The Value of Limited Altruism*

by

Charles Blackorby, Walter Bossert, and David Donaldson

August 1998

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

* We thank Peter Vallentyne for comments and criticisms. Research support from the Social Sciences and Humanities Research Council of Canada is gratefully acknowledged.

August 20, 1998

Abstract

Discounting the utilities of future people or giving smaller weights to groups other than one's own is often criticized on the grounds that the resulting objective function differs from the ethically appropriate one. This paper investigates the consequences of changes in the discount factor and weights when they are moved toward the warranted ones. Using the utilitarian value function, it is shown that, except in restrictive special cases, those moves do not necessarily lead to social improvements. We suggest that limitations to altruism are better captured by maximizing the appropriate value function subject to lower bounds on some utilities. *Journal of Economic Literature* Classification Nos.: D63, D71.

Keywords: Social Choice, Altruism, Consequentialism.

The Value of Limited Altruism

by

Charles Blackorby, Walter Bossert, and David Donaldson

1. Introduction

When economists investigate policies that affect the well-being of several generations, they often employ objective functions that discount the well-being (utility) of future people. This common practice has been criticized by many, either on the grounds that the implied ethics are not impartial, or on the grounds that the absence of discounting is a consequence of the axiom Pareto indifference.[1] Alternatively, some argue that, although discounting is ethically warranted, the levels chosen are too high.[2] If these views are correct, a natural recommendation is to use the appropriate value function. If, however, social decision-makers are unwilling to do that, a response might be that moving the discount factor closer to one would have good consequences.

A justification for using an objective function that discounts the utilities of future generations at a higher rate than the one that is appropriate is that the best policies, according to the value function, may require very great sacrifices by the present generation. Disagreements about the proper level of discounting may therefore take two forms: in the first, the disagreement is about the appropriate value function; in the second, it is about the level of altruism on the part of the present generation that can legitimately be expected.

In this paper, we ask whether moving the discount rate toward the one that is ethically warranted results in better policies. Suppose that the appropriate value function is the undiscounted sum of utilities.[3] It can be used to assess the merits of choices made with other objective functions. Thus, the objective-function-maximizing policies that correspond to different discount rates may be ranked by comparing total utilities in the states of affairs that arise.

[1] See Blackorby, Bossert, and Donaldson [1995, 1996, 1997a,b], Broome [1992], Cowen [1992], and Cowen and Parfit [1992].

[2] There is a justification for discounting changes in incomes and other net benefits measured in monetary terms (see, for example, Sugden and Williams [1978]), although most cost-benefit analysts argue that market rates are inappropriate and lower 'social' rates should be used when projects affect several generations (Sugden and Williams [1978, ch. 15]).

[3] Throughout the paper, we assume that there is a finite number of generations, so convergence problems with infinite horizons do not arise.

In moral theory, traditional consequentialism requires agents to choose the actions that have the best consequences according to some theory which ranks states of affairs with respect to their goodness.[4] These theories should rank states timelessly rather than in a single period (see Blackorby, Bossert, and Donaldson [1996, 1997b] and Blackorby and Donaldson [1984]) and, if they are based exclusively on individual well-being, should employ comprehensive notions of welfare such as the ones investigated by Griffin [1986] and Sumner [1996].

Traditional consequentialism may demand, however, that individuals make very great sacrifices, such as giving a large percentage of their incomes to malnourished people. It is commonly argued in moral theory that "'ought' implies 'can'" (Griffin [1996, pp. 89–92]), and that actions that are regarded as obligatory should take account of the natural limits of individual altruism. Accordingly, some actions may be regarded as obligatory, while others are 'supererogatory'—beyond the call of duty.[5] Supererogatory actions are morally permissible and have better consequences than obligatory ones, but involve greater sacrifices by the agent.

Because of the limits of individual altruism, moral agents may use an objective function to guide their actions that is different from the one that gives everybody's well-being equal weight. One such suggestion has been made by Scheffler [1982],[6] who advocates an objective function that gives a greater weight to the well-being of the moral agent than to others. More generally, consider an objective function that gives the highest weight to the agent's own well-being, a lesser weight to the well-being of his or her family members, a still smaller weight to the utilities of friends, and so on. In such a setting, it is important to ask whether moving the weights closer to the ethically correct ones will result in actions that have better consequences.

We employ a simple model to investigate the problem. A fixed population is divided into at least two groups. These may be different generations or the result of any other partition of the population, and each group may contain a single person or more. Some members of group one or a single person in the group must choose to take an action from a set of feasible actions. We refer to the agent as 'group one' for convenience. Each of the actions has consequences which, for the normative purposes of this paper, can be described as vectors of utilities for the individuals, including the people in group one. It need not be the case that the actions of others are fixed. All that is necessary is that they be predictable—the agent may take actions that elicit cooperative behaviour by others, for example. We do not consider strategic behaviour.

[4] Although the theories that we employ in this paper are welfarist, it is possible to combine consequentialist morality with any principle that provides social rankings.

[5] See Griffin [1996] and Heyd [1982].

[6] See also Mulgan [1997].

In Sections 2 and 3, we assume that the value function that ranks alternative states of affairs according to their goodness is the utilitarian one—the sum of individual utilities. If group one acts to maximize the utilitarian value function, giving equal weight to its own interests and the interests of others, we say that it is perfectly altruistic. Imperfect or limited altruism is modeled by considering two different objective functions, each of which is used to guide the behaviour of group one. The first of these is applicable to an intertemporal setting and it discounts the utilities of the members of other groups geometrically. The second applies weights to the utilities of other groups which are smaller than the weight assigned to the members of group one. The second formulation therefore contains the first as a special case. Sections 2 and 3 differ in the way the set of feasible actions facing group one is described. In each case, however, the same question is asked. Does a reduction in the discount rate or an increase in the weight assigned to the utilities of groups other than the first lead to better choices? That is, do the actions taken lead to consequences with greater total utility?

Section 2 is concerned with a very simple case, the pure distribution problem. A fixed amount of a single resource is to be distributed to the people in the groups, and group one's set of feasible actions corresponds to all divisions of the resource. In the discounting case, we show that a decrease in the discount rate always has good consequences—total utility rises. In the case of weights, we show that (i) a decrease in the weight on the utilities of the members of group one leads to an increase in total utility, (ii) an increase in the smallest weight leads to greater total utility, (iii) increases in other weights have ambiguous consequences, and (iv) some altruism (positive weights for at least one group other than group one) is better than none (self-interested behaviour by group one).

Section 3 considers the general choice problem. Group one has an arbitrary set of feasible actions to which correspond a set of feasible utility vectors. We show that, with three or more groups, a decrease in the discount rate, a decrease in the weight for group one, and an increase in the weight for any other group have ambiguous consequences: total utility may rise, fall, or remain unchanged. In addition, we show that, although it is true that utilitarian consequentialist behaviour is always best, some altruism may be worse than none at all. If, however, there are only two groups, then decreasing the weight on group one or increasing the weight on group two (which are both equivalent to decreasing the discount factor) have good consequences and, as before, some altruism is always better than none.

In Section 4, we propose another way to describe objective functions for agents whose altruism is limited. Instead of placing weights on the well-being of others, we suggest that utility constraints be used. The constraints place lower limits on the total utility of group one and, possibly, some of the other groups. Group one then acts to maximize the utilitarian value function subject to the constraints. In this formulation, the relaxation of any constraint (moving it to a lower level) never leads to worse actions and may lead to better ones. We show that the two approaches are equivalent in the two-group case.

Section 5 discusses a number of variations on our results. First, they are extended to generalized-utilitarian consequentialism, which allows inequality aversion in utilities. Second, our results are shown to be robust to modifications of the utilitarian and generalized utilitarian principles which discount the well-being of the members of future generations. Third, the results of Section 2 on the pure distribution problem are extended to instances of the general choice problem in which the set of feasible utility vectors associated with feasible actions is described by an additively separable function. Last, we show that all our results can be extended to actions whose consequences are uncertain. In this case, *ex ante* social evaluations are employed.[7] Section 6 concludes.

Throughout the paper, we assume that utilities satisfy the necessary measurability and comparability requirements needed for the value and objective functions employed. In the utilitarian cases of Sections 2, 3, and 4, utilities must be at least cardinally measurable and unit comparable. The generalized utilitarian principles have more stringent measurability and comparability requirements.[8]

2. The Pure Distribution Problem

In this section, we investigate limited altruism in a very simple setting, the pure distribution problem. A given amount of a single resource is to be divided among the members of a fixed set of people. The amount of the resource is $\omega \in \mathcal{R}_{++}$, the set of individuals is $\{1, \ldots, n\}$, $n \geq 2$, and the utility function of person $i \in \{1, \ldots, n\}$ is $U_i : \mathcal{R}_+ \longrightarrow \mathcal{R}$, with

$$u_i = U_i(x_i), \tag{2.1}$$

where $x_i \in \mathcal{R}_+$ is person i's consumption of the resource. We assume that, for each i, U_i is continuous, twice continuously differentiable on $\mathcal{R}_{++}$, and strongly concave (Diewert, Avriel, and Zang [1981]), which implies $U_i''(x_i) < 0$ for all $x_i \in \mathcal{R}_{++}$. We assume $U_i'(x_i) > 0$ for all $x_i \in \mathcal{R}_{++}$ and, in addition, that $\lim_{x_i \to 0} U_i'(x_i) = \infty$; this ensures that each person receives a positive consumption level whenever the weight on his or her utility is positive.[9]

[7] For discussions of *ex ante* and *ex post* social evaluation, see Blackorby, Donaldson, and Weymark [1996, 1997], Broome [1991], and Mongin and d'Aspremont [1998].

[8] See Blackorby, Donaldson, and Weymark [1984], Bossert [1991], Bossert and Weymark [1998], d'Aspremont and Gevers [1977], Roberts [1980a,b], and Sen [1974, 1977].

[9] The utility function U_i could, instead, be defined on the interval $[s_i, \infty)$ where s_i is person i's subsistence level of consumption. The condition $\lim_{x_i \to s_i} U_i'(x_i) = \infty$ would replace the one in the text and, if $\sum_{i=1}^n s_i < \omega$, would ensure that each person receives a consumption level that is greater than subsistence. Because this complicates the presentation without enhancing our understanding (all results are unaffected), we have chosen to work with the simpler model.

