CHAPTER 1

Consequentialist social norms for public decisions

Peter J. Hammond

> A consequence cannot make evil an action that was good nor good an action that was evil.
>
> St. Thomas Aquinas[1]

> That the morality of actions depends on the consequences which they tend to produce, is the doctrine of rational persons of all schools; that the good or evil of these consequences is measured solely by pleasure or pain, is all of the doctrine of the school of utility, which is peculiar to it.
>
> J. S. Mill (from Warnock, 1962, p. 120)

> It must always be the duty of every agent to do that one, among all the actions which he *can* do on any given occasion, whose *total consequences* will have the greatest intrinsic value.
>
> G. E. Moore (1912, p. 121)

1 Motivation and outline

1.1 *Introduction*

Ever since the publication of the first edition of Arrow's *Social choice and individual values,* controversy has surrounded several of the conditions that he showed would lead inexorably to a dictatorship. Some of the controversy was discussed in the second edition (Arrow 1963). Much has happened since then, largely reflected in Sen (1970, 1982, 1984, 1985) as well as Arrow (1983).

The research reported in this essay was supported by the National Science Foundation under Grants SES-8201372 and SES-8302460, by the Institute of Advanced Studies of the Hebrew University in Jerusalem, and by the Österreichische Länderbank. I am grateful to these institutions for their support, and to Menahem Yaari, an anonymous referee, and the members of seminars at the University of Graz, Stanford University, the California Institute of Technology, and the Southern Methodist University for their comments, especially Kenneth Arrow and Norman Schofield.

Reprinted from
Social choice and public decision making
Essays in honor of Kenneth J. Arrow, Volume I
Edited by Walter P. Heller, Ross M. Starr and David A. Starrett

Two conditions in particular were often relaxed in an attempt to escape from dictatorship. One was independence of irrelevant alternatives, and a second was ordinality of the social choice function. Without independence of irrelevant alternatives, a wide class of ranking rules such as the Borda rule will satisfy all the other Arrow conditions, as discussed by Fishburn (1973), Fine and Fine (1974), and Gärdenfors (1973), for example. If ordinality is not insisted upon, the door is left open for rules such as Sen's (1970) "Pareto-extension rule," based on the Pareto quasiordering, and the transitive closure of majority rule.

One purpose of this essay is to show that both of these escapes from the Arrow theorem present fundamental difficulties of their own. Specifically, they lead to social choices that violate what is perhaps the most fundamental axiom of normative choice, that all decisions should be based solely on the consequences of decisions that are feasible. This fundamental axiom I call "consequentialism" following Anscombe (1958) and numerous succeeding moral philosophers;[2] it is an obvious generalization of the condition Arrow (1971) called "valuation of actions by consequences." Examples are presented in Sections 1.1 and 1.2 to illustrate how consequentialism is violated by the Pareto-extension rule and by the Borda rule.

A much more fundamental objection to utilitarianism in particular and to most of social choice theory in general is expressed in the introduction to Sen and Williams (1982). Indeed, their objections apply to any consequentialist theory of ethics, as Williams (1973) earlier made clear. Related objections are considered by Sen (1984) and, in particular, by Parfit (1984), Scheffler (1984), and Slote (1984). They seem to carry most force, however, in the field of personal ethics, where one is asking what constitutes moral behavior for an individual. Arrow and his many successors in social choice theory were more concerned with the making and evaluation of public policy, especially but not exclusively economic policy. Although one looks (all too often in vain) for high standards of personal morality among public policymakers, personal ethics is obviously a very inadequate guide to the public policymaker. Nor is it obvious that the kind of objections Sen and Williams and others have made to consequentialist ethics carry over with much force from personal morality to public choice. In all probability, some such objections will be advanced, though it is unclear that most of them cannot be met by allowing enough relevant ethical considerations to be counted within the space of "consequences." Nevertheless, in this chapter I shall investigate the implications of assuming that anybody responsible for a public decision of any kind

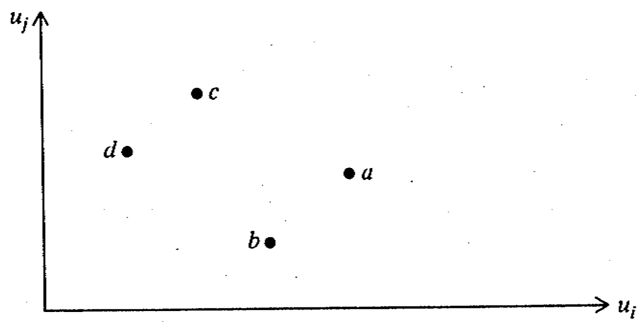

Figure 1. Utility possibilities.

makes it solely on the basis of its social consequences. I shall say a little more about the issues raised by this assumption in Sections 1.4 and 4.1.

