Generalized-Utilitarian principles fail to satisfy other information-invariance conditions such as ratio-scale full comparability or translation-scale full comparability which are not discussed in this chapter. See, for example, Blackorby and Donaldson (1982).

Proof. That Generalized Utilitarianism satisfies information invariance with respect to Φ if (5.1) is satisfied can be verified by substitution.

Now suppose Generalized Utilitarianism, generated by a function g, satisfies information invariance with respect to Φ and Φ contains continuous and increasing transformations only. Information invariance requires that an admissible transformation $\phi = (\phi_1, \ldots, \phi_n) \in \Phi$ must satisfy the condition

$$\sum_{i=1}^{n} g\big(\phi_i(u_i)\big) \geq \sum_{i=1}^{n} g\big(\phi_i(v_i)\big) \iff \sum_{i=1}^{n} g(u_i) \geq \sum_{i=1}^{n} g(v_i) \tag{5.2}$$

for all $u, v \in \mathcal{R}^n$. This is equivalent to the functional equation

$$\sum_{i=1}^{n} g\big((\phi_i(u_i)\big) = H\Big(\sum_{i=1}^{n} g(u_i)\Big) \tag{5.3}$$

for all $u \in \mathcal{R}^n$, where H is increasing. Letting $z_i = g(u_i)$ and $G_i = g \circ \phi_i \circ g^{-1}$ for all $i \in \{1, \ldots, n\}$, (5.3) can be rewritten as

$$\sum_{i=1}^{n} G_i(z_i) = H\Big(\sum_{i=1}^{n} z_i\Big), \tag{5.4}$$

a Pexider equation, and it follows that $G_i(\tau) = a_i + b\tau$ with $b \in \mathcal{R}_{++}$ and $a_i \in \mathcal{R}$ for all $i \in \{1, \ldots, n\}$.[19] Substituting back, we obtain $\phi_i(\tau) = g^{-1}\big(a_i + bg(\tau)\big)$ for all $i \in \{1, \ldots, n\}$. ∎

Condition (5.1) says, in effect, that the information environment must support cardinal unit comparability of transformed utilities $(g(u_1), \ldots, g(u_n))$. It is difficult to justify such an information environment unless the function g is affine, in which case we are back to Utilitarianism. Therefore, the informational difficulties involved in applying Generalized-Utilitarian principles other than Utilitarianism suggest that the Utilitarian social-evaluation functional has an important advantage over its competitors within that class.

Information-invariance assumptions can be used to provide characterizations of Utilitarianism. The following theorem is a strengthening of a result due to d'Aspremont and Gevers (1977) who use the stronger axiom information invariance with respect to cardinal unit comparability instead of information invariance with respect to translation-scale measurability (see also Blackwell and Girshick, 1954, Milnor, 1954, and Roberts, 1980b).

---

[19] See Aczél (1966, Chapter 3) for a detailed discussion of Pexider equations and their solutions. Because the functions ϕ_i and g (and thus the inverse of g) are continuous, the domains of H and the G_i are nondegenerate intervals, which ensures that the requisite functional-equations results apply. Pexider equations are also discussed in Eichhorn (1978).

Theorem 9: *A social-evaluation ordering $\overset{*}{R}$ satisfies A, WP, and information invariance with respect to TSM if and only if $\overset{*}{R}$ is the Utilitarian social-evaluation ordering.*

Proof. That the Utilitarian social-evaluation ordering satisfies the required axioms is easily verified. Suppose, therefore, that $\overset{*}{R}$ satisfies A, WP, and information invariance with respect to TSM. It is sufficient to show that, for all $u, v \in \mathcal{R}^n$,

$$\sum_{i=1}^{n} u_i = \sum_{i=1}^{n} v_i \implies u \overset{*}{I} v; \tag{5.5}$$

once this is established, the implication

$$\sum_{i=1}^{n} u_i > \sum_{i=1}^{n} v_i \implies u \overset{*}{P} v \tag{5.6}$$

for all $u, v \in \mathcal{R}^n$ follows immediately from WP.

Reflexivity of $\overset{*}{R}$ and WP are sufficient to establish (5.5) and (5.6) for the case $n = 1$. Suppose $n \geq 2$ and let $u, v \in \mathcal{R}^n$ be such that $\sum_{i=1}^{n} u_i = \sum_{i=1}^{n} v_i$. Define $u^0 = u$ and $v^0 = v$, and let $\tilde{u}^0$ and $\tilde{v}^0$ be welfare-ranked permutations of u^0 and v^0, that is, $\tilde{u}_1^0 \geq \tilde{u}_2^0 \geq \ldots \geq \tilde{u}_n^0$ and $\tilde{v}_1^0 \geq \tilde{v}_2^0 \geq \ldots \geq \tilde{v}_n^0$. For all $k \in \{1, \ldots, n\}$, let

$$u_i^k = \tilde{u}_i^{k-1} - \min\{\tilde{u}_i^{k-1}, \tilde{v}_i^{k-1}\} \tag{5.7}$$

and

$$v_i^k = \tilde{v}_i^{k-1} - \min\{\tilde{u}_i^{k-1}, \tilde{v}_i^{k-1}\} \tag{5.8}$$

for all $i \in \{1, \ldots, n\}$, and let $\tilde{u}^k$ and $\tilde{v}^k$ be welfare-ranked permutations of u^k and v^k. By A, $u^k \overset{*}{I} \tilde{u}^k$ and $v^k \overset{*}{I} \tilde{v}^k$ for all $k \in \{0, \ldots, n\}$. Because u^k and v^k can be obtained from $\tilde{u}^{k-1}$ and $\tilde{v}^{k-1}$ by means of an admissible transformation for TSM with $a_i = -\min\{\tilde{u}_i^{k-1}, \tilde{v}_i^{k-1}\}$ for all $i \in \{1, \ldots, n\}$, information invariance with respect to TSM implies

$$u^k \overset{*}{R} v^k \iff \tilde{u}^{k-1} \overset{*}{R} \tilde{v}^{k-1} \tag{5.9}$$

for all $k \in \{1, \ldots, n\}$. By transitivity,

$$u^0 \overset{*}{R} v^0 \iff u^n \overset{*}{R} v^n. \tag{5.10}$$

Because, by definition, $u = u^0$, $v = v^0$ and $u^n = v^n = 0\mathbf{1}_n$,

$$u \overset{*}{R} v \iff 0\mathbf{1}_n \overset{*}{R} 0\mathbf{1}_n, \tag{5.11}$$

and reflexivity of $\overset{*}{R}$ implies $u \overset{*}{I} v$. ■

An alternative characterization of Utilitarianism can be obtained for the case $n \geq 3$ by employing information invariance with respect to CFC and the separability axiom IUUI together with C, A and SP. This theorem is due to Maskin (1978)—see also Deschamps and Gevers (1978).

Theorem 10: *Suppose that $n \geq 3$. A social-evaluation ordering $\overset{*}{R}$ satisfies C, A, SP, IUUI, and information invariance with respect to CFC if and only if $\overset{*}{R}$ is the Utilitarian social-evaluation ordering.*

Proof. That Utilitarianism satisfies the required axioms is easy to verify. Now suppose a social-evaluation ordering $\overset{*}{R}$ satisfies C, A, SP, IUUI, and information invariance with respect to CFC. By Theorem 4, $\overset{*}{R}$ is Generalized Utilitarian with a continuous and increasing function g. Because any increasing affine transformation of g leads to the same ordering of utility vectors, we can without loss of generality assume that $g(0) = 0$ and $g(1) = 1$. It remains to be shown that, given this normalization, g must be the identity mapping.

Information invariance with respect to CFC requires

$$\sum_{i=1}^{n} g(a + bu_i) \geq \sum_{i=1}^{n} g(a + bv_i) \iff \sum_{i=1}^{n} g(u_i) \geq \sum_{i=1}^{n} g(v_i) \tag{5.12}$$

for all $u, v \in \mathcal{R}^n$, $a \in \mathcal{R}$, $b \in \mathcal{R}_{++}$. This is equivalent to the functional equation

$$\sum_{i=1}^{n} g(a + bu_i) = H_{a,b}\left(\sum_{i=1}^{n} g(u_i)\right) \tag{5.13}$$

for all $u \in \mathcal{R}^n$, $a \in \mathcal{R}$, $b \in \mathcal{R}_{++}$, where $H_{a,b}$ is increasing. Letting $z_i = g(u_i)$ and $G_{a,b}(z_i) = g(a + bg^{-1}(z_i))$, (5.13) can be rewritten as

$$\sum_{i=1}^{n} G_{a,b}(z_i) = H_{a,b}\left(\sum_{i=1}^{n} z_i\right). \tag{5.14}$$

Our continuity and monotonicity assumptions ensure that all solutions to this Pexider equation are such that $G_{a,b}(z_i) = A(a,b) + B(a,b)z_i$. Substituting back, we obtain the equation

$$g(a + b\tau) = A(a,b) + B(a,b)g(\tau) \tag{5.15}$$

for all $a, \tau \in \mathcal{R}$ and all $b \in \mathcal{R}_{++}$, where we use τ instead of u_i for simplicity. Setting $\tau = 0$ and using the normalization $g(0) = 0$, we obtain $A(a,b) = g(a)$, and choosing $\tau = 1$ in (5.15) yields, together with the normalization $g(1) = 1$, $B(a,b) = g(a+b) - g(a)$. Therefore, (5.15) is equivalent to

$$g(a + b\tau) = g(\tau)\big[g(a + b) - g(a)\big] + g(a) \tag{5.16}$$

for all $a, \tau \in \mathcal{R}$ and all $b \in \mathcal{R}_{++}$. Setting $a = 0$, we obtain

$$g(b\tau) = g(b)g(\tau) \tag{5.17}$$

for all $\tau \in \mathcal{R}$ and all $b \in \mathcal{R}_{++}$. Analogously, choosing $b = 1$ in (5.16) yields

$$g(a + \tau) = g(\tau)\big[g(a + 1) - g(a)\big] + g(a) \tag{5.18}$$

for all $a, \tau \in \mathcal{R}$. This is a special case of Equation 3.1.3(3) in Aczél (1966, p. 150)[20] and, together with the increasingness of g, it follows that either there exists a $c \in \mathcal{R}_{++}$ such that

$$g(\tau) = \frac{e^{c\tau} - 1}{e^c - 1} \tag{5.19}$$

for all $\tau \in \mathcal{R}$, or $g(\tau) = \tau$ for all $\tau \in \mathcal{R}$. Because (5.19) is incompatible with (5.17), this completes the proof. ∎

Continuity plays a crucial role in Theorem 10. Deschamps and Gevers (1978) examine the consequences of dropping C from the list of axioms in the above theorem. Among other results, they show that if a social-evaluation ordering $\overset{*}{R}$ satisfies A, SP, IUUI, and information invariance with respect to CFC, then $\overset{*}{R}$ must be Weakly Utilitarian, Leximin or Leximax. It is a remarkable observation that these axioms narrow down the class of possible social-evaluation orderings to that extent. When minimal equity is added, only Weakly Utilitarian principles and Leximin survive because Leximax obviously violates ME. Therefore, we obtain the following theorem, which is due to Deschamps and Gevers (1978). Because the proof is very lengthy and involved, we state the theorem without proving it and refer interested readers to the appendix of their paper.