The utilitarian solution to the pure distribution problem is $(\overset{*}{x}_1, \ldots, \overset{*}{x}_n)$, and it maximizes $\sum_{i=1}^{n} U_i(x_i)$ subject to $\sum_{i=1}^{n} x_i \leq \omega$. The first-order conditions are

$$U_i'(\overset{*}{x}_i) = \overset{*}{\lambda} \tag{2.2}$$

for all $i = 1, \ldots, n$, where λ is a Lagrange multiplier, and

$$\sum_{i=1}^{n} \overset{*}{x}_i = \omega. \tag{2.3}$$

(2.2) requires the resource to be distributed so that marginal utilities are the same for everyone, and (2.3) requires the whole amount of the resource to be distributed.

In the following subsections, we assume that group one has a set of actions from which it can choose and that, for every feasible allocation of the resource, there is an action which leads to that allocation. Perfect altruism results in an action which maximizes total utility. Limited altruism is characterized by assuming that group one acts to maximize an objective function with utility weights on groups of individuals. In this case, group one receives a higher weight than the others, and at least one other group has a positive weight. Self-interested behaviour gives a weight of zero to the utilities of all groups other than the first. The groups are $N_1, \ldots, N_m$, $m \geq 2$, with at least one person in each, and they form a partition of $\{1, \ldots, n\}$—each person is in exactly one group.

It is convenient to deal with the maximization problems that follow in two stages. Because, in each problem, the utilities of the members of a group receive the same weight, the allocation of resources within the groups must maximize $\sum_{i \in N_j} U_i(x_i)$ subject to $\sum_{i \in N_j} x_i \leq z_j$, where z_j is total consumption of group j, $j = 1, \ldots, m$. The functions $V_j\colon \mathcal{R}_+ \longrightarrow \mathcal{R}$, $j = 1, \ldots, m$, are given by

$$V_j(z_j) = \max_{(x_i)_{i \in N_j}} \left\{ \sum_{i \in N_j} U_i(x_i) \;\middle|\; \sum_{i \in N_j} x_i \leq z_j \right\}. \tag{2.4}$$

Given our assumptions, each V_j is continuous, twice continuously differentiable on $\mathcal{R}_{++}$, $V_j'(z_j) > 0$ and $V_j''(z_j) < 0$ for all $z_j \in \mathcal{R}_{++}$, and $\lim_{z_j \to 0} V_j'(z_j) = \infty$ (see Lemma 1 in the Appendix).

It follows that the utilitarian solution maximizes

$$\sum_{j=1}^{m} V_j(z_j) \tag{2.5}$$

subject to

$$\sum_{j=1}^{m} z_j \leq \omega, \tag{2.6}$$

and we write the solution as $(\overset{*}{z}_1, \ldots, \overset{*}{z}_m)$.

2.1. Geometric Discounting

Suppose that group one is the present generation and that, although the social value function is the utilitarian one, group one's altruism is limited and, in deciding on its actions, it uses an objective function that discounts the utilities of future generations geometrically. Let $\delta \in (0,1)$ be the discount factor. Generation one acts to maximize

$$\sum_{j=1}^{m} \delta^{j-1} \sum_{i \in N_j} U_i(x_i) \tag{2.7}$$

subject to

$$\sum_{i=1}^{n} x_i = \sum_{j=1}^{m} \sum_{i \in N_j} x_i \leq \omega. \tag{2.8}$$

Using the functions $V_1, \ldots, V_m$, this can be converted into a maximization problem which assigns resources to groups. Group one must choose $(z_1, \ldots, z_m)$ to maximize

$$\sum_{j=1}^{m} \delta^{j-1} V_j(z_j) \tag{2.9}$$

subject to the constraint

$$\sum_{j=1}^{m} z_j \leq \omega. \tag{2.10}$$

If $(\mathring{x}_1, \ldots, \mathring{x}_n)$ solves the maximization problem of (2.7) and (2.8), and the solution to the maximization problem of (2.9) and (2.10) is $(\mathring{z}_1, \ldots, \mathring{z}_m)$, then

$$\mathring{z}_j = \sum_{i \in N_j} \mathring{x}_i \tag{2.11}$$

for all $j = 1, \ldots, m$. The utilitarian social value of the solution is

$$\sum_{i=1}^{n} U_i(\mathring{x}_i) = \sum_{j=1}^{m} V_j(\mathring{z}_j), \tag{2.12}$$

the undiscounted sum of utilities.

According to the utilitarian value function, the no-discounting solution to the maximization problem is better than the solution with discounting. The question we ask, however, is the following: if group one were to discount future utilities less, would total (undiscounted) utility at the chosen action rise? That is, if $\tilde{\delta} > \hat{\delta}$, and $(\tilde{x}_1, \ldots, \tilde{x}_n)$ and $(\hat{x}_1, \ldots, \hat{x}_n)$ are the corresponding solutions to (2.7) and (2.8), is it true that

$$\sum_{i=1}^{n} U_i(\tilde{x}_i) > \sum_{i=1}^{n} U_i(\hat{x}_i)? \tag{2.13}$$

This question is equivalent to asking whether

$$\sum_{j=1}^{m} V_j(\tilde{z}_j) > \sum_{j=1}^{m} V_j(\hat{z}_j), \tag{2.14}$$

where $(\tilde{z}_1, \ldots, \tilde{z}_m)$ and $(\hat{z}_1, \ldots, \hat{z}_m)$ are the optimal group consumption levels corresponding to $\tilde{\delta}$ and $\hat{\delta}$. In the case of the pure distribution problem, the answer is yes.

Theorem 1: *In the pure distribution problem with discounting, if $\tilde{\delta} > \hat{\delta}$, then*

$$\sum_{i=1}^{n} U_i(\tilde{x}_i) > \sum_{i=1}^{n} U_i(\hat{x}_i) \tag{2.15}$$

or, equivalently,

$$\sum_{j=1}^{m} V_j(\tilde{z}_j) > \sum_{j=1}^{m} V_j(\hat{z}_j). \tag{2.16}$$

Proof: See the Appendix.

The result of Theorem 1 agrees with the standard intuition of those who regard discounting as inappropriate: less of it is always better. We shall see, however, that this intuition is not robust.

2.2. Weighting Schemes

Suppose, as in the previous subsection, that group one's actions determine the distribution of the resource and that, instead of employing an objective function that discounts the utilities of other groups geometrically, it attaches weights to their utilities which are smaller than the weight it gives itself. This might occur if group one were an agent's family, group two the members of his or her community, and so on. The optimization problem solved by group one is to maximize

$$\sum_{j=1}^{m} \gamma_j \sum_{i \in N_j} U_i(x_i) \tag{2.17}$$

subject to

$$\sum_{i=1}^{n} x_i \leq \omega, \tag{2.18}$$

or, equivalently, to maximize

$$\sum_{j=1}^{m} \gamma_j V_j(z_j) \tag{2.19}$$

subject to

$$\sum_{j=1}^{m} z_j \leq \omega. \tag{2.20}$$

If group one is perfectly altruistic, all the weights are equal and positive. If group one's altruism is limited, $\gamma_1 > \gamma_j \geq 0$ for all $j = 2, \ldots, m$ with $\gamma_j > 0$ for at least one $j > 1$. Because any two groups with the same weight can be combined, we can assume that the group weights are pairwise distinct. Without loss of generality, therefore, we number the groups so that $\gamma_{j-1} > \gamma_j$ for all $j = 2, \ldots, m$. If group one's behaviour is self interested, $\gamma_1 > 0$ and $\gamma_j = 0$ for all $j = 2, \ldots, m$.

In the case of limited altruism with $\gamma_m > 0$, the first-order conditions are

$$\gamma_j V_j'(\bar{z}_j) = \bar{\lambda} \tag{2.21}$$

for all $j = 1, \ldots, m$, and

$$\sum_{j=1}^{m} \bar{z}_j = \omega. \tag{2.22}$$

If $\gamma_m = 0$, (2.21) holds for all $j \neq m$ and $\bar{z}_m = 0$. (2.21) shows that, in this solution, marginal utilities are unequal; group one has the lowest and group m the highest. Consequently, from the utilitarian point of view, transfers from groups with higher weights to groups with lower weights are warranted.

Theorem 2 proves that, in the case of limited altruism, an increase in γ_m, the smallest weight, or a decrease in γ_1, the largest weight, leads to an increase in total utility. Thus, from the (unweighted) utilitarian standpoint, such changes have good consequences.

Theorem 2: *Given limited altruism, in the pure distribution problem with weights, if (i) $\tilde{\gamma}_1 < \hat{\gamma}_1$ and $\tilde{\gamma}_j = \hat{\gamma}_j$ for all $j = 2, \ldots, m$, or (ii) $\tilde{\gamma}_m > \hat{\gamma}_m$ and $\tilde{\gamma}_j = \hat{\gamma}_j$ for all $j = 1, \ldots, m-1$, then*

$$\sum_{i=1}^{n} U_i(\tilde{x}_i) > \sum_{i=1}^{n} U_i(\hat{x}_i) \tag{2.23}$$

or, equivalently,

$$\sum_{j=1}^{m} V_j(\tilde{z}_j) > \sum_{j=1}^{m} V_j(\hat{z}_j). \tag{2.24}$$

Proof: See the Appendix.

Suppose that $\gamma_m > 0$ and that the weight γ_k, $1 < k < m$, is increased. This might be the result of an increase in concern for the agent's friends, with other weights unchanged. Lemma 2, which is used in the proof of Theorem 2, shows that, other weights equal, this increases the consumption of group k and decreases the consumption of all other groups. In addition, an increase in γ_k increases the value of the multiplier $\bar{\lambda}$. This implies that the most deserving group (at the margin)—group m—loses consumption to other groups. Total (unweighted) utility may rise, fall, or remain unchanged in this case.