After the examples of Sections 1.2 and 1.3, Section 1.4 argues why a much simplified model of public choice suffices to explore the implications of consequentialism. Section 1.5 then gives an outline of the heart of this essay.

1.2 *Inconsequentialism of the Pareto rule*

Allowing anything that is Pareto efficient into the social choice set is Sen's (1970) Pareto-extension social choice rule, which I shall call the *Pareto rule.* It is not consequentialist, as the following example shows.

Suppose there are two individuals i and j and four social states, or consequences, a, b, c, and d. Let the two individuals have strict preference rankings such that

$$a P_i b P_i c P_i d \quad \text{and} \quad c P_j d P_j a P_j b.$$

Corresponding utility possibilities are illustrated in Figure 1. Consider first a decision tree in which society chooses between nodes n_1 and n_2 at the initial node n_0. At n_1, a and d are still available, whereas at n_2, b and c are available, as shown in Figure 2(a). Then the Pareto choice set at node n_0 is $\{a, c\}$, since this pair is Pareto efficient whereas b and d are inefficient. Thus the Pareto rule allows the first decision at node n_0 to take the society either to n_1, planning to go on to a, or to n_2, planning to go on to c.

Consider, however, the second decision in the tree. Suppose society does make its first move from n_0 to n_1. Then, at n_1, the choice is between

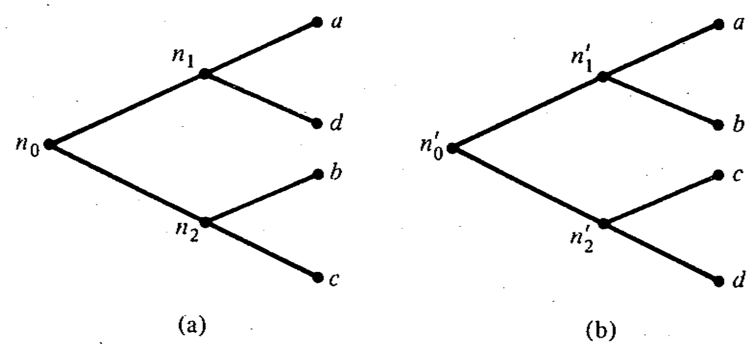

Figure 2.

a and *d*, both of which are Pareto efficient because *c*, which Pareto dominated *d* originally, is no longer available. To exclude *d*, society has to remember that *c* was available previously and Pareto dominated *d*. Or it has to precommit itself somehow to choosing only *a* at node n_1. In either case, one is led to wonder what past alternatives no longer available at node n_0 might have Pareto dominated *a* or *c*, or else what past commitments had been entered into. Moreover, the choice at n_1 of *a* over *d* depends not just on the pair of consequences *a* and *d*, which society faces then; it also depends on other consequences or on other considerations and so violates consequentialism according to the definition I shall give of this term.

Nor can the problem be avoided by moving to n_2 first, intending to continue on to *c*, which is Pareto efficient. For, at n_2, both *b* and *c* are Pareto efficient because *a*, which Pareto dominated *b* originally, is no longer available. Thus an exactly similar problem arises at n_2 as at n_1. At n_1, the Pareto rule chooses both *a* and *d*; at n_2, it chooses both *b* and *c*; therefore, the outcome of following it throughout could be *a*, *b*, *c*, or *d* – nothing whatsoever is ruled out.