Theorem 11: *Suppose that $n \geq 3$. If a social-evaluation ordering $\overset{*}{R}$ satisfies A, SP, ME, IUUI, and information invariance with respect to CFC, then $\overset{*}{R}$ is the Leximin social-evaluation ordering or a Weakly Utilitarian social-evaluation ordering.*

It should be noted that the above theorem is not a characterization result because its statement is an implication rather than an equivalence. The reason is that not all Weakly Utilitarian orderings satisfy all the required axioms.

It is possible to characterize Utilitarianism without information restrictions by employing an axiom that we call incremental equity. In its definition, $\mathbf{1}_n^j$ is the vector $x \in \mathcal{R}^n$ with $x_j = 1$ and $x_i = 0$ for all $i \in \{1, \ldots, n\} \setminus \{j\}$.

Incremental Equity (IE): For all $u \in \mathcal{R}^n$, for all $\delta \in \mathcal{R}$, for all $j, k \in \{1, \ldots, n\}$,

$$\big(u + \delta\mathbf{1}_n^j\big) \overset{*}{I} \big(u + \delta\mathbf{1}_n^k\big). \tag{5.20}$$

IE requires a kind of impartiality with respect to utility increases or decreases. If a single individual's utility level changes by the amount δ, IE requires the change to be ranked as equally good, no matter who receives the increment. Incremental equity and weak Pareto together characterize Utilitarianism.

---

[20] To see this, set $f(x) = k(x) = g(\tau)$ and $h(y) = g(a + 1) - g(a)$ in Aczél (1966, Equation 3.1.3(3)).

Theorem 12: *A social-evaluation ordering $\overset{*}{R}$ satisfies WP and IE if and only if $\overset{*}{R}$ is the Utilitarian social-evaluation ordering.*

Proof. That the Utilitarian social-evaluation ordering satisfies WP and IE is easily checked.

If $n = 1$, WP alone implies the result. Therefore, let $n \geq 2$. Applying IE to $(u - \delta \mathbf{1}_n^j)$, (5.20) implies that

$$u \overset{*}{I} \left(u - \delta \mathbf{1}_n^j + \delta \mathbf{1}_n^k \right). \tag{5.21}$$

For any $u \in \mathcal{R}^n$, (5.21) implies

$$
\begin{aligned}
u \overset{*}{I} & \left(\frac{1}{n} \sum_{i=1}^{n} u_i, u_2, \ldots, u_n + u_1 - \frac{1}{n} \sum_{i=1}^{n} u_i \right) \\
\overset{*}{I} & \left(\frac{1}{n} \sum_{i=1}^{n} u_i, \frac{1}{n} \sum_{i=1}^{n} u_i, \ldots, u_n + u_1 + u_2 - \frac{2}{n} \sum_{i=1}^{n} u_i \right) \\
& \vdots \\
\overset{*}{I} & \left(\frac{1}{n} \sum_{i=1}^{n} u_i, \ldots, \sum_{i=1}^{n} u_i - \frac{n-1}{n} \sum_{i=1}^{n} u_i \right) = \left(\frac{1}{n} \sum_{i=1}^{n} u_i \right) \mathbf{1}_n.
\end{aligned}
\tag{5.22}
$$

Using WP, this implies

$$u \overset{*}{R} v \iff \sum_{i=1}^{n} u_i \geq \sum_{i=1}^{n} v_i \tag{5.23}$$

for all $u, v \in \mathcal{R}^n$. ∎

Equation (5.21) in the proof shows that IE requires social indifference about transfers of utility from one individual to another. Consequently, all distributions of the same total must be regarded as equally good and Utilitarianism results.

6. Variable-Population Extensions

Utilitarian and Generalized-Utilitarian social-evaluation orderings may be extended to a variable-population framework in different ways. As an example, Average and Classical Utilitarianism, which use average and total utility levels of those alive to rank alternatives respectively, coincide on fixed-population rankings but may order alternatives with different population sizes differently.

For each $x \in X$, let $\mathbf{N}(x) = N$ denote the set of individuals alive in x, where $N \subseteq \mathcal{Z}_{++}$ is finite and nonempty.[21] Furthermore, let $\mathcal{Z}_{++}$ be the set of potential people and define

---

[21] If the empty set were included as a possible population, all results in this section would still be valid. See, for example, Blackorby, Bossert and Donaldson (1995) for details.

$X_i = \{x \in X \mid i \in \mathbf{N}(x)\}$ to be the set of all alternatives in which individual $i \in \mathcal{Z}_{++}$ is alive. Individual i's utility function is $U_i \colon X_i \longrightarrow \mathcal{R}$ and a profile of utility functions is $U = (U_i)_{i \in \mathcal{Z}_{++}}$. We follow the standard convention in population ethics and normalize lifetime utilities so that a lifetime-utility level of zero represents neutrality. A life, taken as a whole, is worth living for an individual if and only if lifetime utility is above neutrality. Consequently, a fully informed self-interested and rational person whose lifetime-utility level is below neutrality would prefer not to have any of his or her experiences.[22] We assume that, for each nonempty and finite $\hat{N} \subseteq \mathcal{Z}_{++}$, the set $\{x \in X \mid \mathbf{N}(x) = \hat{N}\}$ contains at least three elements. This assumption, which is analogous to the fixed-population assumption that X contains at least three elements, ensures that a variable-population version of the welfarism theorem is valid. The vector of lifetime utilities of those alive in alternative $x \in X$ is $\left(U_i(x)\right)_{i \in \mathbf{N}(x)} = (u_i)_{i \in N}$. $\mathcal{U}^E$ is the set of all possible utility profiles $(U_i)_{i \in \mathcal{Z}_{++}}$, which extends the domain $\mathcal{U}$ employed in earlier sections to a variable-population framework.

A variable-population social-evaluation functional is a mapping $F^E \colon \mathcal{D}^E \longrightarrow \mathcal{O}$, where $\mathcal{D}^E \subseteq \mathcal{U}^E$ is the set of admissible profiles. For all $U \in \mathcal{D}^E$, the social no-worse-than relation is $R_U^E = F^E(U)$ and I_U^E and P_U^E denote its symmetric and asymmetric components. As in the fixed-population case, variable-population welfarism is the consequence of three axioms.

Population Unrestricted Domain (PUD): $\mathcal{D}^E = \mathcal{U}^E$.

Population Binary Independence of Irrelevant Alternatives (PBI): For all $x, y \in X$, for all $U, V \in \mathcal{D}^E$, if $U_i(x) = V_i(x)$ for all $i \in \mathbf{N}(x)$ and $U_i(y) = V_i(y)$ for all $i \in \mathbf{N}(y)$, then $xR_U^E y$ if and only if $xR_V^E y$.

Population Pareto Indifference (PPI): For all $x, y \in X$ such that $\mathbf{N}(x) = \mathbf{N}(y)$, for all $U \in \mathcal{D}^E$, if $U_i(x) = U_i(y)$ for all $i \in \mathbf{N}(x)$, then $xI_U^E y$.

Note that population Pareto indifference is a fixed-population axiom (it applies to comparisons of alternatives with the same people alive in each only), whereas population binary independence of irrelevant alternatives imposes restrictions on the comparison of alternatives that may involve different populations and population sizes. PBI requires the social ranking of any pair of alternatives to be the same if two profiles coincide on the pair.

Results analogous to Theorems 1 and 2 are valid in this variable-population model— see Blackorby, Bossert and Donaldson (1999a) for details. Because we restrict attention to anonymous variable-population social-evaluation functionals in this section, we do not provide formal statements of the corresponding generalizations and, instead, state a related result that incorporates a variable-population anonymity condition.

---

[22] See Broome (1993) for a discussion of neutrality and its normalization to zero.

Population Anonymity (PA): For all $U, V \in \mathcal{D}^E$, for all bijective mappings $\pi \colon \mathcal{Z}_{++} \longrightarrow \mathcal{Z}_{++}$ such that $U_i = V_{\pi(i)}$ for all $i \in \mathcal{Z}_{++}$,

$$R_U^E = R_V^E. \tag{6.1}$$

Let $\Omega = \cup_{n \in \mathcal{Z}_{++}} \mathcal{R}^n$. An ordering $\overset{*}{R}{}^E$ on Ω is anonymous if and only if the restriction of $\overset{*}{R}{}^E$ to $\mathcal{R}^n$ satisfies A for all $n \in \mathcal{Z}_{++}$.

We now obtain the following anonymous variable-population version of the welfarism theorem. Since the proof of this theorem is analogous to its fixed-population version, it is omitted. See Blackorby, Bossert and Donaldson (1999a) and Blackorby and Donaldson (1984) for details.

Theorem 13: *Suppose that a variable-population social-evaluation functional F^E satisfies PUD. F^E satisfies PBI, PPI, and PA if and only if there exists an anonymous ordering $\overset{*}{R}{}^E$ on Ω such that, for all $x, y \in X$ and for all $U \in \mathcal{D}^E$,*

$$x R_U^E y \iff \left(U_i(x)\right)_{i \in \mathbf{N}(x)} \overset{*}{R}{}^E \left(U_i(y)\right)_{i \in \mathbf{N}(y)}. \tag{6.2}$$

We call $\overset{*}{R}{}^E$ a variable-population social-evaluation ordering, and we use $\overset{*}{I}{}^E$ and $\overset{*}{P}{}^E$ to denote its symmetric and asymmetric components.

We say that the ordering $\overset{*}{R}{}^E$ satisfies the fixed-population axioms continuity, weak Pareto, strong Pareto and independence of the utilities of unconcerned individuals respectively if and only if, for all $n \in \mathcal{Z}_{++}$, the restriction of $\overset{*}{R}{}^E$ to $\mathcal{R}^n$ satisfies the appropriate fixed-population axiom.

C and WP guarantee the existence of a representation of any anonymous $\overset{*}{R}{}^E$ that depends on population size and representative utility only. Note that no variable-population axioms are required for this result, which is due to Blackorby and Donaldson (1984). In the theorem statement, Ξ^n is the representative-utility function for the restriction of $\overset{*}{R}{}^E$ to $\mathcal{R}^n$.