The following is an example in which total utility falls in response to an increase in an intermediate weight. For three groups, total utility TU is $\sum_{j=1}^{3} V_j(\bar{z}_j)$ and, using (2.21), the change in TU with respect to an increase in γ_2 is given by

$$\frac{\partial TU}{\partial \gamma_2} = \sum_{j=1}^{3} V_j'(\bar{z}_j)\frac{\partial \bar{z}_j}{\partial \gamma_2} = \bar{\lambda}\sum_{j=1}^{3}\frac{1}{\gamma_j}\frac{\partial \bar{z}_j}{\partial \gamma_2}. \tag{2.25}$$

From (2.22),

$$\frac{\partial \bar{z}_2}{\partial \gamma_2} = -\frac{\partial \bar{z}_1}{\partial \gamma_2} - \frac{\partial \bar{z}_3}{\partial \gamma_2} \tag{2.26}$$

and it follows that

$$\frac{\partial TU}{\partial \gamma_2} = \left[\frac{1}{\gamma_1} - \frac{1}{\gamma_2}\right]\frac{\partial \bar{z}_1}{\partial \gamma_2} + \left[\frac{1}{\gamma_3} - \frac{1}{\gamma_2}\right]\frac{\partial \bar{z}_3}{\partial \gamma_2}. \tag{2.27}$$

If $\gamma_1 = 1$, $\gamma_2 = 1/2$, and $\gamma_3 = 1/3$, this becomes

$$\frac{\partial TU}{\partial \gamma_2} = -\frac{\partial \bar{z}_1}{\partial \gamma_2} + \frac{\partial \bar{z}_3}{\partial \gamma_2}. \tag{2.28}$$

Because both $\partial \bar{z}_1/\partial \gamma_2$ and $\partial \bar{z}_3/\partial \gamma_2$ are negative, the sign of $\partial TU/\partial \gamma_2$ depends on their relative magnitudes. Differentiating (2.21) with respect to γ_2 for $j = 1, 3$ and setting $\gamma_1 = 1$ and $\gamma_3 = 1/3$,

$$\frac{\partial \bar{z}_1}{\partial \gamma_2} = \frac{\partial \bar{\lambda}/\partial \gamma_2}{V_1''(\bar{z}_1)} \tag{2.29}$$

and

$$\frac{\partial \bar{z}_3}{\partial \gamma_2} = 3\frac{\partial \bar{\lambda}/\partial \gamma_2}{V_3''(\bar{z}_3)}. \tag{2.30}$$

Consequently, (2.28) becomes

$$\frac{\partial TU}{\partial \gamma_2} = \frac{\partial \bar{\lambda}}{\partial \gamma_2}\left[-\frac{1}{V_1''(\bar{z}_1)} + 3\frac{1}{V_3''(\bar{z}_3)}\right]. \tag{2.31}$$

Both $V_1''(\bar{z}_1)$ and $V_3''(\bar{z}_3)$ are negative, but they can take on any magnitude. That is, for any $\omega > 0$, any $(\bar{z}_1, \bar{z}_2, \bar{z}_3) \in \mathcal{R}_{++}^3$ with $\sum_{j=1}^{3} \bar{z}_j = \omega$, and any $(v_1, v_2, v_3) \in \mathcal{R}_{--}^3$, there exist functions V_1, V_2, V_3 such that $(\bar{z}_1, \bar{z}_2, \bar{z}_3)$ maximizes $V_1(z_1)+(1/2)V_2(z_2)+(1/3)V_3(z_3)$ subject to $\sum_{j=1}^{3} z_j \leq \omega$, and $\big(V_1''(\bar{z}_1), V_2''(\bar{z}_2), V_3''(\bar{z}_3)\big) = (v_1, v_2, v_3)$. If $V_1''(\bar{z}_1) = V_3''(\bar{z}_3) = -1$,

$$\frac{\partial TU}{\partial \gamma_2} = -2\frac{\partial \bar{\lambda}}{\partial \gamma_2}. \tag{2.32}$$

Lemma 2 shows that $\partial\bar{\lambda}/\partial\gamma_2$ is positive, and it follows that, in this example, $\partial TU/\partial\gamma_2 < 0$: total utility falls when the weight on group 2 increases.

Is some altruism better than none at all? That is, is giving some weight to at least one other group better than assigning a zero weight to all other groups? In the pure distribution problem with weights, this is the case, and the result is proved in Theorem 3.

Theorem 3: *In the pure distribution problem with weights, limited altruism results in a better outcome, according to the utilitarian value function, than self-interested behaviour.*

Proof: See the Appendix.

Theorem 3 implies that some altruism is always better than none in the pure distribution problem with weights and with discounting. This implies that discounting with any $\delta \in (0, 1]$ is better than purely self-interested behaviour.

3. The General Choice Problem

The pure distribution problem is a special case; in general, constraints may take many forms. For example, moral agents (either single individuals or groups) cannot control the behaviour of other people and technologies are rarely linear. Under certain circumstances, however, it is possible to predict the behaviour of others and the effect that one's actions have on it. Given that, an agent's actions will correspond to a set of feasible utility vectors.

Suppose that the set of feasible utility vectors corresponding to the set of actions available to the agents in group one is $\mathcal{F}$. It need not be convex, but it must be such that solutions to our maximization problems exist. For that reason we assume that $\mathcal{F}$ is compact (this is consistent with $\mathcal{F}$ being a finite set). We call the choice problem for group one the general choice problem.

The best possible behaviour for group one is to choose an action that maximizes the unweighted sum $\sum_{i=1}^{n} u_i$ subject to $(u_1, \ldots, u_n) \in \mathcal{F}$. If group one's altruism is limited, it maximizes the weighted sum

$$\sum_{j=1}^{m} \gamma_j \sum_{i \in N_j} u_i \tag{3.1}$$

subject to the constraint

$$(u_1, \ldots, u_n) \in \mathcal{F}, \tag{3.2}$$

where $\gamma_1 > 0$, $\gamma_{j-1} > \gamma_j$ for all $j = 2, \ldots, m$, and $\gamma_m \geq 0$. Self-interested behaviour on the part of group one is given by the case $\gamma_j = 0$ for all $j = 2, \ldots, m$, and behaviour that is guided by the utilitarian principle corresponds to the case in which $\gamma_j = \gamma_1$ for all $j = 2, \ldots, m$.

3.1. Two Groups

If there are only two groups of people, a very general result is true. In this case, group one acts to maximize

$$\gamma_1 \sum_{i \in N_1} u_i + \gamma_2 \sum_{i \in N_2} u_i \tag{3.3}$$

subject to (3.2). Because there are only two groups, this covers the discounting case with $\delta = \gamma_2/\gamma_1$.

We ask whether, when group one's altruism is limited, increasing γ_2 or decreasing γ_1 results in a better outcome according to the utilitarian value function. In this case, it does.

Theorem 4: *Given limited altruism, in the general choice problem with two groups, if (i) $\tilde{\gamma}_1 = \hat{\gamma}_1$ and $\tilde{\gamma}_2 > \hat{\gamma}_2$ or (ii) $\tilde{\gamma}_1 < \hat{\gamma}_1$ and $\tilde{\gamma}_2 = \hat{\gamma}_2$ then*

$$\sum_{i=1}^{n} \tilde{u}_i \geq \sum_{i=1}^{n} \hat{u}_i. \tag{3.4}$$

Proof: See the Appendix.

Because Theorem 4 is true for the case $\hat{\gamma}_2 = 0$, some altruism is at least as good as none in the two-group case. The inequality of Theorem 4 is weak rather than strict because maxima may not be unique and because the two utility vectors may coincide as well. If we assume, however, that maxima are unique and that the two utility vectors are different, the inequality in (3.4) is strict.

Theorem 4 shows that, in the case of two groups, an increase in the weight on group two or, equivalently, a decrease in the weight on group one results in a social improvement according to the utilitarian principle. We show in the following subsections that this result does not generalize to three or more groups.

3.2. *Geometric Discounting*

In this subsection, we construct two examples to show that the result of Theorem 1 is not true in the general choice problem. To do it, we consider the case in which there is one person in each group.

First we consider the maximization problem with positive weights $(\gamma_1, \ldots, \gamma_n) \in \mathcal{R}^n_{++}$. In that case, $(\bar{u}_1, \ldots, \bar{u}_n)$ maximizes $\sum_{i=1}^n \gamma_i u_i$ subject to $(u_1, \ldots, u_n) \in \mathcal{F}$. We define the function $\Pi \colon \mathcal{R}^n_{++} \longrightarrow \mathcal{R}$ by

$$\Pi(\gamma_1, \ldots, \gamma_n) = \max_{(u_1,\ldots,u_n)} \left\{ \sum_{i=1}^n \gamma_i u_i \;\middle|\; (u_1, \ldots, u_n) \in \mathcal{F} \right\}. \tag{3.5}$$

The function Π is analogous to a profit function, and it is homogeneous of degree one and convex. Standard duality theory shows that, if Π is differentiable, then

$$\frac{\partial \Pi(\gamma_1, \ldots, \gamma_n)}{\partial \gamma_i} = \Pi_i(\gamma_1, \ldots, \gamma_n) = \bar{u}_i \tag{3.6}$$

for all $i = 1, \ldots, n$.

For our example, $n = 3$ and we choose the approximation

$$\Pi(\gamma_1, \gamma_2, \gamma_3) = 2\gamma_1 + 9\gamma_2 + 3\gamma_3 - 6\gamma_1^{\frac{1}{2}}\gamma_2^{\frac{1}{2}} + 3\gamma_1^{\frac{1}{2}}\gamma_3^{\frac{1}{2}} - 8\gamma_2^{\frac{1}{2}}\gamma_3^{\frac{1}{2}}. \tag{3.7}$$

This function is a special case of $\sum_{i=1}^n \sum_{j=1}^n \alpha_{ij} \gamma_i^{\frac{1}{2}} \gamma_j^{\frac{1}{2}}$, and it is capable of approximating any function to the second order at a point (Diewert [1971]).

In the case of discounting, $\gamma_i = \delta^{i-1}$ for all $i = 1, \ldots, n$. We choose $\delta = 1/4$, and at that value, $(\gamma_1, \gamma_2, \gamma_3) = (1, 1/4, 1/16)$. Writing $(\mathring{u}_1, \mathring{u}_2, \mathring{u}_3)$ as the vector that maximizes the discounted sum of utilities, using (3.6), $\mathring{u}_1 = \Pi_1(1, 1/4, 1/16) = 7/8$, $\mathring{u}_2 = \Pi_2(1, 1/4, 1/16) = 1$, and $\mathring{u}_3 = \Pi_3(1, 1/4, 1/16) = 1$.