Suppose now that the decision tree changes but the four consequences *a*, *b*, *c*, and *d* do not. Specifically, suppose that society chooses between the two nodes n'_1, n'_2 at the initial node n'_0 of the new tree. Suppose now that *a* and *b* are available at node n'_1, but at node n'_2, *c* and *d* are available, as shown in Figure 2(b). Then, since *a*, *c* are Pareto efficient and *b*, *d* are Pareto dominated, society's first decision could take it to n'_1 or to n'_2. Now, however, at n'_1, *b* is Pareto dominated by *a* so that only *a* will be chosen, and at n'_2, *d* is Pareto dominated by *c* so that only *c* will be chosen. For this new decision tree, the Pareto rule leads only to *a* or to *c*; *b* and *d* are excluded. Thus, by changing the structure of the decision tree, the social choice set has been changed from $\{a, b, c, d\}$ to just $\{a, c\}$. There is now no danger of choosing a consequence that was originally Pareto inefficient. In particular, the choice of consequences has been shown to depend on the structure of the decision tree rather than just on the set of available consequences. This shows that the Pareto rule does indeed violate consequentialism.

In this simple example, it is easy to construct other decision trees that lead the Pareto rule to the choice set $\{a, b, c\}$ or to the choice set $\{a, c, d\}$, as the reader will easily verify. Of course, a Pareto-efficient consequence can never be rejected by applying the Pareto rule within a decision tree.

To avoid such inconsequentialist choices, there must be a social preference ordering rather than just a quasiordering such as the Pareto rule applies. This was already shown implicitly in Hammond (1977) and explicitly in Hammond (1985b) and is discussed in Section 3.2.

1.3 *Inconsequentialism of the Borda rule*

Next I shall consider the Borda rule, which does indeed maximize a social ordering and has a number of other desirable properties, such as Pareto efficiency, anonymity, and so on. The Borda rule, however, violates Arrow's independence of irrelevant alternatives condition and this, we shall see, leads to inconsequentialism.

Suppose that there are two individuals, i and j as before, but that there are now five possible consequences a, b, c, d, and e. Suppose that the two individuals have strict preference rankings given by

$$a\,P_i\,b\,P_i\,c\,P_i\,d\,P_i\,e \quad \text{and} \quad d\,P_j\,e\,P_j\,a\,P_j\,b\,P_j\,c.$$

According to the Borda rule, each consequence is given a Borda count equal to the sum of the ranks given to it by both the individuals, with preferred consequences ranking higher. Thus i gives a a rank of 4 and j gives a a rank of 2, for a total Borda count $B(a)=4+2=6$. Similarly, $B(b)=3+1=4$, $B(c)=2+0=2$, $B(d)=1+4=5$, and $B(e)=0+3=3$. Since a has the highest Borda count, it is the consequence chosen by the Borda rule.

Suppose, however, there is a decision tree in which the choice of a leads to a second decision node n at which d and e are still available as alternatives to a, but b and c are no longer possible. Then, applying the Borda rule once again and ignoring the irrelevant consequences b and c, the preference rankings are

$$a\,P_j\,d\,P_i\,e \quad \text{and} \quad d\,P_j\,e\,P_j\,a$$

so the total Borda counts are

$$B(a) = 2+0 = 2, \qquad B(d) = 1+2 = 3, \qquad B(e) = 0+1 = 1.$$

This leads to the social choice of d rather than a. Clearly, then, the Borda rule gives rise to a dynamic inconsistency in the sense of Strotz (1956). If d is indeed the final outcome, it also shows that the Borda rule is inconsequentialist. If the decision tree only has one decision node so that the society is forced to choose at once between a, b, c, d, and e without any chance to revise the choice at a later node, the choice is a. But for the decision tree just considered, the choice is d even though the available consequences are exactly the same.

1.4 *Modeling normative public choice*

Since Arrow's *Social choice and individual values* appeared in 1951, normative social choice theory has concerned itself with devising good rules for deciding public policy questions. Arrow and most of his successors considered feasible sets of possible *social states* and examined various ways of choosing from such sets. Then a policy would be chosen, presumably, that led to good social states. This, however, is an intrinsically static model of social choice and also one that allows no uncertainty, imperfect information, and so forth, of the kind that besets most practical public policy questions. To allow fully for all such considerations, public choice needs to be modeled in an extensive game in which both individual members of the society and the public policymaker have moves to make. In fact, not even the usual von Neumann and Morgenstern (1953) apparatus of an extensive game or the generalization due to Kuhn (1953) are really quite sufficient; as Dubey and Kaneko (1982) and others have noticed, an extensive game in which two or more players really do move simultaneously may not be the same as one that uses the standard representation of *simultaneous* moves, with players moving in sequence but in ignorance of what other players have chosen. Allowing such simultaneous moves leads to consideration of *generalized* extensive games.