Theorem 14: *If an anonymous variable-population social-evaluation ordering $\overset{*}{R}{}^E$ satisfies C and WP, then there exists a value function $W \colon \mathcal{Z}_{++} \times \mathcal{R} \longrightarrow \mathcal{R}$, continuous and increasing in its second argument, such that, for all $n, m \in \mathcal{Z}_{++}$, for all $u \in \mathcal{R}^n$, for all $v \in \mathcal{R}^m$,*

$$u \overset{*}{R}{}^E v \iff W\big(n, \Xi^n(u)\big) \geq W\big(m, \Xi^m(v)\big). \tag{6.3}$$

Proof. The existence of the representative-utility function Ξ^n for all $n \in \mathcal{Z}_{++}$ is guaranteed by C and WP (see Theorem 3). Define the ordering $\mathbf{R}$ on $\mathcal{Z}_{++} \times \mathcal{R}$ by letting, for all $(n, \xi), (m, \zeta) \in \mathcal{Z}_{++} \times \mathcal{R}$, $(n, \xi)\mathbf{R}(m, \zeta)$ if and only if $\xi \mathbf{1}_n \overset{*}{R}{}^E \zeta \mathbf{1}_m$. By definition of the representative utility, $u \overset{*}{R}{}^E v$ if and only if $\big(n, \Xi^n(u)\big)\mathbf{R}\big(m, \Xi^m(v)\big)$ for all $n, m \in \mathcal{Z}_{++}$ for all $u \in \mathcal{R}^n$, for all $v \in \mathcal{R}^m$. By C, for any $n \in \mathcal{Z}_{++}$, the restriction of $\mathbf{R}$ to $\{(n, \xi) \mid \xi \in \mathcal{R}\}$ can be represented by a function the image of which is an open interval of length one. Because $\mathcal{Z}_{++}$ is countable, it follows that $\mathbf{R}$ can be represented by a function the image of which is a countable union of such intervals. This establishes the existence of the function W, and continuity and increasingness in its second argument follow immediately from C and WP. ■

A natural generalization of IUUI requires this separability property to hold for all comparisons, not only those that involve utility vectors of the same dimension. See Blackorby, Bossert and Donaldson (1998) for a discussion.

Extended Independence of the Utilities of Unconcerned Individuals (EIUUI): For all $u, v, w \in \Omega$, $u \overset{*}{R}{}^E v$ if and only if $(u, w) \overset{*}{R}{}^E (v, w)$.

The next axiom establishes a link between different population sizes. It requires, for each possible alternative, the existence of another alternative with an additional individual alive that is as good as the original alternative, where the individuals alive in both alternatives are unaffected by the population augmentation. This assumption rules out social orderings that always declare an alternative with a larger population better (worse) than an alternative with a smaller population, thereby ensuring nontrivial trade-offs between population size and well-being.

Expansion Equivalence (EE): For all $u \in \Omega$, there exists $c \in \mathcal{R}$ such that $(u, c) \overset{*}{I}{}^E u$.

The number c in the definition of EE is a critical level of lifetime utility for $u \in \Omega$. If $\overset{*}{R}{}^E$ satisfies SP and EE, the critical level for any $u \in \Omega$ is unique and can be written as $c = C(u)$, where $C : \Omega \longrightarrow \mathcal{R}$ is a critical-level function.

For most of the results in this section, it is not necessary to impose EE—a weaker version which requires the existence of a critical level only for at least one $\bar{u} \in \Omega$ is sufficient. Therefore, we define

Weak Expansion Equivalence (WEE): There exist $\bar{u} \in \Omega$ and $c \in \mathcal{R}$ such that $(\bar{u}, c) \overset{*}{I}{}^E \bar{u}$.

If a variable-population social-evaluation ordering $\overset{*}{R}{}^E$ satisfies EIUUI and WEE, then $\overset{*}{R}{}^E$ satisfies EE. Moreover, if c is a critical level for $\bar{u}$, then c is a critical level for all $u \in \Omega$. See Blackorby, Bossert and Donaldson (1995) for an analogous result in an intertemporal model.

Theorem 15: *If an anonymous variable-population social-evaluation ordering $\overset{*}{R}{}^E$ satisfies EIUUI and WEE, then $\overset{*}{R}{}^E$ satisfies EE and there exists $\alpha \in \mathcal{R}$ such that α is a critical level for all $u \in \Omega$.*

Proof. By WEE, there exist $\bar{u} \in \Omega$ and $c \in \mathcal{R}$ such that $(\bar{u}, c)\overset{*}{I}{}^E\bar{u}$. Let $u \in \Omega$ be arbitrary. By EIUUI, $(u, \bar{u}, c)\overset{*}{I}{}^E(u, \bar{u})$, and applying EIUUI again, it follows that $(u, c)\overset{*}{I}{}^Eu$. Letting $\alpha = c$, the theorem is established. ∎

If SP is added to EIUUI and WEE, Theorem 15 implies that the critical-level function C is constant.[23]

There are several possibilities for extending Utilitarianism to a variable-population framework. For example, Average Utilitarianism is defined as follows. For all $n, m \in \mathcal{Z}_{++}$, for all $u \in \mathcal{R}^n$, for all $v \in \mathcal{R}^m$,

$$u \overset{*}{R}{}^E v \iff \frac{1}{n}\sum_{i=1}^{n} u_i \geq \frac{1}{m}\sum_{i=1}^{m} v_i. \tag{6.4}$$

Average Utilitarianism satisfies EE with average utility as the critical level for any utility vector $u \in \Omega$. In addition, Average Utilitarianism satisfies IUUI but not EIUUI.[24]

Critical-Level Utilitarianism (Blackorby, Bossert and Donaldson, 1995, Blackorby and Donaldson, 1984) uses the sum of the differences between individual utility levels and a fixed critical level of utility as its value function. If a person is added to a population that is unaffected in terms of utilities, the critical level is that level of lifetime utility that makes the two alternatives equally good. $\overset{*}{R}{}^E$ is a Critical-Level Utilitarian social-evaluation ordering if and only if there exists $\alpha \in \mathcal{R}$ such that, for all $n, m \in \mathcal{Z}_{++}$, for all $u \in \mathcal{R}^n$, for all $v \in \mathcal{R}^m$,

$$u \overset{*}{R}{}^E v \iff \sum_{i=1}^{n}\big[u_i - \alpha\big] \geq \sum_{i=1}^{m}\big[v_i - \alpha\big]. \tag{6.5}$$

Critical-Level Utilitarianism satisfies EE with a constant critical level α and, in addition, satisfies EIUUI. Classical Utilitarianism is a special case of Critical-Level Utilitarianism with α equal to zero, the level of lifetime utility representing neutrality.

We believe that a positive value (that is, a value above neutrality) should be chosen for the critical level α. If α is at or below the critical level, the 'repugnant conclusion' results (Parfit, 1976, 1982, 1984). A variable-population social-evaluation ordering leads

---

[23] Hammond (1988, 1999) and Dasgupta (1993) argue for constant critical levels and normalize them to zero. Both set the critical level above neutrality, which implies that the utility level that represents neutrality is negative. Hammond uses an axiom that is similar to IUD.

[24] See Blackorby and Donaldson (1984, 1991), Bossert (1990b,c) and Hurka (1982) for discussions of Average Utilitarianism.

to the repugnant conclusion if, for any level of utility experienced by each individual in a society of a given size (no matter how far above neutrality this utility level is) and for any utility level μ above neutrality but arbitrarily close to it, there exists a larger population size m so that an alternative in which each of m individuals has a utility of μ is considered better than the former alternative. Therefore, a situation with mass poverty is superior to an alternative in which every person leads a good life. Because we consider the repugnant conclusion ethically unacceptable, we recommend a level of α above neutrality.

Positive critical levels are incompatible with the Pareto plus principle (Sikora, 1978), which requires the ceteris paribus addition of an individual above neutrality to a given population to be desirable. Classical Utilitarianism satisfies the Pareto plus principle and leads to the repugnant conclusion. On the other hand, Average Utilitarianism does not lead to the repugnant conclusion and it does not satisfy the Pareto plus principle. Although it performs well in this regard, it possesses the ethically unattractive property of declaring the ceteris paribus addition of a person below neutrality desirable in some situations. The incompatibility between avoiding the repugnant conclusion and the Pareto plus principle is not restricted to Utilitarian principles, however. More general impossibility results are reported in Arrhenius (1997), Blackorby, Bossert, Donaldson and Fleurbaey (1998), Blackorby and Donaldson (1991), Carlson (1998) and Ng (1989).

The variable-population versions of Generalized Utilitarianism use the same transformation of utilities g for all population sizes. Moreover, because this function is unique up to increasing affine transformations only, we can, without loss of generality, assume that $g(0) = 0$; this ensures that the utility level that represents neutrality is preserved when applying the transformation.

$\overset{*}{R}{}^{E}$ is the Average Generalized-Utilitarian social-evaluation ordering if and only if there exists a continuous and increasing function $g \colon \mathcal{R} \longrightarrow \mathcal{R}$ with $g(0) = 0$ such that, for all $n, m \in \mathcal{Z}_{++}$, for all $u \in \mathcal{R}^n$, for all $v \in \mathcal{R}^m$,

$$u \overset{*}{R}{}^{E} v \iff \frac{1}{n}\sum_{i=1}^{n} g(u_i) \geq \frac{1}{m}\sum_{i=1}^{m} g(v_i). \tag{6.6}$$

$\overset{*}{R}{}^{E}$ is a Critical-Level Generalized-Utilitarian social-evaluation ordering if and only if there exist $\alpha \in \mathcal{R}$ and a continuous and increasing function $g \colon \mathcal{R} \longrightarrow \mathcal{R}$ satisfying $g(0) = 0$ such that, for all $n, m \in \mathcal{Z}_{++}$, for all $u \in \mathcal{R}^n$, for all $v \in \mathcal{R}^m$,

$$u \overset{*}{R}{}^{E} v \iff \sum_{i=1}^{n} \big[g(u_i) - g(\alpha)\big] \geq \sum_{i=1}^{m} \big[g(v_i) - g(\alpha)\big]. \tag{6.7}$$

Setting $\alpha = 0$ yields Classical Generalized Utilitarianism.

As is the case for Average Utilitarianism, Average Generalized Utilitarianism satisfies EE with average utility as the critical level and satisfies IUUI but violates EIUUI. Critical-Level Generalized Utilitarianism satisfies EE with the constant critical level α and EIUUI. Classical Generalized Utilitarianism satisfies the Pareto plus principle and leads to the repugnant conclusion but Average Generalized Utilitarianism and Critical-Level Generalized Utilitarianism with a positive critical level avoid the repugnant conclusion and violate Pareto plus.

Analogously to Theorem 4, EIUUI can be used to characterize Critical-Level Generalized Utilitarianism in the variable-population case. The following theorem is due to Blackorby, Bossert and Donaldson (1998).