The Hessian matrix of Π at $(1, 1/4, 1/16)$ is

$$H = \begin{pmatrix} 9/16 & -3 & 3 \\ -3 & 16 & -16 \\ 3 & -16 & 16 \end{pmatrix}. \tag{3.8}$$

It can be checked that H satisfies the standard conditions for convexity and homogeneity of Π.

Total utility is

$$TU = \sum_{i=1}^3 \mathring{u}_i = \Pi_1 + \Pi_2 + \Pi_3, \tag{3.9}$$

and

$$\frac{\partial TU}{\partial \delta} = \Pi_{12} + 2\delta\Pi_{13} + \Pi_{22} + 2\delta\Pi_{23} + \Pi_{32} + 2\delta\Pi_{33}$$
$$= -3 + 3/2 + 16 - 8 - 16 + 8 = -3/2. \tag{3.10}$$

Thus, an increase in δ decreases total utility.

If $\delta = 1$, TU is maximized. What the example shows is that TU is not necessarily monotonically increasing in δ throughout its range.

A second example is concerned with discrete choice. Suppose that $n = 3$ and that the set $\mathcal{F}$ consists of just two utility vectors: $(100, 100, 100)$ and $(90, 150, 50)$. Action a_1 results in the first of these and action a_2 leads to the second. Total utility is greater in $(100, 100, 100)$, so utilitarian consequentialism recommends a_1. If person one discounts the utilities of the other two with $\delta = 1/4$, $(100, 100, 100)$ has a discounted value of $100 + 25 + 6.25 = 131.25$ and $(90, 150, 50)$ has a discounted value of $90 + 37.5 + 3.125 = 130.625$ and a_1 will be chosen. If, however, the level of discounting is reduced, with $\delta = 1/2$, discounted utilities are $100 + 50 + 25 = 175$ and $90 + 75 + 12.5 = 177.5$, and a_2 will be chosen. A decrease in the amount of discounting leads to an action that is worse from the utilitarian point of view.

3.3. Weighting Schemes

If there are more than two groups, increasing the weight on the utilities of the people in group m or decreasing the weight on the utilities of the people in group one has ambiguous consequences. In addition, it is not necessarily true that some altruism is better than none.

As in Subsection 3.2, we construct examples for the case where there is one person in each group. Analogously to (3.6), we know that

$$\frac{\partial \Pi(\gamma_1, \ldots, \gamma_n)}{\partial \gamma_i} = \bar{u}_i, \tag{3.11}$$

where $(\bar{u}_1, \ldots, \bar{u}_n)$ maximizes $\sum_{i=1}^{n} \gamma_i u_i$ subject to $(u_1, \ldots, u_n) \in \mathcal{F}$. It follows that the derivative of the change in total utility with respect to a change in γ_k is

$$\frac{\partial TU}{\partial \gamma_k} = \sum_{i=1}^{n} \Pi_{ik}(\gamma_1, \ldots, \gamma_n). \tag{3.12}$$

Using the example of Subsection 3.2,

$$\frac{\partial TU}{\partial \gamma_1} = \sum_{i=1}^{3} \Pi_{i1}(1, 1/4, 1/16) = 9/16 - 3 + 3 = 9/16 > 0. \tag{3.13}$$

Consequently, in the general choice problem, property (i) of Theorem 2 does not hold: a decrease in the weight on the utility of group one, other weights unchanged, can lead to a decrease in total utility. Similar examples can be found in which an increase in the smallest weight leads to a decrease in total utility.

A second example can be described as follows. Suppose that $n = 3$ and that the set $\mathcal{F}$ consists of two utility vectors: $(100, 100, 100)$ and $(80, 150, 50)$. Action a_1 results in the first vector and action a_2 leads to the second. Total utility is greater in $(100, 100, 100)$, so utilitarian consequentialism recommends a_1. If person one is completely selfish, he or she chooses actions using the weights $(\gamma_1, \gamma_2, \gamma_3) = (1, 0, 0)$. In this case, the weighted value of $(100, 100, 100)$ is 100 and the weighted value of $(80, 150, 50)$ is 80, so a_1—the better action—is chosen. If, however, person one uses the weights $(\gamma_1, \gamma_2, \gamma_3) = (1, .8, .2)$, weighted values of the two utility vectors are $100 + 80 + 20 = 200$ and $80 + 120 + 10 = 210$, so a_2 is chosen. An increase in the weights on the utilities of others leads to a worse action from the utilitarian point of view.

In the same example, suppose that weights are $(1, .3, .2)$. Then the weighted value of $(100, 100, 100)$ is $100 + 30 + 20 = 150$ and the weighted value of $(80, 150, 50)$ is $80 + 45 + 10 = 135$. Consequently, a_1 is chosen. Because a_2 is chosen when the weights are $(1, .8, .2)$ an increase in the middle weight (γ_2) results in an action which is worse according to utilitarianism.

Now suppose that the set $\mathcal{F}$ consists of the utility vectors $(100, 200, 0)$ and $(200, 10, 80)$. Action a_1 results in the first vector and a_2 results in the second. According to utilitarian consequentialism, $(100, 200, 0)$ is better and a_1 is the best action person one can take. Suppose that person one uses the weights $(1, .6, .1)$ in his or her objective function. Then the weighted value of the two outcomes is $100 + 120 = 220$ and $200 + 6 + 8 = 214$ so a_1 is chosen. If, however, the weight on person three is increased to .5, weighted values are 220 and $200 + 6 + 40 = 246$ and a_2 is chosen: an increase in the smallest weight leads to an action that is worse.

Given our assumptions, maximized weighted utility is continuous in $(\gamma_1, \ldots, \gamma_n)$. Consequently, if the weights are sufficiently close to equality, moving them toward equality results in actions that are no worse. Therefore, if departures from utilitarian consequentialism are small, there is no problem.

4. Solutions

If there is any reason for attaching lower weights to the interests of others than to oneself (or one's group), it must be that one wishes to limit the sacrifices that morally based action demands. If this practice is to be a reasonable one, it ought to be true that any change in the weights that brings them closer to the ones that represent the social good should move behaviour in that direction as well. But, as we have seen, this is not the case, both for reductions in discount rates and for changes in individual or group weights that move them closer to the weight on one's own well-being. These results are obtained under the assumption that the social good is represented by the utilitarian value function, but the next section shows that they have a much broader application.

We suggest, therefore, that a new approach is needed. It is legitimate to expect that agents may be unwilling to make very great sacrifices and to take actions that would harm people or groups with which they are personally or emotionally involved. Our suggestion is that principles for guiding action should take this into account explicitly rather than attempting to mimic it with weighting schemes.

Suppose that group one is to take an action from a feasible set whose consequences, in utility terms, consist of the vectors in the set $\mathcal{F}$. In order to avoid untoward sacrifices, group one chooses a set of utility levels that serve as constraints on its actions. Let C be a proper subset of the m groups which includes group one and, for each $j \in C$, let c_j be a total-utility floor for the group. We assume that there is a feasible utility vector $(u_1, \ldots, u_n) \in \mathcal{F}$ in which each group $j \in C$ has at least a total utility of c_j. Group one chooses actions to maximize total utility subject to these constraints. Therefore, an action is chosen such that its associated utility vector $(\check{u}_1, \ldots, \check{u}_n)$ maximizes

$$\sum_{i=1}^{n} u_i \tag{4.1}$$

subject to

$$(u_1, \ldots, u_n) \in \mathcal{F} \tag{4.2}$$

and

$$\sum_{i \in N_j} u_i \geq c_j \tag{4.3}$$

for all $j \in C$.

This formulation of the problem has a significant advantage over the discounting/weights one. If any of the constraints is relaxed, the action chosen is at least as good as the one chosen before the change. The best action(s) is (are) unchanged if the constraint did not bind in the original problem. If it did bind, the action that was chosen is still available because its associated utility vector satisfies the new constraints. Therefore, the new action can be no worse and may be better because more utility vectors and their associated actions are now at hand.

In our suggestion in this section, all the actions that are better than the chosen action, according to the utilitarian value function, must violate the constraints. If the constraints on group utility levels are regarded as describing obligatory actions for an individual agent, supererogatory actions are the ones that lead to utility vectors in $\mathcal{F}$ with greater total utilities and, at the same time, violate the constraints. Our theorems show that no similar formulation is possible (for the general choice problem) in the discounting/weights approach except in the two-group case.

In the general choice problem with two groups, the above formulation of the problem is equivalent to the discounting/weights formulation. If a utility vector is chosen when the weight on the utilities of people in group one is greater than the weight for group two and both are positive, then the *same* utility vector maximizes total utility subject to the constraint that the total utility of group one be no less than its total utility in the chosen utility vector in the weights case. This is shown in the following theorem.

Theorem 5: *In the general choice problem with two groups, for any* (γ_1, γ_2) *with* $\gamma_1 > \gamma_2 > 0$, $(\check{u}_1, \ldots, \check{u}_n)$ *maximizes*

$$\gamma_1 \sum_{i \in N_1} u_i + \gamma_2 \sum_{i \in N_2} u_i \tag{4.4}$$

subject to

$$(u_1, \ldots, u_n) \in \mathcal{F} \tag{4.5}$$

if and only if $(\check{u}_1, \ldots, \check{u}_n)$ *maximizes*

$$\sum_{i=1}^{n} u_i = \sum_{i \in N_1} u_i + \sum_{i \in N_2} u_i \tag{4.6}$$

subject to

$$(u_1, \ldots, u_n) \in \mathcal{F} \tag{4.7}$$

and

$$\sum_{i \in N_1} u_i \geq \sum_{i \in N_1} \check{u}_i. \tag{4.8}$$

Proof: See the Appendix.

The maximizing vectors in the two problems in Theorem 5 may not be unique. The theorem indicates, however, that each of them is a solution to both problems.

Our new formulation of the problem leaves open the question of how to set the constraints. It might be argued, for example, that, for a single agent, a utility floor equal to some fraction of the utility that he or she would get with selfish behaviour would be appropriate. These rules of thumb could be extended to other groups such as immediate family, members of the agent's community, and so on.