A complete model of a public decision problem, then, is likely to be an immensely complicated generalized extensive game whose players are the individual members of the society as well as at least one public agent whose behavior is the ultimate subject of our normative analysis. That is, we are particularly interested in recommendations or evaluations concerning the decisions of this public agent. A complete theory will specify how the public agent should behave in each and every possible society and extensive game, taking full account of all the strategic considerations that arise in such a game. In particular, such a theory requires a complete theory of solutions to extensive games; such a theory is lacking even in the standard but perhaps rather special case where all players are known to be expected utility maximizers, as von Neumann and Morgenstern originally assumed.

In this chapter, I shall circumvent the major difficulties such general extensive games present by restricting attention to trivial games under complete certainty in which all moves are made by the public agent whose behavior is the object of our normative analysis. Such games are, of course, one person games. Assuming that the public agent has perfect recall, the game can then be represented by a certain decision tree of the kind discussed in Hammond (1976a, 1977), for instance.

This is, of course, a very serious restriction indeed, which will have to be relaxed later on if any significant public decisions at all are to fall within the scope of the theory. Nevertheless, I shall show that considering only such a severely restricted class of public decision problems, consequentialism has very strong implications. These will be implications concerning the implicit choice of consequences in any such public decision tree. Later on, when the theory can be extended to broader classes of public decision problems under uncertainty in which also individuals other than the public agent have strategic choices to make, these results for such restricted decision problems will still be of use. In effect, they will tell us about the objectives the public agent should pursue subject to the constraints that arise because of individuals' strategic behavior. However, this last observation rather anticipates results concerning the likely nature of consequentialist behavior in extensive games – results in a theory yet to be developed. So it should be viewed with caution if not skepticism for the time being.

1.5 *Outline of chapter*

I hope the reader is convinced that a reexamination of the axiomatic foundations of social choice theory is still worthwhile. That, at any rate, is what I shall present in the remaining sections. Section 2 prepares the ground by presenting a formal description of general social norms for behavior in certain public decision trees. A new feature will be the definition of an individual's welfare norm in Section 2.3 as the norm for a society of identical individuals. This makes the concept of individual welfare explicitly normative and removes many of the usual objections to the Pareto criterion or to utilitarianism. An Arrow social norm is defined in Section 2.4

as one whose prescriptions in any decision tree depend only on the prescriptions of individuals' welfare norms in the same tree.

Section 3 then considers the implications of consequentialism for Arrow social norms. Assuming only consequentialism and an otherwise unrestricted domain of possible individual welfare norms, all the conditions of Arrow's general possibility theorem are shown to follow, which implies that there must be a dictator. That is, the social norm must always prescribe behavior that is acceptable to the individual welfare norm of a single dictator; the society is essentially treated as though all individuals had the same welfare norm as the dictator.

Section 4 discusses three important limitations and possible extensions of these results. First are the limitations of consequentialism itself. Second is the assumption of an unrestricted domain of individual norms and its inapplicability in economics. The third limitation discussed is the exclusion of interpersonal comparisons.

2 General social norms for certain decision trees

2.1 *Certain decision trees and behavior*

Certain decision trees describe decision problems in the absence of uncertainty. A *finite certain decision tree* $T=(N, N^{+1}, X)$ consists of a set of nodes N, a set of terminal nodes (or outcomes) X, and a successor correspondence $N^{+1}: N \twoheadrightarrow N$ defined so that $N^{+1}(n)$ is the set of nodes that immediately succeed n in the tree T. The set N includes an initial node n_0. The correspondence N^{+1} has to be suitably restricted so that T really is a tree (cf. Hammond 1976a, 1977, 1983, 1985b).