Theorem 16: *An anonymous variable-population social-evaluation ordering $\overset{*}{R}{}^E$ satisfies C, SP, EIUUI, and WEE if and only if $\overset{*}{R}{}^E$ is a Critical-Level Generalized-Utilitarian social-evaluation ordering.*

Proof: By Theorem 4, fixed-population comparisons for population sizes $n \geq 3$ must be made according to fixed-population Generalized Utilitarianism with continuous and increasing functions g^n. Because each g^n is unique up to increasing affine transformations only, we can without loss of generality assume $g^n(0) = 0$ for all $n \geq 3$. EIUUI requires

$$\sum_{i=1}^{n} g^{n+m}(u_i) \geq \sum_{i=1}^{n} g^{n+m}(v_i) \iff \sum_{i=1}^{n} g^n(u_i) \geq \sum_{i=1}^{n} g^n(v_i) \tag{6.8}$$

for all $n, m \geq 3$, for all $u, v \in \mathcal{R}^n$, which implies that the g^n can be chosen independently of n, and we define $g = g^n$ for all $n \geq 3$. By Theorem 15 and SP, there exists a unique constant critical level $\alpha \in \mathcal{R}$. Consider $u \in \mathcal{R}^n$ and $v \in \mathcal{R}^m$ with $n \geq 3$ and, without loss of generality, $n \geq m$. Because α is a critical level for v, it follows that

$$u \overset{*}{R}{}^E v \iff u \overset{*}{R}{}^E (v, \alpha \mathbf{1}_{n-m}). \tag{6.9}$$

Because u and $(v, \alpha \mathbf{1}_n)$ are of the same dimension $n \geq 3$, it follows that

$$
\begin{aligned}
u \overset{*}{R}{}^E v &\iff u \overset{*}{R}{}^E (v, \alpha \mathbf{1}_n) \\
&\iff \sum_{i=1}^{n} g(u_i) \geq \sum_{i=1}^{m} g(v_i) + (n-m)g(\alpha) \\
&\iff \sum_{i=1}^{n} \big[g(u_i) - g(\alpha)\big] \geq \sum_{i=1}^{m} \big[g(v_i) - g(\alpha)\big].
\end{aligned}
\tag{6.10}
$$

If $n < 3$, the definition of a critical level can be used again to conclude $u \overset{*}{I}{}^E (u, \alpha \mathbf{1}_{3-n})$, and the above argument can be repeated with u replaced by $(u, \alpha \mathbf{1}_{3-n})$. ∎

If the requirement that the repugnant conclusion be avoided is added to the axioms of Theorem 16, it follows immediately that the critical level must be positive.

As in the fixed-population case, Critical-Level Generalized Utilitarianism can be characterized in an intertemporal model with a variable-population version of independence of the utilities of the dead. This extended version of IUD is obtained from IUD in the same way EIUUI is obtained from IUUI; see Blackorby, Bossert and Donaldson (1995) for details.

Alternative intertemporal consistency conditions are explored in Blackorby, Bossert and Donaldson (1996). In the intertemporal setting, individuals are assumed to experience utilities in each period which aggregate into lifetime utilities. Forward-looking consistency requires that, in any period, future utilities are separable from past utilities. In Blackorby, Bossert and Donaldson (1996), it is shown that consistency between forward-looking social evaluations and intertemporal social evaluations implies, together with some other axioms, Classical Generalized Utilitarianism and thus the repugnant conclusion. The same results are obtained for a full intertemporal consistency requirement which is stronger than forward-looking consistency by itself but is equivalent to it in the presence of other axioms. The consequences of weakening the intertemporal strong Pareto principle are examined in Blackorby, Bossert and Donaldson (1997a,b), where versions of Critical-Level Generalized Utilitarianism and Classical Generalized Utilitarianism that allow for discounting are characterized.

We conclude this section with a discussion of information-invariance assumptions in the variable-population framework. Let $\sim^E$ denote an equivalence relation defined on the domain $\mathcal{D}^E$ of a variable-population social-evaluation functional F^E. Information invariance with respect to $\sim^E$ is defined as follows.

Information Invariance with Respect to $\sim^E$: For all $U, V \in \mathcal{D}^E$, if $U \sim^E V$, then $R_U^E = R_V^E$.

In the presence of welfarism, information invariance can alternatively be defined in terms of $\overset{*}{R}{}^E$.

As in the fixed-population case, one possible way to define an information assumption is to specify a set of admissible vectors of utility transformations. In the variable-population case, the elements of such a set Φ can be written as $\phi = (\phi_i)_{i \in \mathcal{Z}_{++}}$, where each ϕ_i is a function $\phi_i \colon \mathcal{R} \longrightarrow \mathcal{R}$ that transforms individual i's utility u_i into $\phi_i(u_i)$.

The following information assumptions are used in this section.

Ordinal Full Comparability (OFC): $\phi \in \Phi$ if and only if there exists an increasing function $\phi_0 \colon \mathcal{R} \longrightarrow \mathcal{R}$ such that $\phi_i = \phi_0$ for all $i \in \mathcal{Z}_{++}$.

Cardinal Measurability (CM): $\phi \in \Phi$ if and only if there exist $a_i \in \mathcal{R}$ and $b_i \in \mathcal{R}_{++}$ for each $i \in \mathcal{Z}_{++}$ such that $\phi_i(\tau) = a_i + b_i\tau$ for all $\tau \in \mathcal{R}$ and all $i \in \mathcal{Z}_{++}$.

Cardinal Full Comparability (CFC): $\phi \in \Phi$ if and only if there exist $a \in \mathcal{R}$ and $b \in \mathcal{R}_{++}$ such that $\phi_i(\tau) = a + b\tau$ for all $\tau \in \mathcal{R}$ and all $i \in \mathcal{Z}_{++}$.

Numerical Full Comparability (NFC): $\phi \in \Phi$ if and only if $\phi_i(\tau) = \tau$ for all $\tau \in \mathcal{R}$ and all $i \in \mathcal{Z}_{++}$.

In the variable-population framework, information assumptions are considerably more restrictive than in the fixed-population case. In the presence of anonymity and the weak Pareto principle, for example, the possibility of level comparisons is necessary for the existence of a variable-population social-evaluation functional. This implies that stronger information-invariance requirements than information invariance with respect to OFC cannot be satisfied. See Blackorby, Bossert and Donaldson (1999a) for details.

Furthermore, some information-invariance assumptions impose significant restrictions on ethical parameters such as critical levels, which clearly is undesirable. It can be shown that information invariance with respect to CFC leads to Average Utilitarianism in the presence of some other axioms including IUUI, and if IUUI is strengthened to EIUUI, an impossibility result is obtained. These results are proved and discussed in Blackorby, Bossert and Donaldson (1999a).

Because of these negative observations, we suggest an alternative way of formulating information invariance in a variable-population framework. The fundamental difficulty appears to be that the standard welfarist framework with an unrestricted domain is inadequate to define norms, such as neutrality, which permit interpersonal comparisons of utility at a single utility level. Therefore, we suggest the use of a systematic procedure for incorporating such norms (see Blackorby, Bossert and Donaldson, 1999a). In particular, we propose the use of norms to restrict the domain of admissible utility profiles.[25]

For $U \in \mathcal{U}^E$ and $i \in \mathcal{Z}_{++}$, let $\eta_i(U_i)$ denote the level of utility individual i assigns to a neutral life, given the utility function U_i. Suppose, in addition, that a second norm denotes a life above neutrality at some satisfactory or 'excellent' level (not necessarily a critical level). It is possible, given these norms, to represent the value of a neutral life with a utility level of zero and the value of an excellent life with a utility level of one. Letting $\varepsilon_i(U_i)$ denote the utility level representing an excellent life according to i's utility function U_i, the restricted domain that respects both normalizations is given by

$$\mathcal{D}^E = \mathcal{U}^E_{\eta\varepsilon} = \left\{ U \in \mathcal{U}^E \mid \eta_i(U_i) = 0 \ \text{ and } \ \varepsilon_i(U_i) = 1 \ \ \forall i \in \mathcal{Z}_{++} \right\}. \qquad (6.11)$$

These normalizations allow us to start with very demanding information-invariance assumptions on the unrestricted domain $\mathcal{U}^E$ and yet have remarkable flexibility in designing social-choice rules if we restrict attention to the profiles respecting our normalizations. We obtain the following theorem (see Blackorby, Bossert and Donaldson, 1999a).

---

[25] See Tungodden (1999) for a discussion of a single norm in combination with ordinally measurable utilities.

Theorem 17: *If a variable-population social-evaluation functional F^E satisfies information invariance with respect to CM and the utility levels representing a neutral life and an excellent life are normalized to zero and one respectively, then the restriction of F^E to $\mathcal{D}^E = \mathcal{U}^E_{\eta\varepsilon}$ satisfies information invariance with respect to NFC.*

Proof. Suppose F^E satisfies information invariance with respect to CM. Let $U \in \mathcal{U}^E_{\eta\varepsilon}$. By definition, $\eta_i(U_i) = 0$ and $\varepsilon_i(U_i) = 1$, and it follows that $\phi_i(0) = 0$ and $\phi_i(1) = 1$ for all $i \in \mathcal{Z}_{++}$. Consequently, $a_i = 0$ and $b_i = 1$ for all $i \in \mathcal{Z}_{++}$. ∎

Note that only cardinal measurability is required in Theorem 17; full interpersonal comparability is provided by the two norms. Therefore, the theorem shows that, if utilities are cardinally measurable and two norms are employed, utilities on the resulting restricted domain are numerically measurable and fully interpersonally comparable. Therefore, the norms generate sufficient additional information to apply any social-evaluation functional. Similar results involving a single norm can be found in Blackorby, Bossert and Donaldson (1999a).

7. Uncertainty

Suppose that a government or an individual must take an action from a set of feasible actions. If the agent knows with certainty the alternative that results from each action, a social-evaluation functional can be used to rank them. In that case, a function f maps actions into alternatives and, for any two actions a and b, action a is at least as good as action b if and only if $f(a)$ is socially no worse than $f(b)$.[26]

In most cases, the consequences of actions are not known with certainty at the time a choice of action has to be made. It may be possible, however, to attach probabilities to the outcomes that may materialize and, in that case, actions can be ranked by ranking prospects. Prospects can be identified with vectors of social alternatives if probabilities are fixed. For such an approach to make normative sense, probabilities may be subjective but must be based on the best information available at the time decisions are taken. If probabilities represent uninformed individual beliefs, the normative force of this approach is weakened substantially.

In any normative investigation, rationality plays an important role and this suggests that both social and individual preferences should satisfy the expected-utility hypothesis (von Neumann and Morgenstern, 1944). Given that, two different versions of welfarism are possible. Ex-ante welfarism bases social evaluations of prospects on individual valuations while ex-post welfarism orders alternatives after the uncertainty has been resolved and aggregates these judgements into a social ordering of prospects.

---

[26] See Broome (1991b) for a discussion.