5. Generalizations

The results of this paper are not limited to the models that we have chosen. The presentation is significantly simplified with the utilitarian value function but other value functions preserve the results. These include the generalized utilitarian principles, which allow for inequality aversion in utilities (Subsection 5.1), and principles which discount the utilities of future generations or weight the interests of different groups unequally (Subsection 5.2). We do not endorse principles of the latter type, but include a discussion for completeness.

Some of the results of Section 2 on the pure distribution problem can also be generalized to additively separable constraints (Subsection 5.3). The section concludes with a discussion of uncertainty (Subsection 5.4). All our results on the general choice problem can be extended to uncertain environments. In addition, combinations of these generalizations, applied to the appropriate models, work without difficulty.

5.1. Generalized Utilitarianism

Utilitarianism is often criticized on the grounds that it exhibits no aversion to utility inequality.[10] The value function for generalized utilitarianism (GU)[11] employs transformed utility levels and is given by

$$\sum_{i=1}^{n} g(u_i) \tag{5.1}$$

where g is a continuous and increasing function. GU is weakly inequality averse if g is concave and strictly inequality averse if g is strictly concave. We assume in the discussion that follows that g is concave (which includes the case of strict concavity) and twice continuously differentiable.

The pure distribution problem can be adapted to GU by defining a transformed utility function $U_i^g = g \circ U_i$ for each $i = 1, \ldots, n$. Weights and discount factors are applied to transformed utilities $U_i^g(x_i)$, the analysis is unchanged, and the theorems indicate the direction of change in the GU value function.

The general choice problem using GU as the value function requires the selection of a utility vector in $\mathcal{F}$ that maximizes

$$\sum_{i=1}^{n} g(u_i). \tag{5.2}$$

[10] Utilitarianism does possess aversion to income inequality as long as individual (indirect) utility functions have the property of decreasing marginal utility of income.

[11] See Blackorby, Bossert, and Donaldson [1995, 1996, 1997a,b] and Broome [1992] for discussions.

This problem can be rewritten by defining transformed utilities $\upsilon_i = g(u_i)$. The set of feasible transformed utilities is

$$\mathcal{F}^g = \Big\{(\upsilon_1, \ldots, \upsilon_n) \mid \upsilon_i = g(u_i) \text{ for all } i = 1, \ldots, n \text{ and } (u_1, \ldots, u_n) \in \mathcal{F}\Big\}, \tag{5.3}$$

and the maximization problems use unweighted, discounted, or weighted transformed utilities and the feasible set $\mathcal{F}^g$.

5.2. Discounting in the Value Function

It might be thought that discounting or weighting is appropriate in the value function that represents the social good. We discuss the discounting case in this subsection and note that the case of weights is analogous. Either case can be extended to GU without difficulty.

Suppose that the social good is represented by the value function

$$\sum_{j=1}^{m} \bar{\delta}^{j-1} \sum_{i \in N_j} u_i \tag{5.4}$$

where $\bar{\delta}$ is the ethically appropriate discount factor. We define $\bar{\upsilon}_i$ to be $\bar{\delta}^{j-1} u_i$ for all $i \in N_j$ and all $j = 1, \ldots, m$. In the pure distribution problem, $\bar{\upsilon}_i = \bar{U}_i(x_i) := \bar{\delta}^{j-1} U_i(x_i)$ for all $i \in N_j$ and all $j = 1, \ldots, m$. In addition, we define $\check{\delta} = \delta / \bar{\delta}$ where δ is the actual discount factor in group one's objective function.

In the pure distribution problem, the best actions maximize

$$\sum_{j=1}^{m} \bar{\delta}^{j-1} \sum_{i \in N_j} U_i(x_i) = \sum_{j=1}^{m} \sum_{i \in N_j} \bar{U}_i(x_i) = \sum_{i=1}^{n} \bar{U}_i(x_i), \tag{5.5}$$

and, when utilities are discounted at a level that is more than the ethically appropriate level ($\delta < \bar{\delta}$), chosen actions will maximize the objective function

$$\sum_{j=1}^{m} \delta^{j-1} \sum_{i \in N_j} U_i(x_i) = \sum_{j=1}^{m} \left(\frac{\delta}{\bar{\delta}}\right)^{j-1} \sum_{i \in N_j} \bar{\delta}^{j-1} U_i(x_i) = \sum_{j=1}^{m} \check{\delta}^{j-1} \sum_{i \in N_j} \bar{U}_i(x_i). \tag{5.6}$$

Moving δ closer to $\bar{\delta}$ is equivalent to moving $\check{\delta}$ toward one, our analysis is unchanged and Theorem 1 can be interpreted as describing the consequences of moving δ toward $\bar{\delta}$.

Similar techniques permit the reinterpretation of Theorems 2 and 3. In those cases, actual weights are less than the ethically appropriate weights and are moved toward them.

In the general choice problem, the set of feasible discounted utilities is

$$\mathcal{F}^{\bar{\delta}} = \Big\{(\bar{\upsilon}_1, \ldots, \bar{\upsilon}_n) \mid \bar{\upsilon}_i = \bar{\delta}^{j-1} u_i \text{ for all } i \in N_j,\ j = 1, \ldots, m, \text{ and } (u_1, \ldots, u_n) \in \mathcal{F}\Big\}, \tag{5.7}$$

and the results of Sections 3 and 4 can be reinterpreted without difficulty.

5.3. *Additive Separability and the Feasible Set*

Some of the results of Section 2 (on the pure distribution problem) can be extended to the general choice problem if the feasible set $\mathcal{F}$ can be described by an additively separable function. In this case, there exist functions $h_1, \ldots, h_n$ and a constant $\bar{\omega}$ such that

$$\mathcal{F} = \left\{ (u_1, \ldots, u_n) \;\Big|\; \sum_{i=1}^{n} h_i(u_i) \le \bar{\omega} \right\}. \tag{5.8}$$

We assume that each function h_i is increasing, twice continuously differentiable, and strongly convex. This ensures that the set $\mathcal{F}$ is strictly convex.

For each i, we define $\nu_i = h_i(u_i)$ so that $u_i = h_i^{-1}(\nu_i)$. Because each h_i is strongly convex, h_i^{-1} is strongly concave. Defining $\Upsilon_i = h_i^{-1}$ for all $i = 1, \ldots, n$, the utilitarian solution to the general choice problem requires the choice of an action that leads to a vector $(\overset{*}{\nu}_1, \ldots, \overset{*}{\nu}_n)$ that maximizes

$$\sum_{i=1}^{n} \Upsilon_i(\nu_i) \tag{5.9}$$

subject to

$$\sum_{i=1}^{n} \nu_i \le \bar{\omega} \tag{5.10}$$

with similar maximization problems for the discounting and weights cases. Provided that solutions to the maximization problems exist, it is straightforward to reinterpret Theorem 1 and, with positive weights, Theorem 2 in this generalization.

5.4. *Uncertainty*

The results of Sections 3 and 4 can be generalized to the case of uncertainty. Let $\mathcal{S}$ be a finite set of contingent states of affairs with state-contingent feasible utility vectors. Thus $(u_1^s, \ldots, u_n^s)$ is the utility vector that occurs in state $s \in \mathcal{S}$ and the set $\mathcal{F}^{\mathcal{S}}$ consists of vectors of the form $((u_1^1, \ldots, u_n^1), \ldots, (u_1^S, \ldots, u_n^S))$ where $S = |\mathcal{S}|$. Individual i's *ex ante* utility is $u_i = U_i^e(u_i^1, \ldots, u_i^S)$ where U_i^e is an increasing function. We do not assume that it satisfies the expected-utility hypothesis, although that is covered as a special case. It is possible to describe the set of feasible *ex ante* utilities $\mathcal{F}^e$ given the functions U_i^e, $i = 1, \ldots, n$. It is

$$\begin{aligned}\mathcal{F}^e = \Big\{ (u_1, \ldots, u_n) \;\Big|\; & u_i = U_i^e(u_i^1, \ldots, u_i^S) \;\text{ for all }\; i = 1, \ldots, n \\ & \text{and } \big((u_1^1, \ldots, u_n^1), \ldots, (u_1^S, \ldots, u_n^S)\big) \in \mathcal{F}^{\mathcal{S}} \Big\}.\end{aligned} \tag{5.11}$$

This reduces the general choice problem under uncertainty to the mathematical equivalent of the general choice problem without uncertainty.

6. Conclusion

This paper suggests that objective functions that attach weights to or discount the utilities of others perform poorly as guides to action when altruism is limited. In the idealized environment of the pure distribution problem, the discounting formulation performs in a reasonable way—decreases in the discount rate result in better choices. In the weights formulation, however, only decreases in the weight for group one or increases in the weight for group m lead to the same result; all other changes have ambiguous consequences. In addition, a move from self-interested behaviour for group one to any amount of altruism leads to a better outcome.

In the general choice problem with three or more groups, however, none of these changes in objective functions has similar implications. Although utilitarian consequentialist behaviour, in which the value function and objective functions coincide, leads to the best outcome, increases in the amount of altruism may lead to better or worse outcomes. In addition, self-interested behaviour by group one may be better than incomplete altruism. In the case of two groups, any increase in the weight on the utilities of the members of the second group improves the choices of the first.

We have shown that these results are robust to a number of generalizations, which include employment of the generalized utilitarian value functions, value functions that allow discounting, and environments that allow the consequences of actions to be uncertain.[12] In addition, we show that some of the results of Section 2 on the pure distribution problem can be extended to the general choice problem if the feasible set of utility vectors $\mathcal{F}$ has an additively separable representation. We cannot reasonably expect feasible sets to take this form, however. Technologies are more general except in the simplest of models and moral agents cannot control the behaviour of others completely.

We suggest, therefore, that constraints on the utility levels of groups in which the agent is especially interested provide a better way to guide actions when altruism is less than perfect. In the case of intertemporal economic policies, the present generation might select total utility levels for itself and, possibly, several others that follow it. No policy would be chosen that has the consequence of pushing group utilities below these floors. In the case of individual moral agents, no action would be taken that fails to keep the well-being of the agent and the groups with which he or she is especially concerned at or above the constraint levels. In both cases, any relaxation of the constraints leads to outcomes that are no worse and, possibly, better. It is true, of course, that the best actions are the ones in which the constraints are absent.