An *intended behavior* in the decision tree T is a *behavior strategy correspondence* $B: N \twoheadrightarrow N$ that determines at every node n a nonempty *behavior set* $B(n)$ that is a subset of $N^{+1}(n)$. Thus an intended behavior selects a set of acceptable moves at each decision node. Notice that any selection $\beta: N \to N$ from the correspondence B is a behavior strategy satisfying $\beta(n) \in B(n) \subset N^{+1}(n)$ for all $n \in N$.

However, intended behavior may easily depart from actual behavior, as the literature on naive choice certainly testifies (see, e.g., Strotz 1956; Hammond 1976a; Elster 1979). It is reasonable to assert that actual behavior matches intended behavior only at the initial node n_0 of the tree T. At any node n other than n_0, the agent no longer faces the decision tree T but only the *continuation decision tree* $T(n)$ whose initial node is n

itself and whose set of nodes $N(n)$ consists exactly of those nodes in N that eventually succeed n in T (including n itself).

Whereas intended behavior is of psychological interest, a normative theory of behavior is concerned with prescribing actual behavior rather than mere intentions. An agent's actual behavior in the choice tree T is described by the behavior set $B_{T(n)}(n)$ at the initial node of $T(n)$ for each decision node $n \in N$. So we naturally *define* $B_T(n) := B_{T(n)}(n)$ (all $n \in N$). In other words, an agent's actual behavior is described not by original intentions but by the intentions the agent has later on when he or she is about to make a move at the initial node n of the continuation tree $T(n)$. Then the agent's actual behavior must be *dynamically consistent* in the sense that whenever $n \in N$ for a decision tree T and $n' \in N(n)$ is a decision node of the continuation tree $T(n)$, then $B_{T(n)}(n') = B_T(n')$ because both must be equal to $B_{T(n')}(n')$.

Before coming to the definition of a behavior rule, two preliminary definitions are required. First, if T is a decision tree, say that T' is a *subtree* of T, and write $T' \subset T$ if (i) the set of nodes of N' of the tree T' is a subset of N, the set of nodes of T and (ii) T' is the decision tree with the same initial node n_0 whose successor correspondence $N_{T'}^{+1}$ is the obvious restriction of N^{+1} to the set of nodes N'. Second, a collection $\mathfrak{T}$ of finite decision trees is *complete* if

(i) whenever $T' \subset T$, then $T' \in \mathfrak{T}$, and
(ii) whenever $T \in \mathfrak{T}$ and $n \in N_T$, then $T(n) \in \mathfrak{T}$;

that is, a complete collection of decision trees includes all subtrees and continuation trees of its member trees.

A *behavior rule or norm* B is mapping whose domain $\mathfrak{T}$ is a complete set of decision trees, such that

(i) for each $T \in \mathfrak{T}$, B_T is a behavior in the decision tree T and
(ii) B is dynamically consistent (as defined above).

2.2 *Social norms for public decision trees*

A membership M is a finite set of individuals. For each individual member $i \in M$, assume there is a set Θ_i of possible characteristics θ_i of i. Let Θ^M denote the Cartesian product space $X_{i \in M} \Theta_i$ of possible *profiles* of individual characteristics, with typical member $\theta^M = (\theta_i)_{i \in M}$. A *society* S then consists of a pair (M, θ^M). I shall assume throughout this essay that the society is *exogenous* in the sense that it cannot be affected by behavior within the public decision tree. Where behavior does affect the society, the society itself becomes a matter for public decision. Such *endogenous* societies are considered in Hammond (1985a; in press). Here, I assume that M is fixed and that a society is sufficiently described by the profile θ^M.

A *social norm* is then a mapping $B(\theta^M)$ defined on a domain of societies S such that in each society θ^M of S, $B(\theta^M)$ is a behavior rule for a complete collection of decision trees $\mathfrak{T}$. Then $B_T(\theta^M)$ will denote behavior in decision tree T when the profile is θ^M.

2.3 *Individuals' welfare norms*

A general social norm $B(\theta^M)$ will have no regard for the behavior, intentions, choices, or preferences of the individual members of M. Such preferences may be part of the characteristic profile θ^M, but this has not yet been postulated. I shall now do so, but by an indirect route that enables the ethical value judgment of consumers' sovereignty so common in welfare economics to be dispensed with when it is appropriate to do so. It will also circumvent all the usual objections to the Pareto criterion. Since our concern is normative behavior, it is appropriate to have a normative concept of individual behavior too, and this is what I shall assume that we have.