Harsanyi (1955, 1977) investigates ex-ante welfarism and shows that it has surprising consequences for social evaluation. In his formulation, individuals have ex-ante utility functions that satisfy the Bernoulli hypothesis (Broome, 1991a), a condition that is stronger than the expected-utility hypothesis (see the discussion below).[27] There are $m \geq 2$ 'states of nature' with probabilities $p = (p_1, \ldots, p_m) \in \mathcal{R}_+^m$, $\sum_{j=1}^m p_j = 1$, and they are agreed upon by individuals and by the social evaluator. Individual ex-ante utilities are given by

$$u_i^A = \sum_{j=1}^m p_j u_i^j = \sum_{j=1}^m p_j U_i(x^j) \tag{7.1}$$

for all $i \in \{1, \ldots, n\}$, where u_i^j is individual i's utility in state j and x^j is the social alternative that occurs in state j. There is a single profile of utility functions and x^j is fixed for all $j \in \{1, \ldots, m\}$. Consequently, the utility level u_i^j is fixed for all $i \in \{1, \ldots, n\}$ and all $j \in \{1, \ldots, m\}$. Each probability vector p is called a lottery and all lotteries $p \in \mathcal{R}_+^m$ with $\sum_{j=1}^m p_j = 1$ are permitted. Social ex-ante preferences are represented by

$$u_0^A = \sum_{j=1}^m p_j U_0(x^j) = \sum_{j=1}^m p_j u_0^j \tag{7.2}$$

where u_0^j is social utility in state j and $U_0 \colon X \longrightarrow \mathcal{R}$ is a social utility function. Harsanyi requires social preferences over lotteries to satisfy ex-ante Pareto indifference, which requires society to rank any two lotteries as equally good whenever they are equally valuable for each individual. This axiom alone has the consequence that there exist $\gamma \in \mathcal{R}^n$ and $\delta \in \mathcal{R}$ such that social utilities are weighted sums of individual utilities with $u_0^j = \sum_{i=1}^n \gamma_i u_i^j + \delta$ for all $j \in \{1, \ldots, m\}$ and $u_0^A = \sum_{i=1}^n \gamma_i u_i^A + \delta$. If p and q are any two lotteries, then

$$p \succeq q \iff \sum_{i=1}^n \gamma_i \sum_{j=1}^m p_j u_i^j \geq \sum_{i=1}^n \gamma_i \sum_{j=1}^m q_j u_i^j \tag{7.3}$$

where $p \succeq q$ means that p is socially at least as good as q. This result is called Harsanyi's (1955) social-aggregation theorem.[28] The weights $(\gamma_1, \ldots, \gamma_n)$ in (7.3) are, in general, not unique and need not be positive. The imposition of stronger Pareto conditions implies some restrictions on their signs, however. If strong Pareto is satisfied, (7.3) can be satisfied with positive weights[29] and if weak Pareto is satisfied, (7.3) can be satisfied with non-negative

---

[27] Arrow (1964) provides an account of the expected-utility hypothesis that is consistent with our approach.

[28] Border (1981) presents an elegant proof which is reproduced in expanded form in Weymark (1994). See also Blackorby, Bossert and Donaldson (1998), Blackorby, Donaldson and Weymark (1999a,b), Broome (1990, 1991a), Coulhon and Mongin (1989), Domotor (1979), Fishburn (1984), Hammond (1981, 1983), Mongin (1994, 1995, 1998), and Mongin and d'Aspremont (1998).

[29] See Domotor (1979), De Meyer and Mongin (1995), Weymark (1993) and Zhou (1997).

weights, at least one of which is positive.[30] Because the model employs a single profile of utility functions, application of the anonymity axiom is impossible and, in addition, the weights are not necessarily independent of utilities. For that reason, we do not include a proof.

Instead, we present a variant of Harsanyi's theorem that uses the basic model employed in the rest of this survey. It is a multi-profile model which allows for interpersonal comparisons of utilities and permits the application of the anonymity axiom.[31] X is a set of alternatives with at least four elements. A prospect is a vector $\mathbf{x} = (x^1, \ldots, x^m) \in X^m$ with $m \geq 2$ and the prospect $\mathbf{x}_c = (x, \ldots, x) \in X^m$ is one in which $x \in X$ occurs for certain. The vector of positive probabilities is fixed at $p = (p_1, \ldots, p_m) \in \mathcal{R}^m_{++}$ with $\sum_{j=1}^m p_j = 1$.[32]

As in Harsanyi (1955, 1977), we assume that individual utilities satisfy the Bernoulli hypothesis with the ex-ante utility function $U_i^A \colon X^m \longrightarrow \mathcal{R}$ given by

$$U_i^A(\mathbf{x}) = EU_i(\mathbf{x}) = \sum_{j=1}^m p_j U_i(x^j). \tag{7.4}$$

$U_i^A(\mathbf{x})$ is the value of the prospect $\mathbf{x}$ to person i, $U_i \colon X \longrightarrow \mathcal{R}$ is individual i's von Neumann – Morgenstern (vNM) utility function and $EU_i(\mathbf{x})$ is i's expected utility for prospect $\mathbf{x}$. (7.4) implies that, for all $x \in X$, $U_i^A(\mathbf{x}_c) = EU_i(\mathbf{x}_c) = U_i(x)$. A profile of ex-ante utility functions is $U^A = (U_1^A, \ldots, U_n^A)$ and a profile of vNM utility functions is $U = (U_1, \ldots, U_n)$. Writing $U^A(\mathbf{x}) = (U_1^A(\mathbf{x}), \ldots, U_n^A(\mathbf{x})) = EU(\mathbf{x}) = (EU_1(\mathbf{x}), \ldots, EU_n(\mathbf{x}))$ for all $\mathbf{x} \in X^m$ and $U(\mathbf{x}) = (U_1(x), \ldots, U_n(x))$ for all $x \in X$,

$$U^A(\mathbf{x}) = EU(\mathbf{x}) = \sum_{j=1}^m p_j U(x^j) \tag{7.5}$$

for all $\mathbf{x} \in X^m$. The functions $U_1, \ldots, U_n$ do double duty in this formulation: they are the individuals' vNM utility functions and they measure individual well-being.

An ex-ante social-evaluation functional $F^A \colon \mathcal{D}^A \longrightarrow \mathcal{O}^A$ is a function which maps each profile of ex-ante utility functions into an ordering of X^m. We say that the domain $\mathcal{D}^A$ is the Bernoulli domain $\mathcal{D}_B^A$ if and only if it consists of all profiles of ex-ante utility functions $U^A = (U_1^A, \ldots, U_n^A)$ such that (7.5) is satisfied for some $U = (U_1, \ldots, U_n) \in \mathcal{U}$.

F^A is (ex-ante) welfarist if and only if there exists an ordering $\overset{*}{R}{}^A$ on $\mathcal{R}^n$ such that

$$\mathbf{x}\, R_U^A\, \mathbf{y} \iff \left(U_1^A(\mathbf{x}), \ldots, U_n^A(\mathbf{x})\right) \overset{*}{R}{}^A \left(U_1^A(\mathbf{y}), \ldots, U_n^A(\mathbf{y})\right) \tag{7.6}$$

---

[30] See Weymark (1993, 1994, 1995) and Zhou (1997).

[31] Multi-profile models are presented in Blackorby, Donaldson and Weymark (1999b), Hammond (1981, 1983), Mongin (1994) and Mongin and d'Aspremont (1998).

[32] Because probabilities are assumed to be fixed, any state of nature with a probability of zero may be dropped.

for all $\mathbf{x}$, $\mathbf{y} \in X^m$, where $R_U^A = F^A(U^A)$ is the social ordering of prospects. On the Bernoulli domain, ex-ante welfarism is a consequence of the assumptions binary independence of irrelevant alternatives and Pareto indifference (applied to ex-ante utilities).[33]

Social preferences satisfy the expected-utility hypothesis if and only if there exists a function $U_0 \colon X \times \mathcal{D}^A \longrightarrow \mathcal{R}$ such that, for all $\mathbf{x}$, $\mathbf{y} \in X^m$,

$$\mathbf{x}\, R_U^A\, \mathbf{y} \iff \sum_{j=1}^m p_j U_0\big(x^j, U^A\big) \geq \sum_{j=1}^m p_j U_0\big(y^j, U^A\big). \tag{7.7}$$

Note that this is somewhat weaker than (7.4) because there is no need to measure a social ex-ante utility level. The social vNM function is allowed be profile-dependent. It is possible that it may disregard utility information, in which case a single ordering of X^m is produced. In our multi-profile setting, if the social vNM utility function were written without U^A, an impossibility theorem would result. In Harsanyi's lottery problem, there is only a single profile of vNM utility functions and the second argument of U_0 is not needed. If $R_U^A = F^A(U^A)$ satisfies the expected-utility hypothesis for every $U^A \in \mathcal{D}^A$ we say that the range of F^A is $\mathcal{O}_{EU}^A$.

Now suppose that the domain of the welfarist social-evaluation functional F^A is $\mathcal{D}_B^A$, so that individual utilities satisfy the Bernoulli hypothesis, and its range is $\mathcal{O}_{EU}^A$, so that social preferences satisfy the expected-utility hypothesis. Then it must be true that, for all $\mathbf{x}$, $\mathbf{y} \in X^m$,

$$
\begin{aligned}
\mathbf{x}\, R_U^A\, \mathbf{y} &\iff \sum_{j=1}^m p_j U_0\big(x^j, U^A\big) \geq \sum_{j=1}^m p_j U_0\big(y^j, U^A\big) \\
&\iff \Big(EU_1(\mathbf{x}), \ldots, EU_n(\mathbf{x})\Big)\, \mathring{R}^A \Big(EU_1(\mathbf{y}), \ldots, EU_n(\mathbf{y})\Big) \\
&\iff EU(\mathbf{x})\, \mathring{R}^A\, EU(\mathbf{y}).
\end{aligned}
\tag{7.8}
$$

Setting $\mathbf{x} = \mathbf{x}_c$ and $\mathbf{y} = \mathbf{y}_c$ in (7.8) results in

$$
\begin{aligned}
\mathbf{x}_c\, R_U^A\, \mathbf{y}_c &\iff U_0\big(x, U^A\big) \geq U_0\big(y, U^A\big) \\
&\iff \Big(U_1(x), \ldots, U_n(x)\Big)\, \mathring{R}^A \Big(U_1(y), \ldots, U_n(y)\Big) \\
&\iff U(x)\, \mathring{R}^A\, U(y).
\end{aligned}
\tag{7.9}
$$

This implies that there is a single social-evaluation ordering which is the same for all states, and it is the ex-ante social-evaluation ordering $\mathring{R}^A$. The ordering $\mathring{R}^A$ orders prospects and it also orders alternatives once the uncertainty has been resolved. Such a social-evaluation functional is both ex-ante and ex-post welfarist.