[12] Uncertainty is sometimes used as a justification for discounting. In our formulation, however, it does not provide one because risk aversion is captured by individual utility functions.

APPENDIX

Lemma 1: *For each $j = 1, \ldots, m$, the function V_j, given by (2.4), is twice differentiable,*

$$V_j'(z_j) > 0 \text{ and } V_j''(z_j) < 0 \tag{A.1}$$

for all $z_j \in \mathcal{R}_{++}$, and

$$\lim_{z_j \to 0} V_j'(z_j) = \infty. \tag{A.2}$$

Proof: For all $z_j \in \mathcal{R}_{++}$, the first-order conditions for the maximization problem in (2.4) are

$$U_i'(\hat{x}_i) = \hat{\lambda} \tag{A.3}$$

and

$$\sum_{i \in N_j} \hat{x}_i = z_j. \tag{A.4}$$

From (A.3), $\hat{\lambda} > 0$.

The function V_j satisfies

$$V_j(z_j) = \sum_{i \in N_j} U_i(\hat{x}_i). \tag{A.5}$$

Consequently,

$$V_j'(z_j) = \sum_{i \in N_j} U_i'(\hat{x}_i) \frac{\partial \hat{x}_i}{\partial z_j} = \hat{\lambda} \sum_{i \in N_j} \frac{\partial \hat{x}_i}{\partial z_j}. \tag{A.6}$$

From (A.4),

$$\sum_{i \in N_j} \frac{\partial \hat{x}_i}{\partial z_j} = 1, \tag{A.7}$$

so

$$V_j'(z_j) = \hat{\lambda} > 0. \tag{A.8}$$

From (A.3),

$$U_i''(\hat{x}_i) \frac{\partial \hat{x}_i}{\partial z_j} = \frac{\partial \hat{\lambda}}{\partial z_j}. \tag{A.9}$$

Because $U_i''(\hat{x}_i) < 0$ for all $i \in N_j$, if $\partial \hat{\lambda} / \partial z_j \geq 0$, it follows that $\partial \hat{x}_i / \partial z_j \leq 0$ for all $i \in N_j$ which contradicts (A.7). Consequently, $\partial \hat{\lambda} / \partial z_j < 0$ and $\partial \hat{x}_i / \partial z_j > 0$.

(A.8) implies

$$V_j''(z_j) = \frac{\partial \hat{\lambda}}{\partial z_j} < 0. \tag{A.10}$$

As $z_j \to 0$, $\hat{x}_i \to 0$ and, using (A.3), $\lim_{z_j \to 0} \hat{\lambda} = \infty$. From (A.8), $\lim_{z_j \to 0} V_j'(z_j) = \infty$.
■

Theorem 1: *In the pure distribution problem with discounting, if $\tilde{\delta} > \hat{\delta}$, then*

$$\sum_{i=1}^{n} U_i(\tilde{x}_i) > \sum_{i=1}^{n} U_i(\hat{x}_i) \tag{A.11}$$

or, equivalently,

$$\sum_{j=1}^{m} V_j(\tilde{z}_j) > \sum_{j=1}^{m} V_j(\hat{z}_j). \tag{A.12}$$

Proof: It suffices to show that the assumptions imply (A.12). If $(\overset{\circ}{z}_1, \ldots, \overset{\circ}{z}_m)$ solves the maximization problem of (2.9) and (2.10), the first-order conditions are

$$\delta^{j-1} V_j'(\overset{\circ}{z}_j) - \overset{\circ}{\lambda} = 0 \tag{A.13}$$

and

$$-\sum_{j=1}^{m} \overset{\circ}{z}_j + \omega = 0 \tag{A.14}$$

which implies

$$\sum_{j=1}^{m} \frac{\partial \overset{\circ}{z}_j}{\partial \delta} = 0 \tag{A.15}$$

for all $j = 1, \ldots, m$. (A.13) implies $V_1'(\overset{\circ}{z}_1) = \overset{\circ}{\lambda}$ and $(j-1)\delta^{j-2} V_j'(\overset{\circ}{z}_j) = (j-1)\overset{\circ}{\lambda}/\delta$ for all $j \geq 2$. Furthermore,

$$V_j'(\overset{\circ}{z}_j) = \delta^{m-j} V_m'(\overset{\circ}{z}_m) \tag{A.16}$$

for all $j < m$.

Differentiating the first-order conditions with respect to δ, we obtain

$$\begin{bmatrix} 0 & -1 & \cdots & -1 \\ -1 & V_1'' & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ -1 & 0 & \cdots & V_m'' \end{bmatrix} \begin{bmatrix} \partial \overset{\circ}{\lambda}/\partial \delta \\ \partial \overset{\circ}{z}_1/\partial \delta \\ \vdots \\ \partial \overset{\circ}{z}_m/\partial \delta \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ -V_2'(\overset{\circ}{z}_2) \\ \vdots \\ -(m-1)\delta^{m-2} V_m'(\overset{\circ}{z}_m) \end{bmatrix} = \frac{\overset{\circ}{\lambda}}{\delta} \begin{bmatrix} 0 \\ 0 \\ -1 \\ \vdots \\ -(m-1) \end{bmatrix} \tag{A.17}$$

where V_j'' is used instead of $V_j''(\overset{\circ}{z}_j)$ for simplicity.

Let $y_0 := (\delta/\mathring{\lambda})(\partial\mathring{\lambda}/\partial\delta)$, $y_j := (\delta/\mathring{\lambda})(\partial\mathring{z}_j/\partial\delta)$ for all $j = 1, \ldots, m$, and

$$y = \begin{bmatrix} y_0 \\ y_1 \\ \vdots \\ y_m \end{bmatrix}. \tag{A.18}$$

Using these definitions and interchanging the first and last rows of (A.17), this system of equations is equivalent to

$$\begin{bmatrix} -1 & 0 & \cdots & 0 & V_m'' \\ -1 & V_1'' & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ -1 & 0 & \cdots & V_{m-1}'' & 0 \\ 0 & -1 & \cdots & -1 & -1 \end{bmatrix} y = \begin{bmatrix} -(m-1) \\ 0 \\ -1 \\ \vdots \\ -(m-2) \\ 0 \end{bmatrix}. \tag{A.19}$$

Multiplying the first equation by -1 and adding it to all rows but the last,

$$\begin{bmatrix} 1 & 0 & \cdots & 0 & -V_m'' \\ 0 & V_1'' & \cdots & 0 & -V_m'' \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & V_{m-1}'' & -V_m'' \\ 0 & -1 & \cdots & -1 & -1 \end{bmatrix} y = \begin{bmatrix} m-1 \\ m-1 \\ m-2 \\ \vdots \\ 1 \\ 0 \end{bmatrix}. \tag{A.20}$$

Dividing all but the first and last rows by $V_1'', \ldots, V_{m-1}''$ respectively and adding all of them to the last row,

$$\begin{bmatrix} 1 & 0 & \cdots & 0 & -V_m'' \\ 0 & 1 & \cdots & 0 & -V_m''/V_1'' \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -V_m''/V_{m-1}'' \\ 0 & 0 & \cdots & 0 & -\left[1 + \sum_{h=1}^{m-1} V_m''/V_h''\right] \end{bmatrix} y = \begin{bmatrix} m-1 \\ (m-1)/V_1'' \\ \vdots \\ 1/V_{m-1}'' \\ \sum_{h=1}^{m-1} (m-h)/V_h'' \end{bmatrix}. \tag{A.21}$$

Finally, we divide the last row by $-\left[1 + \sum_{h=1}^{m-1} V_m''/V_h''\right]$ to obtain

$$\begin{bmatrix} 1 & 0 & \cdots & 0 & -V_m'' \\ 0 & 1 & \cdots & 0 & -V_m''/V_1'' \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -V_m''/V_{m-1}'' \\ 0 & 0 & \cdots & 0 & 1 \end{bmatrix} y = \begin{bmatrix} m-1 \\ (m-1)/V_1'' \\ \vdots \\ 1/V_{m-1}'' \\ -\left[\sum_{h=1}^{m-1} (m-h)/V_h''\right] \Big/ \left[1 + \sum_{h=1}^{m-1} V_m''/V_h''\right] \end{bmatrix}. \tag{A.22}$$

Therefore,

$$y_m = -\frac{\sum_{h=1}^{m-1}(m-h)/V_h''}{1+\sum_{h=1}^{m-1}V_m''/V_h''} > 0. \tag{A.23}$$

For $1 \le j < m$, we obtain

$$y_j = \frac{m-j}{V_j''} + \frac{V_m''}{V_j''}y_m = \frac{1}{V_j''}\big[(m-j) + V_m'' y_m\big]. \tag{A.24}$$

Now we differentiate (A.16) with respect to δ to obtain

$$V_j''\frac{\partial \mathring{z}_j}{\partial \delta} = (m-j)\delta^{m-j-1}V_m'(\mathring{z}_m) + \delta^{m-j}V_m''\frac{\partial \mathring{z}_m}{\partial \delta}. \tag{A.25}$$

Substituting $(\mathring{\lambda}/\delta)y_j$ for $\partial \mathring{z}_j/\partial\delta$, and using (A.13) for $V_m'(\mathring{z}_m)$,

$$V_j''\frac{\mathring{\lambda}}{\delta}y_j = (m-j)\delta^{m-j-1}\frac{\mathring{\lambda}}{\delta^{m-1}} + \delta^{m-j}V_m''\frac{\mathring{\lambda}}{\delta}y_m, \tag{A.26}$$

which is the same as

$$V_j''\frac{\mathring{\lambda}}{\delta}y_j = (m-j)\delta^{m-j-1}\frac{\mathring{\lambda}}{\delta}\frac{1}{\delta^{m-2}} + \delta^{m-j}V_m''\frac{\mathring{\lambda}}{\delta}y_m. \tag{A.27}$$