First, I assume that there is a fixed set of possible individual characteristics Θ such that, for each individual i in M, $\Theta^i = \Theta$. This really amounts to assuming that the set of possible societies S is broad enough to allow any individual to have any possible characteristic.

Next, given any membership M and any characteristic θ of Θ, write $\theta 1^M$ for the society whose membership is M all of whom share the same characteristic θ. In other words, $\theta 1^M$ is a society of individuals with identical characteristics. Define the *welfare behavior rule* of an individual with characteristic θ in a membership M as $B(\theta 1^M)$. Notice that, by definition, individuals with the same characteristic have the same welfare behavior rule. This is not unreasonable; two individuals with different welfare behavior rules should be described by different characteristics, after all. Notice, however, that an individual's welfare behavior rule may depend on the membership M. In particular, the size of the membership may be important. For example, if all individuals are identical, equality of treatment is appealing unless there is so much scarcity that equality precludes survival. In any case, the very specification of behavior may depend on the membership M – one cannot ask fifty people to perform a certain task unless the society includes at least fifty people.

In the following, I shall write $B(\theta)$ for the welfare behavior rule of an individual with characteristic θ, treating the membership M as fixed throughout.

2.4 *Arrow social norms and unanimity*

Let M be any membership and T any decision tree in the domain $\mathfrak{T}$ of the social norm B. Then $B_T(\cdot)$ describes how social behavior within the tree T varies as society varies, with its membership fixed at M, however. Say that $B_T(\cdot)$ is *based purely on individuals' welfare behavior* within the subtrees of T if, whenever in the two societies $\theta^M, \bar{\theta}^M$ all the individuals' welfare behavior is identical for all subtrees T' of the tree T, so that

$$B_{T'}(\theta_i) = B_{T'}(\bar{\theta}_i) \qquad (\text{all } i \in M; \text{ all } T' \subset T),$$

then social behavior in the tree T in the two societies is the same too, so $B_T(\theta^M) = B_T(\bar{\theta}^M)$. Thus, within the tree T, each individual $i \in M$ is characterized just by welfare behavior $B_T(\theta_i)$ in all subtrees $T' \subset T$ for any profile θ^M.

As will be seen in Section 3.3, Arrow's original theory of social choice involves a social norm based purely on individuals' (welfare) behavior (or "values," though perhaps in a different sense from the one propounded here) within each tree in precisely this way. Thus I shall refer to such norms as *Arrow social norms.*

Let T be any decision tree and θ^M any profile. Say that individuals' welfare behavior is *unanimous within the subtrees of* T if, for every $T' \subset T$, $B_{T'}(\theta_i)$ is the same for all $i \in M$.

Lemma 1 (Unanimity). Let B be an Arrow social norm, T any decision tree, and θ^M any profile in which individuals' welfare behavior is unanimous within the subtrees T. Then $B_T(\theta^M) = B_T(\theta_i)$ (all $i \in M$); that is, social behavior within the tree T is identical to the individuals' unanimous behavior.

Proof: Because of unanimity, there is at least one characteristic $\bar{\theta} \in \Theta$ such that $B_{T'}(\theta_i) = B_{T'}(\bar{\theta})$ for every $i \in M$ and $T' \subset T$. Define $\bar{\theta}^M := \bar{\theta}1^M$. Then $B_{T'}(\theta_i) = B_{T'}(\bar{\theta}_i)$ for every $i \in M$ and $T' \subset T$. Because B is an Arrow social norm, it follows that $B_T(\theta^M) = B_T(\bar{\theta}^M)$. But $B_T(\bar{\theta}^M) = B_T(\bar{\theta}1^M) = B_T(\bar{\theta})$, where the last equality follows from the definition of an individual's behavior norm. Also, by definition of $\bar{\theta}$, $B_T(\bar{\theta}) = B_T(\theta_i)$ (all $i \in M$). So

$$B_T(\theta^M) = B_T(\bar{\theta}^M) = B_T(\bar{\theta}) = B_T(\theta_i) \qquad (\text{all } i \in M). \qquad \text{Q.E.D.}$$