---

[33] See Blackorby, Donaldson and Weymark (1999b), Mongin (1994) and Mongin and d'Aspremont (1998).

Next, we prove a theorem that shows that any welfarist ex-ante social-evaluation functional on the Bernoulli domain whose social preferences satisfy the expected-utility hypothesis must satisfy a property that is equivalent to the requirement that $\overset{*}{R}{}^A$ satisfy information invariance with respect to translation-scale measurability (TSM).[34] In order to find the largest class of functions satisfying our axioms, no information restriction is placed on the social-evaluation functional.

Theorem 18: *Suppose that $|X| \geq 4$. If an ex-ante social-evaluation functional $F^A: \mathcal{D}^A_B \longrightarrow \mathcal{O}^A_{EU}$ is welfarist, then, for all $u, v, a \in \mathcal{R}^n$,*

$$u \overset{*}{R}{}^A v \iff (u + a) \overset{*}{R}{}^A (v + a). \tag{7.10}$$

Proof. For any $u, v, a \in \mathcal{R}^n$, choose a profile $U^A \in \mathcal{D}^A_B$ such that there exist $x, y, z, w \in X$ with $U(x) = u/p_1$, $U(y) = v/p_1$, $U(z) = 0\mathbf{1}_n$ and $U(w) = a/\sum_{j=2}^m p_j$ where U is the vNM profile corresponding to U^A. Consider $\mathbf{x}, \mathbf{y} \in X^m$ with $x^1 = x$, $y^1 = y$ and $x^j = y^j = z$ for all $j \in \{2, \ldots, m\}$. Because $EU(\mathbf{x}) = u$ and $EU(\mathbf{y}) = v$, (7.8) implies

$$u \overset{*}{R}{}^A v \iff p_1 U_0(x, U^A) + \sum_{j=2}^m p_j U_0(z, U^A) \geq p_1 U_0(y, U^A) + \sum_{j=2}^m p_j U_0(z, U^A)$$
$$\iff p_1 U_0(x, U^A) \geq p_1 U_0(y, U^A). \tag{7.11}$$

Now consider $\mathbf{w}, \mathbf{z} \in X^m$ with $w^1 = x$, $z^1 = y$ and $w^j = z^j = w$ for all $j \in \{2, \ldots, m\}$. Because $EU(\mathbf{w}) = (u + a)$ and $EU(\mathbf{z}) = (v + a)$, (7.8) implies

$$(u + a) \overset{*}{R}{}^A (v + a) \iff p_1 U_0(x, U^A) + \sum_{j=2}^m p_j U_0(w, U^A) \geq p_1 U_0(y, U^A) + \sum_{j=2}^m p_j U_0(w, U^A)$$
$$\iff p_1 U_0(x, U^A) \geq p_1 U_0(y, U^A). \tag{7.12}$$

Because the second lines of (7.11) and (7.12) are identical, (7.10) is immediate. ∎

The property described by (7.10) is the same as information invariance with respect to translation-scale measurability and we use the result of Theorem 9 to show that, given anonymity and weak Pareto, $\overset{*}{R}{}^A$ must be the Utilitarian ordering.

---

[34] Mongin (1994) and Mongin and d'Aspremont (1998) prove a similar theorem for lotteries.

Theorem 19: *Suppose that $|X| \geq 4$. An ex-ante social-evaluation functional $F^A \colon \mathcal{D}_B^A \longrightarrow \mathcal{O}_{EU}^A$ is welfarist and satisfies A and WP if and only if, for all $u, v \in \mathcal{R}^n$,*

$$u \overset{*}{R}{}^A v \iff \sum_{i=1}^n u_i \geq \sum_{i=1}^n v_i \tag{7.13}$$

and, for all $\mathbf{x}, \mathbf{y} \in X^m$ and all profiles $U^A \in \mathcal{D}_B^A$,

$$\mathbf{x} R_U^A \mathbf{y} \iff \sum_{i=1}^n EU_i(\mathbf{x}) \geq \sum_{i=1}^n EU_i(\mathbf{y})$$

$$\iff \sum_{i=1}^n \sum_{j=1}^m p_j U_i(x^j) \geq \sum_{i=1}^n \sum_{j=1}^m p_j U_i(y^j). \tag{7.14}$$

Proof. Necessity follows from Theorems 9 and 18. Sufficiency is immediate. ∎

Note that continuity of $\overset{*}{R}{}^A$ is not needed in Theorem 19. Although it is true that the social-evaluation functional is welfarist and the social vNM utility function given by $\overset{\circ}{U}_0(x, U^A) = \sum_{i=1}^n U_i(x)$ satisfies all our axioms, other social vNM utility functions are possible. An example is $U_0(x, U^A) = U_1(z)(\sum_{i=1}^n U_i(x))$ where $z \in X$ is fixed. The inclusion of $U_1(z)$ affects the vNM utility function but not the ordering $\overset{*}{R}{}^A$.

A variant of Theorem 19 can be proved by adding continuity and dropping anonymity. In that case, (7.13) becomes

$$u \overset{*}{R}{}^A v \iff \sum_{i=1}^n \gamma_i u_i \geq \sum_{i=1}^n \gamma_i v_i \tag{7.15}$$

where $\gamma \in \mathcal{R}_+^n \setminus 01_n$. The social-evaluation ordering $\overset{*}{R}{}^A$ is Weighted Utilitarian and the weights are nonnegative with at least one that is positive. Because anonymity is such an important axiom in welfarist social ethics, however, we have presented the theorem that uses it.

An objection that is sometimes made to Harsanyi's theorem is that vNM utility functions are not unique (increasing affine transformations represent the same preferences) and, thus, equal weights on utilities are meaningless. Our framework does not suffer from this difficulty. Because of our assumptions about the measurement and comparability of individual utilities, a particular vNM function is selected for each person. The result of the theorem implies that the information structure (for both vNM utility functions and the ex-ante utility functions $(U_1^A, \ldots, U_n^A)$) must support cardinal unit comparability.

An interesting question of interpretation is whether the theorem provides a convincing argument for Utilitarianism. If the answer is 'yes', it should be noted that the theorem

requires the Bernoulli hypothesis to be satisfied, a stronger requirement than the expected-utility hypothesis. Thus, the utility functions $(U_1, \ldots, U_n)$ must represent people's good, free of the irrationalities of compulsive gambling, for example. It is not usual to present arguments in favour of the Bernoulli hypothesis, over and above the requirements of the expected-utility hypothesis, but it has been done by Broome (1991a). If individual probabilities are subjective and can differ across individuals, it is possible to prove impossibility theorems.[35] Thus, probabilities must be the same for all individuals and for society. This is a demanding requirement but it might be justified by regarding probabilities as 'best-information' probabilities.

Suppose that, instead of the Bernoulli hypothesis, individual ex-ante utilities satisfy the expected-utility hypothesis. In that case, writing U_i^{NM} as individual i's vNM utility function,

$$U_i^A(\mathbf{x}) \geq U_i^A(\mathbf{y}) \iff \sum_{j=1}^m p_j U_i^{NM}(x^j) \geq \sum_{j=1}^m p_j U_i^{NM}(y^j) \qquad (7.16)$$

for all $\mathbf{x}, \mathbf{y} \in X^m$ and all $i \in \{1, \ldots, n\}$. It follows that there exist increasing functions $h_1, \ldots, h_n$ such that, for all $i \in \{1, \ldots, n\}$ and all $\mathbf{x} \in X^m$,

$$U_i^A(\mathbf{x}) = h_i\left(\sum_{j=1}^m p_j U_i^{NM}(x^j) \right). \qquad (7.17)$$

If $\mathbf{x} = \mathbf{x}_c$ in (7.17), $U_i^A(\mathbf{x}_c) = h_i(U_i^{NM}(x))$ and this utility level can be regarded as person i's *actual* utility level when alternative x is realized. Writing U_i as i's actual utility function, $U_i(x) = h_i(U_i^{NM}(x))$ and (7.17) becomes

$$U_i^A(\mathbf{x}) = h_i\left(\sum_{j=1}^m p_j U_i^{NM}(x^j) \right) = h_i\left(\sum_{j=1}^m p_j h_i^{-1}\big(U_i(x^j)\big) \right). \qquad (7.18)$$

The domain of an ex-ante social-evaluation functional is the expected-utility domain $\mathcal{D}_{EU}^A$ if and only if, for each $i \in \{1, \ldots, n\}$, (7.18) is satisfied for some vNM utility function U_i^{NM} and some increasing function h_i. It is possible to show that, on such a domain, no welfarist ex-ante social-evaluation functional exists.

Theorem 20: *Suppose that $|X| \geq 4$. There exists no ex-ante social-evaluation functional $F^A: \mathcal{D}_{EU}^A \longrightarrow \mathcal{O}_{EU}^A$ that is welfarist and satisfies A and WP.*

---

[35] See Hammond (1981, 1983) and Mongin (1995).

Proof. First consider the subdomain $\mathcal{D}_B^A$ of $\mathcal{D}_{EU}^A$ and assume that F^A is welfarist and satisfies A and WP. Theorem 19 implies that, for all $u, v \in \mathcal{R}^n$

$$u \overset{*}{R}{}^A v \iff \sum_{i=1}^n u_i \geq \sum_{i=1}^n v_i. \tag{7.19}$$

Now consider the subdomain $\mathcal{D}_h^A$ of $\mathcal{D}_{EU}^A$ in which, for all $i \in \{1, \ldots, n\}$,

$$U_i^A(\mathbf{x}) = h\left(\sum_{j=1}^m p_j U_i^{NM}(x^j) \right) \tag{7.20}$$

where h is increasing but not affine. Define the ordering $\overset{*}{R}{}_h^A$ on $\mathcal{R}^n$ by

$$u \overset{*}{R}{}_h^A v \iff \left(h(u_1), \ldots, h(u_n) \right) \overset{*}{R}{}^A \left(h(v_1), \ldots, h(v_n) \right) \tag{7.21}$$

for all $u, v \in \mathcal{R}^n$. Writing $EU_i^{NM}(\mathbf{x}) = \sum_{j=1}^m p_j U_i^{NM}(x^j)$ for all $i \in \{1, \ldots, n\}$, we know that

$$
\begin{aligned}
U^A(\mathbf{x}) &\overset{*}{R}{}^A U^A(\mathbf{y}) \\
&\iff \left(h\big(EU_1^{NM}(\mathbf{x})\big), \ldots, h\big(EU_n^{NM}(\mathbf{x})\big) \right) \overset{*}{R}{}^A \left(h\big(EU_1^{NM}(\mathbf{y})\big), \ldots, h\big(EU_n^{NM}(\mathbf{y})\big) \right) \\
&\iff \left(EU_1^{NM}(\mathbf{x}), \ldots, EU_n^{NM}(\mathbf{x}) \right) \overset{*}{R}{}_h^A \left(EU_1^{NM}(\mathbf{y}), \ldots, EU_n^{NM}(\mathbf{y}) \right).
\end{aligned}
\tag{7.22}
$$