Dividing by $\mathring{\lambda}/\delta$ and simplifying,

$$V_j''y_j = (m-j)\delta^{1-j} + \delta^{m-j}V_m''y_m. \tag{A.28}$$

By (A.24), $V_j''y_j = (m-j) + V_m''y_m$. Therefore, by (A.28),

$$(m-j)\delta^{1-j} + \delta^{m-j}V_m''y_m = (m-j) + V_m''y_m. \tag{A.29}$$

Solving for $V_m''y_m$, we obtain

$$V_m''y_m = \frac{(m-j)(\delta^{1-j}-1)}{1-\delta^{m-j}} \tag{A.30}$$

for all $j = 1, \ldots, m-1$. Substituting (A.30) into (A.24),

$$\begin{aligned} y_j &= \frac{1}{V_j''}\left[(m-j) + \frac{(m-j)(\delta^{1-j}-1)}{1-\delta^{m-j}}\right] \\ &= \frac{1}{V_j''}\left[\frac{(m-j)(\delta^{1-j}-\delta^{m-j})}{1-\delta^{m-j}}\right] < 0 \end{aligned} \tag{A.31}$$

for all $j = 1, \ldots, m-1$. By (A.15), $y_m = -\sum_{j=1}^{m-1} y_j$. Therefore, using (A.13),

$$\sum_{j=1}^{m} V_j'(\mathring{z}_j)y_j = \mathring{\lambda}\sum_{j=1}^{m} \frac{1}{\delta^{j-1}}y_j = \mathring{\lambda}\left[\sum_{j=1}^{m-1} \frac{1}{\delta^{j-1}}y_j - \frac{1}{\delta^{m-1}}\sum_{j=1}^{m-1} y_j\right] \\ = \mathring{\lambda}\sum_{j=1}^{m-1}(\delta^{1-j} - \delta^{1-m})y_j > 0, \tag{A.32}$$

where the last line follows from (A.31). Substituting back $(\delta/\mathring{\lambda})(\partial\mathring{z}_j/\partial\delta)$ for y_j,

$$\sum_{j=1}^{m} V_j'(\mathring{z}_j)\frac{\delta}{\mathring{\lambda}}\frac{\partial\mathring{z}_j}{\partial\delta} = \mathring{\lambda}\sum_{j=1}^{m}(\delta^{1-j} - \delta^{1-m})\frac{\delta}{\mathring{\lambda}}\frac{\partial\mathring{z}_j}{\partial\delta}, \tag{A.33}$$

and, therefore, writing total utility $\sum_{j=1}^{m} V_j(\mathring{z}_j)$ as TU,

$$\frac{\partial TU}{\partial\delta} = \sum_{j=1}^{m} V_j'(\mathring{z}_j)\frac{\partial\mathring{z}_j}{\partial\delta} = \mathring{\lambda}\sum_{j=1}^{m-1}(\delta^{1-j} - \delta^{1-m})\frac{\partial\mathring{z}_j}{\partial\delta} > 0 \tag{A.34}$$

because $y_j < 0$ implies $(\mathring{\lambda}/\delta)y_j = \partial\mathring{z}_j/\partial\delta < 0$.
■

Lemma 2: *Given limited altruism, in the pure distribution problem with positive weights, for any $k = 1, \ldots, m$,*

$$\frac{\partial\bar{z}_k}{\partial\gamma_k} > 0, \tag{A.35}$$

$$\frac{\partial\bar{z}_j}{\partial\gamma_k} < 0 \tag{A.36}$$

for all $j = 1, \ldots, m$ such that $j \neq k$, and

$$\frac{\partial\bar{\lambda}}{\partial\gamma_k} > 0. \tag{A.37}$$

Proof: Differentiating the first-order conditions (2.21) and (2.22),

$$V_k'(\bar{z}_k) + \gamma_k V_k''(\bar{z}_k)\frac{\partial \bar{z}_k}{\partial \gamma_k} = \frac{\partial \bar{\lambda}}{\partial \gamma_k}, \tag{A.38}$$

$$\gamma_j V_j''(\bar{z}_j)\frac{\partial \bar{z}_j}{\partial \gamma_k} = \frac{\partial \bar{\lambda}}{\partial \gamma_k} \tag{A.39}$$

for all $j = 1, \ldots, m$ such that $j \neq k$, and

$$\sum_{j=1}^{m} \frac{\partial \bar{z}_j}{\partial \gamma_k} = 0. \tag{A.40}$$

If $\partial\bar{\lambda}/\partial\gamma_k \leq 0$, it follows that $\partial\bar{z}_j/\partial\gamma_k \geq 0$ for all $j \neq k$ (from (A.39)), and

$$\gamma_k V_k''(\bar{z}_k)\frac{\partial \bar{z}_k}{\partial \gamma_k} = \frac{\partial \bar{\lambda}}{\partial \gamma_k} - V_k'(\bar{z}_k) < 0 \tag{A.41}$$

(from (A.38)), implying $\partial\bar{z}_k/\partial\gamma_k > 0$, which contradicts (A.40). Consequently, $\partial\bar{\lambda}/\partial\gamma_k > 0$. From (A.39), $\partial\bar{z}_j/\partial\gamma_k < 0$ for all $j \neq k$ and, from (A.40), $\partial\bar{z}_k/\partial\gamma_k > 0$.
■

Theorem 2: *Given limited altruism, in the pure distribution problem with weights, if (i) $\tilde{\gamma}_1 < \hat{\gamma}_1$ and $\tilde{\gamma}_j = \hat{\gamma}_j$ for all $j = 2, \ldots, m$, or (ii) $\tilde{\gamma}_m > \hat{\gamma}_m$ and $\tilde{\gamma}_j = \hat{\gamma}_j$ for all $j = 1, \ldots, m-1$, then*

$$\sum_{i=1}^{n} U_i(\tilde{x}_i) > \sum_{i=1}^{n} U_i(\hat{x}_i) \tag{A.42}$$

or, equivalently,

$$\sum_{j=1}^{m} V_j(\tilde{z}_j) > \sum_{j=1}^{m} V_j(\hat{z}_j). \tag{A.43}$$

Proof: Suppose that $\gamma_m > 0$. Total utility is $TU = \sum_{j=1}^{m} V_j(\bar{z}_j)$ and we show $\partial TU/\partial\gamma_1 < 0$ and $\partial TU/\partial\gamma_m > 0$. From Lemma 2, $\partial\bar{z}_1/\partial\gamma_1 > 0$, $\partial\bar{z}_j/\partial\gamma_1 < 0$ for all $j = 2, \ldots, m$, and $\partial\bar{\lambda}/\partial\gamma_1 > 0$. Using (2.21),

$$\frac{\partial TU}{\partial \gamma_1} = \sum_{j=1}^{m} V_j'(\bar{z}_j)\frac{\partial \bar{z}_j}{\partial \gamma_1} = \bar{\lambda}\sum_{j=1}^{m} \frac{1}{\gamma_j}\frac{\partial \bar{z}_j}{\partial \gamma_1}. \tag{A.44}$$

From (A.40),

$$\frac{\partial \bar{z}_1}{\partial \gamma_1} = -\sum_{j=2}^{m} \frac{\partial \bar{z}_j}{\partial \gamma_1} \tag{A.45}$$

and, substituting into (A.44),

$$\frac{\partial TU}{\partial \gamma_1} = \bar{\lambda} \sum_{j=2}^{m} \left[\frac{1}{\gamma_j} - \frac{1}{\gamma_1} \right] \frac{\partial \bar{z}_j}{\partial \gamma_1} < 0 \tag{A.46}$$

because $[1/\gamma_j - 1/\gamma_1] > 0$ for all $j = 2, \ldots, m$. Consequently, an increase in γ_1 decreases TU and a decrease in γ_1 increases TU. From Lemma 2, $\partial \bar{z}_m / \partial \gamma_m > 0$, $\partial \bar{z}_j / \partial \gamma_m < 0$ for all $j \neq m$, and $\partial \bar{\lambda} / \partial \gamma_m > 0$. A slight reworking of (A.44), (A.45), and (A.46) shows that $\partial TU / \partial \gamma_m > 0$. Consequently, if $\bar{\gamma}_m > 0$, (A.42) and (A.43) are true.

In case (i), if $\gamma_m = 0$, $m \geq 3$ and $\bar{z}_m = \hat{z}_m = 0$. Consequently, the above analysis applies to groups $1, \ldots, m-1$ and (A.42) and (A.43) are satisfied in this case.

In case (ii), because $\partial TU / \partial \gamma_m > 0$ for all $\gamma_m > 0$ and because, given our assumptions, TU is continuous, (A.42) and (A.43) are satisfied when $\hat{\gamma}_m = 0$.

■

Theorem 3: *In the pure distribution problem with weights, limited altruism results in a better outcome, according to the utilitarian value function, than self-interested behaviour.*

Proof: In the case where $\hat{\gamma}_1 > 0$ and $\hat{\gamma}_j = 0$ for all $j = 2, \ldots, m$, it is true that $\hat{z}_1 > 0$ and $\hat{z}_j = 0$ for all $j = 2, \ldots, m$. This solution is the same as the one obtained when $\gamma_1 = \tilde{\gamma}_1$ and $\gamma_j = 0$ for all $j = 2, \ldots, m$. Now consider the case in which $\gamma_1 = \tilde{\gamma}_1$, $\gamma_2 = \tilde{\gamma}_2$, and $\gamma_j = 0$ for all $j = 3, \ldots, m$. In this solution, $z_j = 0$ for all $j = 3, \ldots, m$, and the change is equivalent to one in which there are only two groups. Because the weight on group 2 has increased, total utility rises by Theorem 2 (ii). The weights on groups $3, \ldots, m$ can be increased to $\tilde{\gamma}_3, \ldots, \tilde{\gamma}_m$, one at a time. Because each z_j with a zero weight remains at zero, the result of Theorem 2 (ii) may be used as many times as necessary, and the theorem is established.