3 Consequentialist Arrow social norms and dictatorship

3.1 *Consequential decision trees and the consequences of behavior*

Suppose that there is a set Y of possible *consequences*. Set Y should be sufficiently comprehensive to include everything that is relevant to any public decision. In ethics, many criticisms of utilitarianism in particular and of consequentialism in general take the form of introducing pertinent considerations that, it is alleged, do not affect utility or else are not consequences. Obviously, such criticisms lose much of their force if the space of consequences is expanded to accommodate *all* relevant considerations, though at the risk of having a theory so general that it loses virtually all empirical content [cf. the discussion in Broome (1984)]. Such grand issues are best left for discussion elsewhere; however, I shall say a little bit more about them in Section 4.1.

A decision tree $T=(N, N^{+1}, X)$ in a society with membership M is *consequential* if there is a *consequence mapping* $\gamma: X \to Y$ from terminal nodes to consequences. In future, it will be assumed that $\mathfrak{I}$, the domain of the Arrow social norm, consists of all such consequential decision trees. For each T of $\mathfrak{I}$, the existence of the consequence mapping γ will be implicitly assumed. Notice that $\mathfrak{I}$ is then a complete set of decision trees.

Given any consequential decision tree T in $\mathfrak{I}$, which is the domain of the Arrow social norm $B(s)$ for all societies s in S, the value of the norm in tree T is $B_T(s)$. This behavior norm leads to a set of nodes $N_T^B(s)$ in T that can be constructed recursively according to the rules

(i) $n_0 \in N_T^B(s)$ and
(ii) if $n \in N_T^B(s)$ and $n' \in B_T(s)(n) (\subset N^{+1}(n))$, then $n' \in N_T^B(s)$.

Ultimately, there is a set of terminal nodes,

$$X_T^B(s) := N_T^B(s) \cap X_T,$$

to which the norm gives rise. Corresponding to these terminal nodes are the consequences of the set $\gamma(X_T^B(s))$. These are the consequences that the norm B recommends, effectively, given the tree T and the society s.

3.2 *Consequentialist social norms and social welfare orderings*

As in Hammond (1985b), behavior is said to be consequentialist when it can be predicted solely on the basis of consequences, no matter what the decision tree may be. More specifically, knowledge of the set of consequences $\gamma(X_T)$ available in the decision tree should suffice to determine the consequences of behavior. Consequentialism is intended as a normative standard of behavior, and it retains its appeal for social norms as well as for individual behavior.

Two consequential decision trees T, T' are said to be *consequentially equivalent* if $\gamma(X_T) = \gamma'(X_{T'})$, so that the set of available consequences is the same in both trees, given the respective consequence mappings γ, γ'. The social norm B is said to be *consequentialist* if (i) whenever T, T' are two consequentially equivalent decision trees with consequence mappings γ, γ' and for any fixed society s, the sets of consequences $\gamma(X_T^B(s))$ and $\gamma'(X_{T'}^B(s))$ of the norm B are identical and (ii) the same is true whenever T or T' are continuations of consequential decision trees (as defined in Section 2.1). This is the obvious extension to social norms of the concept of consequentialist individual behavior. As with individual behavior, it has far-reaching implications, as will now be seen.

First, it is at once evident from the inclusion of all consequential decision trees in the domain $\mathfrak{I}$ that, given any nonempty finite subset Z of the set of consequences Y, there exists a decision tree T in $\mathfrak{I}$ and a consequence mapping $\gamma: X_T \to Y$ such that $\gamma(X_T) = Z$. Moreover, from the definition of consequentialism above, for each society s, there is a uniquely specified *consequence choice set* $C(s)(Z)$ whenever the decision tree T and consequence mapping γ together satisfy $\gamma(X_T) = Z$. So the consequentialist social norm B induces and indeed corresponds to a *consequence social choice function* $C(s)(\cdot)$ in each society s of S. The choice function $C(s)(\cdot)$ is defined on the set of all nonempty finite subsets of Y in the society s with membership M and takes values $C(s)(Z)$ that are nonempty subsets of Z for all such Z. In fact, $B(s)$ and $C(s)(\cdot)$ are related by the identity

$$\gamma(X_T^B(s)) = C(s)[\gamma(X_T)]$$

for all consequential decision trees T with consequence mappings $\gamma: X_T \to Y$.