Because h is the same for each individual, $\overset{*}{R}{}_h^A$ is anonymous and Theorem 19 implies that, for all $u, v \in \mathcal{R}^n$,

$$u \overset{*}{R}{}_h^A v \iff \sum_{i=1}^n u_i \geq \sum_{i=1}^n v_i. \tag{7.23}$$

Because h is increasing, (7.21) can be rewritten as

$$u \overset{*}{R}{}^A v \iff \left(h^{-1}(u_1), \ldots, h^{-1}(u_n) \right) \overset{*}{R}{}_h^A \left(h^{-1}(v_1), \ldots, h^{-1}(v_n) \right), \tag{7.24}$$

and (7.19), (7.23) and (7.24) together imply

$$u \overset{*}{R}{}^A v \iff \sum_{i=1}^n u_i \geq \sum_{i=1}^n v_i \iff \sum_{i=1}^n h^{-1}(u_i) \geq \sum_{i=1}^n h^{-1}(v_i) \tag{7.25}$$

for all $u, v \in \mathcal{R}^n$. (7.25) can be satisfied if and only if h^{-1} is affine. This requires h to be affine, and a contradiction results. ∎

The proof of Theorem 20 shows that, when the transforms $h_1, \ldots, h_n$ in (7.18) are identical, prospects must be ranked by using the sums of expected utilities. This means, however, that the social-evaluation ordering $\overset{*}{R}{}^A$ must be depend on h, and welfarism, which requires a single social-evaluation ordering, is contradicted. A variant of Theorem 20 can

be proved by dropping anonymity and adding continuity.[36] A way out of this impossibility result can be provided by restricting the domain of the social-evaluation functional to a single information set. If the expected-utility hypothesis is satisfied, Blackorby, Donaldson and Weymark (1999b) prove that, on a single information set, $\overset{*}{R}{}^A$ must be Generalized Utilitarian if the Bernoulli hypothesis is not satisfied and Utilitarian if it is satisfied.

One possible escape from Harsanyi's theorem is to embrace ex-ante welfarism without requiring social preferences to satisfy the expected-utility hypothesis, a move that has been suggested by Diamond (1967) and Sen (1976, 1977b, 1986). They argue that social preferences that satisfy the expected-utility hypothesis cannot take account of the fairness of procedures by which outcomes are generated (see also Weymark, 1991). An ex-ante social-evaluation functional of the type they advocate is given by

$$\mathbf{x}\, R_U^A\, \mathbf{y} \iff \Xi^A\big(U^A(\mathbf{x})\big) \geq \Xi^A\big(U^A(\mathbf{y})\big) \tag{7.26}$$

for all $\mathbf{x}, \mathbf{y} \in X^m$ and all profiles $U^A \in \mathcal{D}^A$. In the equation, Ξ^A is an ex-ante representative-utility function. If the social-evaluation functional satisfies anonymity, Ξ^A must be symmetric. See Weymark (1991) for a discussion.

An important question to consider, however, is whether it is appropriate to require ex-ante welfarism. This form of welfarism is not applied to actual well-being, and that suggests that ex-post welfarism may be more appropriate and ethically more basic. It is true of course that, given the Bernoulli hypothesis, ex-ante welfarism implies ex post, but the converse is not true. A second 'way out' of the result of Theorem 19, therefore, is provided by requiring ex-post welfarism only. Suppose, for example, that Ξ^P is an ex-post representative-utility function, which expresses a social attitude toward utility inequality. Then ex-post welfarism is satisfied by a principle given by

$$\mathbf{x}\, R_U^P\, \mathbf{y} \iff \sum_{j=1}^m p_j U_0^P\left(\Xi^P\big(U(x^j)\big)\right) \geq \sum_{j=1}^m p_j U_0^P\left(\Xi^P\big(U(y^j)\big)\right) \tag{7.27}$$

for all $\mathbf{x}, \mathbf{y} \in X^m$ and all profiles $U \in \mathcal{D}$. In (7.27), U_0^P is a social vNM utility function which expresses a social attitude toward representative-utility uncertainty. Even if society is neutral toward such uncertainty (U_0^P is affine), such a principle is not consistent, in general, with ex-ante Pareto indifference if individual ex-ante utilities satisfy the expected-utility hypothesis. This means that $\mathbf{x}$ may be regarded as better than $\mathbf{y}$ even though the same standard of rationality that is used socially ranks the prospect $\mathbf{y}$ as better for each person. With such a principle, therefore, social rationality trumps individual rationality.

---

[36] See also Blackorby, Donaldson and Weymark (1999a,b), Roemer (1996), Sen (1976) and Weymark (1991).

8. Conclusion

The idea that a just society is a good society can be an attractive one if the good receives an adequate account. Welfarist social-evaluation functionals are capable of performing well as long as the notion of well-being that they employ captures everything of value to individual people. Given that, principles for social evaluation that are non-welfarist are bound to recommend some social changes in which everyone's life is made worse. This is the lesson of Theorem 2.

If a welfarist principle is to be used to rank alternatives which are complete histories of the world, a single social-evaluation ordering is sufficient to do the job for every profile of utility functions. Although such orderings can be used to rank changes which affect a population subgroup, such as the citizens of a single country or the people of a particular generation, the induced ordering over their utilities is not, in general, independent of the utilities of others. Independence is guaranteed by the axiom independence of the utilities of unconcerned individuals and, in conjunction with continuity, anonymity and strong Pareto, that axiom results in Generalized Utilitarianism. In a dynamic framework, the same result is the consequence of independence of the utilities of the dead together with intertemporal versions of continuity, anonymity and strong Pareto.

Generalized-Utilitarian social-evaluation functionals are ethically attractive, but some of them may require utility information that is difficult to acquire. In parsimonious information environments, Utilitarianism itself may prove to be more attractive than the other members of that class of orderings. The only Generalized-Utilitarian social-evaluation ordering that satisfies information invariance with respect to cardinal full comparability is Utilitarianism. And if individual utilities are translation-scale measurable, anonymity and weak Pareto alone imply that the social-welfare ordering must be Utilitarian. Information restrictions are not the only axioms that generate Utilitarianism, however. Incremental equity is an axiom that requires a kind of impartiality with respect to utility increases or decreases. If one person's utility increases, the axiom requires the change to be equally good no matter who the fortunate person is. This axiom, together with weak Pareto, characterizes Utilitarianism.

The Utilitarian and Generalized-Utilitarian social-evaluation functionals can be extended to environments in which population size and composition may vary across alternatives. Two properties can be considered particularly desirable in this framework: extended independence of the utilities of unconcerned individuals and avoidance of the repugnant conclusion. Given continuity, anonymity and strong Pareto, extended independence of the utilities of unconcerned individuals implies that the social-evaluation ordering must be Critical-Level Generalized Utilitarian. The critical level is a parameter that represents the smallest utility level above which additions to a utility-unaffected population have value. The repugnant conclusion is avoided if and only if the critical level is above neutrality.

If utilities are cardinally measurable, interpersonal comparisons at any two norms are sufficient to produce numerical full comparability. We might choose norms, for example, at neutrality and at a utility level that represents a satisfactory or excellent life. Utility numbers such as zero and one may be chosen for these, and NFC results. It follows that cardinal measurability and two norms are sufficient to employ any welfarist social-evaluation functional.

Social-evaluation functionals can be extended to rank prospects as long as probabilities can be attached to the various states of nature. If individual ex-ante utilities satisfy the Bernoulli hypothesis, social preferences satisfy the expected-utility hypothesis and all subjective probabilities coincide, the only ex-ante social-evaluation functional that satisfies anonymity and weak Pareto is the Utilitarian one. If, however, the domain is expanded to include all individual ex-ante utility functions that satisfy the expected-utility hypothesis or if subjective probabilities can be different for different people, an impossibility results. If the von Neumann – Morgenstern utility functions, in addition to representing people's good, express an attitude toward uncertainty that is rational and has some normative standing, it can be argued that the Bernoulli hypothesis is a reasonable assumption. In that case, Harsanyi's social-aggregation theorem provides support for Utilitarianism.

Together, these results make a strong case for Utilitarian and Generalized-Utilitarian social evaluation. When these social-evaluation functionals are coupled with an adequate account of lifetime well-being, the resulting principles are ethically attractive and perform well in environments in which other principles perform poorly. Social-contract theories, for example, are not able to give an adequate account of justice between generations when the existence of people in one generation is contingent on decisions made by another. On the other hand, the Critical-Level Generalized-Utilitarian principles can cope with fully dynamic environments in which history has a branching structure and the identities of those alive, their numbers, quality of life and length of life can vary across alternatives.

REFERENCES

Aczél, J. (1966), Lectures on Functional Equations and Their Applications (Academic Press, New York).

Anderson, E. (1999), "What is the point of equality?", Ethics 109:287–337.

Arneson, R. (1989), "Equality of opportunity for welfare", Philosophical Studies 56:77–93.

Arrhenius, G. (1997), "An impossibility theorem for welfarist axiologies", unpublished manuscript, University of Toronto, Department of Philosophy.

Arrow, K. (1951, second ed. 1963), Social Choice and Individual Values (Wiley, New York).

Arrow, K. (1964), "The role of securities in the optimal allocation of risk-bearing", Review of Economic Studies 31:91–96.

Atkinson, A. (1970), "On the measurement of inequality", Journal of Economic Theory 2:244–263.

Basu, K. (1983), "Cardinal utility, utilitarianism, and a class of invariance axioms in welfare analysis", Journal of Mathematical Economics 12:193–206.

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 (1995), "Intertemporal population ethics: critical-level utilitarian principles", Econometrica 63:1303–1320.

Blackorby, C., W. Bossert and D. Donaldson (1996), "Intertemporally consistent population ethics: classical utilitarian principles", in: K. Arrow, A. Sen and K. Suzumura, eds., Social Choice Re-Examined, vol. 2 (Macmillan, London) 137–162.

Blackorby, C., W. Bossert and D. Donaldson (1997a), "Birth-date dependent population ethics: critical-level principles", Journal of Economic Theory 77:260–284.

Blackorby, C., W. Bossert and D. Donaldson (1997b), "Intertemporally consistent population ethics: birth-date dependent classical principles", Japanese Economic Review 48:267–292.

Blackorby, C., W. Bossert and D. Donaldson (1998), "Uncertainty and critical-level population principles", Journal of Population Economics 11:1–20.

Blackorby, C., W. Bossert and D. Donaldson (1999a), "Information invariance in variable-population social-choice problems", International Economic Review 40:403–422.

Blackorby, C., W. Bossert and D. Donaldson (1999b), "Functional equations and population ethics", Aequationes Mathematicae, forthcoming.