■

Theorem 4: *Given limited altruism, in the general choice problem with two groups, if (i) $\tilde{\gamma}_1 = \hat{\gamma}_1$ and $\tilde{\gamma}_2 > \hat{\gamma}_2$ or (ii) $\tilde{\gamma}_1 < \hat{\gamma}_1$ and $\tilde{\gamma}_2 = \hat{\gamma}_2$, then*

$$\sum_{i=1}^{n} \tilde{u}_i \geq \sum_{i=1}^{n} \hat{u}_i. \tag{A.47}$$

Proof: (i) Without loss of generality, let $\tilde{\gamma}_1 = \hat{\gamma}_1 = 1$, $\tilde{\gamma}_2 = \tilde{\gamma}$, and $\hat{\gamma}_2 = \hat{\gamma}$. Because $(\tilde{u}_1, \ldots, \tilde{u}_n) \in \mathcal{F}$ and $(\hat{u}_1, \ldots, \hat{u}_n) \in \mathcal{F}$,

$$\sum_{i \in N_1} \tilde{u}_i + \tilde{\gamma} \sum_{i \in N_2} \tilde{u}_i \geq \sum_{i \in N_1} \hat{u}_i + \tilde{\gamma} \sum_{i \in N_2} \hat{u}_i \tag{A.48}$$

and

$$\sum_{i \in N_1} \hat{u}_i + \hat{\gamma} \sum_{i \in N_2} \hat{u}_i \geq \sum_{i \in N_1} \tilde{u}_i + \hat{\gamma} \sum_{i \in N_2} \tilde{u}_i. \tag{A.49}$$

Adding and simplifying,

$$(\tilde{\gamma} - \hat{\gamma}) \sum_{i \in N_2} \tilde{u}_i \geq (\tilde{\gamma} - \hat{\gamma}) \sum_{i \in N_2} \hat{u}_i. \tag{A.50}$$

Because $\tilde{\gamma} > \hat{\gamma}$, $(\tilde{\gamma} - \hat{\gamma}) > 0$ and

$$\sum_{i \in N_2} \tilde{u}_i \geq \sum_{i \in N_2} \hat{u}_i. \tag{A.51}$$

Making use of this and (A.48),

$$\begin{aligned}
\sum_{i=1}^{n} \tilde{u}_i &= \sum_{i \in N_1} \tilde{u}_i + \sum_{i \in N_2} \tilde{u}_i \\
&= \sum_{i \in N_1} \tilde{u}_i + \tilde{\gamma} \sum_{i \in N_2} \tilde{u}_i + (1 - \tilde{\gamma}) \sum_{i \in N_2} \tilde{u}_i \\
&\geq \sum_{i \in N_1} \hat{u}_i + \tilde{\gamma} \sum_{i \in N_2} \hat{u}_i + (1 - \tilde{\gamma}) \sum_{i \in N_2} \hat{u}_i \\
&= \sum_{i \in N_1} \hat{u}_i + \sum_{i \in N_2} \hat{u}_i \\
&= \sum_{i=1}^{n} \hat{u}_i,
\end{aligned} \tag{A.52}$$

which proves case (i). The proof of case (ii) is analogous.
■

Theorem 5: *In the general choice problem with two groups, for any (γ_1, γ_2) with $\gamma_1 > \gamma_2 > 0$, $(\breve{u}_1, \ldots, \breve{u}_n)$ maximizes*

$$\gamma_1 \sum_{i \in N_1} u_i + \gamma_2 \sum_{i \in N_2} u_i \tag{A.53}$$

subject to

$$(u_1, \ldots, u_n) \in \mathcal{F} \tag{A.54}$$

if and only if $(\check{u}_1, \ldots, \check{u}_n)$ *maximizes*

$$\sum_{i=1}^{n} u_i = \sum_{i \in N_1} u_i + \sum_{i \in N_2} u_i \tag{A.55}$$

subject to

$$(u_1, \ldots, u_n) \in \mathcal{F} \tag{A.56}$$

and

$$\sum_{i \in N_1} u_i \geq \sum_{i \in N_1} \check{u}_i. \tag{A.57}$$

Proof: Without loss of generality, let $\gamma_1 = 1$ and $\gamma_2 = \gamma$. Define $\mathbf{U}_1 = \sum_{i \in N_1} u_i$ and $\mathbf{U}_2 = \sum_{i \in N_2} u_i$. Suppose that $(\check{u}_1, \ldots, \check{u}_n)$ maximizes $\sum_{i \in N_1} u_i + \gamma \sum_{i \in N_2} u_i$ subject to $(u_1, \ldots, u_n) \in \mathcal{F}$, and $(\overset{*}{\check{u}}_1, \ldots, \overset{*}{\check{u}}_n)$ maximizes

$$\sum_{i=1}^{n} u_i = \sum_{i \in N_1} u_i + \sum_{i \in N_2} u_i \tag{A.58}$$

subject to $\sum_{i \in N_1} u_i \geq \sum_{i \in N_1} \check{u}_i$ and $(u_1, \ldots, u_n) \in \mathcal{F}$. Then

$$\overset{*}{\check{\mathbf{U}}}_1 + \overset{*}{\check{\mathbf{U}}}_2 \geq \check{\mathbf{U}}_1 + \check{\mathbf{U}}_2 \tag{A.59}$$

and

$$\check{\mathbf{U}}_1 + \gamma \check{\mathbf{U}}_2 \geq \overset{*}{\check{\mathbf{U}}}_1 + \gamma \overset{*}{\check{\mathbf{U}}}_2. \tag{A.60}$$

Adding (A.59) and (A.60) and simplifying,

$$\overset{*}{\check{\mathbf{U}}}_2 + \gamma \check{\mathbf{U}}_2 \geq \check{\mathbf{U}}_2 + \gamma \overset{*}{\check{\mathbf{U}}}_2 \tag{A.61}$$

which implies

$$\overset{*}{\check{\mathbf{U}}}_2 (1 - \gamma) \geq \check{\mathbf{U}}_2 (1 - \gamma). \tag{A.62}$$

Because $0 < \gamma < 1$, $(1 - \gamma) > 0$, and

$$\overset{*}{\check{\mathbf{U}}}_2 \geq \check{\mathbf{U}}_2. \tag{A.63}$$

(A.63) implies

$$\gamma \overset{*}{\check{\mathbf{U}}}_2 \geq \gamma \check{\mathbf{U}}_2 \tag{A.64}$$

and, adding (A.60) and (A.64) and simplifying,

$$\check{\mathbf{U}}_1 \geq \overset{*}{\check{\mathbf{U}}}_1. \tag{A.65}$$

Because $\sum_{i\in N_1} \overset{*}{\check{u}}_i \geq \sum_{i\in N_1} \check{u}_i$, $\overset{*}{\check{\mathbf{U}}}_1 \geq \check{\mathbf{U}}_1$ and, hence, $\overset{*}{\check{\mathbf{U}}}_1 = \check{\mathbf{U}}_1$. Therefore, (A.59) implies

$$\overset{*}{\check{\mathbf{U}}}_2 \geq \check{\mathbf{U}}_2 \tag{A.66}$$

and (A.60) implies

$$\gamma\check{\mathbf{U}}_2 \geq \gamma\overset{*}{\check{\mathbf{U}}}_2 \tag{A.67}$$

or

$$\check{\mathbf{U}}_2 \geq \overset{*}{\check{\mathbf{U}}}_2. \tag{A.68}$$

(A.66) and (A.68) together imply

$$\overset{*}{\check{\mathbf{U}}}_2 = \check{\mathbf{U}}_2. \tag{A.69}$$

Because, in addition, $\overset{*}{\check{\mathbf{U}}}_1 = \check{\mathbf{U}}_1$,

$$\left(\sum_{i\in N_1} \overset{*}{\check{u}}_i, \sum_{i\in N_2} \overset{*}{\check{u}}_i\right) = \left(\sum_{i\in N_1} \check{u}_i, \sum_{i\in N_2} \check{u}_i\right) \tag{A.70}$$

and the theorem is proved.

■

REFERENCES

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 *Social Choice Re-Examined*, vol. 2, K. Arrow, A. Sen, and K. Suzumura, eds., 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. and D. Donaldson, 1984, Social criteria for evaluating population change, *Journal of Public Economics* **25**, 13–33.

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, 1996, Harsanyi's social aggregation theorem for state-contingent alternatives, Discussion Paper 96–26, University of British Columbia, Department of Economics.

Blackorby, C., D. Donaldson, and J. Weymark, 1997, Social aggregation and the expected utility hypothesis, Discussion Paper 97–05, University of British Columbia, Department of Economics.

Bossert, W., 1991, On intra- and interpersonal utility comparisons, *Social Choice and Welfare* **8**, 207–219.

Bossert, W. and J. Weymark, 1998, Utility in social choice, in *Handbook of Utility Theory,* S. Barberà, P. Hammond, and C. Seidl, eds., Kluwer, Dordrecht, forthcoming.

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

Broome, J., 1992, *Counting the Cost of Global Warming,* White Horse, Cambridge.

Cowen, T., 1992, Consequentialism implies a zero rate of intergenerational discount, in *Justice Between Age Groups and Generations,* P. Laslett and J. Fishkin, eds., Yale University Press, New Haven/London, 162–168.

Cowen, T. and D. Parfit, 1992, Against the social discount rate, in *Justice Between Age Groups and Generations,* P. Laslett and J. Fishkin, eds., Yale University Press, New Haven/London, 144–161.

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

Diewert, W.E., 1971, An application of the Shephard duality theorem: a generalized Leontieff production function, *Journal of Political Economy* **79**, 481–507.

Diewert, W.E., M. Avriel, and I. Zang, 1981, Nine kinds of quasiconcavity and concavity, *Journal of Economic Theory* **25**, 397–420.

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

Griffin, J., 1996, *Value Judgement: Improving Our Ethical Beliefs,* Clarendon, Oxford.

Heyd, D., 1982, *Supererogation: Its Status in Ethical Theory,* Cambridge University Press, Cambridge.

Mongin, P. and C. d'Aspremont, 1998, Utility theory in ethics, in *Handbook of Utility Theory,* S. Barberà, P. Hammond, and C. Seidl, eds., Kluwer, Dordrecht, forthcoming.

Mulgan, T., 1997, A non-proportional hybrid moral theory, *Utilitas* **9**, 291–306.

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.

Scheffler, S., 1982, *The Rejection of Consequentialism,* 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., 1977, On weights and measures: informational constraints in social welfare analysis, *Econometrica* **45**, 1539–1572.

Sugden, R. and A. Williams, 1978, *The Principles of Practical Cost-Benefit Analysis,* Oxford University Press, Oxford.

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