Blackorby, C., W. Bossert and D. Donaldson and M. Fleurbaey (1998), "Critical levels and the (reverse) repugnant conclusion", Journal of Economics 67:1–15.

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

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 (1991), "Normative population theory: a comment", Social Choice and Welfare 8:261–267.

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 (1990), "A welfarist proof of Arrow's theorem", Recherches Economique de Louvain 56:259–286.

Blackorby, C., D. Donaldson and J. Weymark (1999a), "Harsanyi's social aggregation theorem for state-contingent alternatives", Journal of Mathematical Economics, forthcoming.

Blackorby, C., D. Donaldson and J. Weymark (1999b), "Social aggregation and the expected utility hypothesis", in: M. Salles and J. Weymark, eds., Justice, Political Liberalism, and Utilitarianism: Themes from Harsanyi and Rawls (Cambridge University Press, Cambridge) forthcoming.

Blackorby, C., D. Primont, D. and R. Russell (1978), Duality, Separability, and Functional Structure: Theory and Economic Applications (North-Holland, Amsterdam).

Blackwell, D. and M. Girshick (1954), Theory of Games and Statistical Decisions (Wiley, New York).

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

Border, K. (1981), "Notes on von Neumann – Morgenstern social welfare functions", unpublished manuscript, California Institute of Technology, Department of Economics.

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. (1990a), "An axiomatization of the single-series Ginis", Journal of Economic Theory 50:82–92.

Bossert, W. (1990b), "Maximin welfare orderings with variable population size", Social Choice and Welfare 7:39–45.

Bossert, W. (1990c), "Social evaluation with variable population size: an alternative concept", Mathematical Social Sciences 19:143–158.

Bossert, W. (1991), "On intra- and interpersonal utility comparisons", Social Choice and Welfare 8:207–219.

Bossert, W. (1999), "Welfarism and information invariance", Social Choice and Welfare, forthcoming.

Bossert, W. and F. Stehling (1992), "A remark on admissible transformations for interpersonally comparable utilities", International Economic Review 33:739–744.

Bossert, W. and F. Stehling (1994), "On the uniqueness of cardinally interpreted utility functions", in: W. Eichhorn, ed., Models and Measurement of Welfare and Inequality (Springer, Berlin) 537–551.

Bossert, W. and J. Weymark (1999), "Utility in social choice", in: S. Barberà, P. Hammond and C. Seidl, eds., Handbook of Utility Theory, vol. 2: Extensions (Kluwer, Dordrecht) forthcoming.

Broome, J. (1990), "Bolker-Jeffrey expected utility theory and axiomatic utilitarianism", Review of Economic Studies 57:477–502.

Broome, J. (1991a), Weighing Goods (Basil Blackwell, Oxford).

Broome, J. (1991b), "The structure of good: decision theory and ethics", in: M. Bacharach and S. Hurley, eds., Foundations of Decision Theory (Basil Blackwell, Oxford) 123–146.

Broome, J. (1993), "Goodness is reducible to betterness: the evil of death is the value of life", in: P. Koslowski, ed., The Good and the Economical: Ethical Choices in Economics and Management (Springer, Berlin) 69–83.

Carlson, E. (1998), "Mere addition and the two trilemmas of population ethics", Economics and Philosophy 14:283–306.

Coulhon, T. and P. Mongin (1989), "Social choice theory in the case of von Neumann – Morgenstern utilities", Social Choice and Welfare 6:175–187.

Dalton, H. (1920), "The measurement of the inequality of incomes", Economic Journal 30:348–361.

Dasgupta, P. (1993), An Inquiry into Well-Being and Destitution (Clarendon, Oxford).

d'Aspremont, C. (1985), "Axioms for social welfare orderings", in: L. Hurwicz, D. Schmeidler and H. Sonnenschein, eds., Social Goals and Social Organizations: Essays in Memory of Elisha Pazner (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 (1999), "Interpersonal comparability of welfare and social choice theory", in: K. Arrow, A. Sen and K. Suzumura, eds., Handbook of Social Choice and Welfare (Elsevier, Amsterdam) forthcoming.

Debreu, G. (1954), "Representation of a preference ordering by a numerical function", in: R. Thrall, C. Coombs and R. Davis, eds., Decision Processes (Wiley, New York) 159–166.

Debreu, G. (1960), "Topological methods in cardinal utility theory", in: K. Arrow, S. Karlin and P. Suppes, eds., Mathematical Methods in the Social Sciences, 1959 (Stanford University Press, Stanford) 16–26.

De Meyer, F. and C. Plott (1971), "A welfare function using 'relative intensity' of preference", Quarterly Journal of Economics 85:179–186.

De Meyer, F, and P. Mongin (1995), "A note on affine aggregation", Economics Letters 47:177–183.

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

Diamond, P. (1967), "Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility: comment", Journal of Political Economy 75:765–766.

Dixit, A. (1980), "Interpersonal comparisons and social welfare functions", unpublished manuscript, University of Warwick, Department of Economics.

Domotor, Z. (1979), "Ordered sum and tensor product of linear utility structures", Theory and Decision 11:375–399.

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

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

Falmagne, J.-C. (1981), "On a recurrent misuse of a classical functional equation result", Journal of Mathematical Psychology 23:190–193.

Fishburn, P. (1984), "On Harsanyi's utilitarian cardinal welfare theorem", Theory and Decision 17:21–28.

Fishburn, P., H. Marcus-Roberts and F. Roberts (1988), "Unique finite difference measurement", SIAM Journal of Discrete Mathematics 1:334–354.

Fishburn, P. and F. Roberts (1989), "Uniqueness in finite measurement", in: F. Roberts, ed., Applications of Combinatorics and Graph Theory to the Biological and Social Sciences (Springer, New York) 103–137.

Fleming, M. (1952), "A cardinal concept of welfare", Quarterly Journal of Economics 66:366–384.

Gevers, L. (1979), "On interpersonal comparability and social welfare orderings", Econometrica 47:75–89.

Gorman, W. (1968), "The structure of utility functions", Review of Economic Studies 32:369–390.

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. (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.

Hammond, P. (1981), "Ex-ante and ex-post welfare optimality under uncertainty", Economica 48:235–250.

Hammond, P. (1983), "Ex-post optimality as a dynamically consistent objective for collective choice under uncertainty", in: P. Pattanaik and M. Salles, eds., Social Choice and Welfare (North-Holland, Amsterdam) 175–205.

Hammond, P. (1988), "Consequentialist demographic norms and parenting rights", Social Choice and Welfare 5:127–146.

Hammond, P. (1999), "Interpersonally comparable utility", in: S. Barberà, P. Hammond and C. Seidl, eds., Handbook of Utility Theory, vol. 2: Extensions (Kluwer, Dordrecht) forthcoming.

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

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).

Hurka, T. (1982), "Average utilitarianism", Analysis 42:65–69.

Kolm, S.-C. (1969), "The optimal production of social justice", in: J. Margolis and S. Guitton, eds., Public Economics (Macmillan, London) 145–200.

Krantz, D., R. Luce, P. Suppes and A. Tversky (1971), Foundations of Measurement, vol. I: Additive and Polynomial Representations (Academic Press, New York).

Maskin, E. (1978), "A theorem on utilitarianism", Review of Economic Studies 45:93–96.

Mill, J. (1969, originally published in 1861), "Utilitarianism", in: Collected Works of John Stuart Mill, vol. X (University of Toronto Press, Toronto) 203–259.

Milnor, J. (1954), "Games against nature", in: R. Thrall, C. Coombs and R. Davis, eds., Decision Processes (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.

Mongin, P. (1995), "Consistent Bayesian aggregation", Journal of Economic Theory 66:313–351.

Mongin, P. (1998), "The paradox of the Bayesian experts and state-dependent utility theory", Journal of Mathematical Economics 29:331–361.

Mongin, P. (1999), "State-of-nature and impartial-observer theories of justice", in: K. Arrow, A. Sen and K. Suzumura, eds., Handbook of Social Choice and Welfare (Elsevier, Amsterdam) forthcoming.

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

Murphy, L. (1998), "Institutions and the demands of justice", Philosophy and Public Affairs 27:251–291.

Ng, Y.-K. (1989), "What should we do about future generations? Impossibility of Parfit's Theory X", Economics and Philosophy 5:235–253.

Parfit, D. (1976), "On doing the best for our children", in: M. Bayles, ed., Ethics and Population (Schenkman, Cambridge) 100–102.

Parfit, D. (1982), "Future generations, further problems", Philosophy and Public Affairs 11:113–172.

Parfit, D. (1984), Reasons and Persons (Oxford University Press, Oxford).

Rawls, J. (1971), A Theory of Justice (Harvard University Press, Cambridge).

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.

Roemer, J. (1996), Theories of Distributive Justice (Harvard University Press, Cambridge).

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

Sen, A. (1973), On Economic Inequality (Oxford University Press, Oxford).

Sen, A. (1974), "Informational bases of alternative welfare approaches: aggregation and income distribution", Journal of Public Economics 3:387–403.

Sen, A. (1976), "Welfare inequalities and Rawlsian axiomatics", Theory and Decision 7:243–262.

Sen, A. (1977a), "On weights and measures: informational constraints in social welfare analysis", Econometrica 45:1539–1572.

Sen, A. (1977b), "Non-linear social welfare functions: a reply to Professor Harsanyi", in: R. Butts and J. Hintikka, eds., Foundational Problems in the Special Sciences (Reidel, Dordrecht) 297–302.

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. (1986), "Social choice theory", in: K. Arrow and M. Intriligator, eds., Handbook of Mathematical Economics, vol. III (North-Holland, Amsterdam) 1073–1181.

Sidgwick, H. (1966, originally published in 1907), The Methods of Ethics (Dover, New York).

Sikora, R. (1978), "Is it wrong to prevent the existence of future generations?", in: R. Sikora and B. Barry, eds., Obligations to Future Generations (Temple, Philadelphia) 112–166.

Sumner, L. (1996), Welfare, Happiness, and Ethics (Clarendon, Oxford).

Tungodden, B. (1999), "Social choices with independent norm levels", unpublished manuscript, Norwegian School of Economics and Business Administration.

von Neumann, J. and O. Morgenstern (1944), Theory of Games and Economic Behavior (Princeton University Press, Princeton).

Weymark, J. (1981), "Generalized Gini inequality indices", Mathematical Social Sciences 1:409–430.

Weymark, J. (1991), "A reconsideration of the Harsanyi-Sen debate on utilitarianism", in: J. Elster and J. Roemer, eds., Interpersonal Comparisons of Well-Being (Cambridge University Press) 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: W. Eichhorn, ed., Models and Measurement of Welfare and Inequality (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.

Weymark, J. (1998), "Welfarism on economic domains", Mathematical Social Sciences 36:251–268.

Zhou, L. (1997), "Harsanyi's utilitarianism theorems: general societies", Journal of Economic Theory 72:198–207.

■   ■   ■