The crucial implications of consequentialism discussed in Hammond (1977, 1983, 1985b) come about because of the application of consequentialism to each continuation decision tree $T(n)$ of a given tree T. It has been suggested that this represents an extra assumption, since, in principle, continuation trees (as defined in Section 2.1) could be treated quite differently from decision trees. If this were conceded, consequentialism would lose nearly all its force.

Continuation decision trees, however, *are* very obviously decision trees. In the absence of precommitment, decisions have to be taken sequentially at successive nodes of the tree, and the continuation decision tree describes the decision problem that is then faced. If precommitment is possible, then it really ought to be modeled *within* the decision tree – Odysseus' opportunity to precommit himself and his crew before confronting the Sirens is a wonderful example of this, as Strotz (1956) recognized (see also Elster 1979). Once all possibilities for precommitment are modeled as available decisions within the tree, then the sequence of decisions remains exactly as I have described it.

What may still be true, however, is that behavior within a continuation tree depends on consequences available from the *whole* tree and not just on consequences available within the continuation. Then continuation trees would have to be regarded as different from entire decision trees in formulating the consequentialist hypothesis, even though they may be consequentially equivalent according to the above definition. Put more simply, counterfactual consequences that past decisions made infeasible are still allowed to count.

I want to argue that such counterfactual consequences are not relevant to choices in continuation decision trees. Here, I feel that I can claim support from Arrow himself, who wrote as follows:

> The social welfare function approach, whether in Bergson's version or in mine, and "populistic democracy," as Dahl terms it, both imply that the social choice at any moment is determined by the range of alternative social states available (given the preferences of individuals); there is no special role given to one alternative because it happens to be identical to or derived from a historically given one. . . .
>
> It is against this background that the importance of the transitivity condition becomes clear. Those familiar with the integrability controversy in the field of consumer's demand theory will observe that the basic problem is the same: the independence of the final choice from the path to it. Transitivity will insure this independence; from any environment, there will be a chosen alternative, and, in the absence of a deadlock, no place for the historically given alternative to be chosen by default. . . .
>
> Collective rationality in the social choice mechanism is not then merely an illegitimate transfer from the individual to society, but an important attribute of a genuinely democratic system capable of full adaptation to varying environments. (1963, pp. 119–20 *passim*)

Although admitting the relevance of counterfactual consequences is rather different in principle from allowing a "place for the historically given alternative to be chosen by default," it has an identical effect. Certainly, it flies in the face of the claim "that the social choice at any moment is determined by the range of alternative social states available." Moreover, counterfactual consequences pose problems at the start of any contemporary decision tree. That is how lawyers earn their keep, after all. Good law, however, looks forward to future consequences rather than backward to what might have been. Of course, property rights are upheld and criminals punished because of what happened in the past, in part; but the ultimate justification, as many lawyers would recognize, are the future consequences of not maintaining property rights or of failing to punish criminals (cf. Harsanyi 1986). Deterrence is preferred to retribution. Future consequences, including the consequences of upholding the law, are relevant; past consequences are not. To quote Moore (1942, p. 559),[3] "Among the consequences of A nothing is included but what is the case *subsequently* to the occurrence of A."

Applying consequentialism to all consequential decision trees and to their continuations yields the following strong result. Its proof can be found in Hammond (1977, 1985b).

Lemma 2 (Ordinality). If the social norm $B_T(s)$ on the domain of profiles S and of consequential decision trees $\mathfrak{J}$ is consequentialist, then there exists a preference ordering $R(s)$ (complete, reflexive, and transitive) on Y for each $s \in S$ such that the consequence social choice function $C(s)$ satisfies

$$C(s)(Z) \equiv \{y \in Z \mid y' \in Z \text{ implies } yR(s)y'\}$$

for all nonempty and finite $Z \subset Y$.

Thus the consequentialist social norm must maximize a *social welfare ordering* $R(s)$ on Y in every society s of S.

3.3 *Consequentialist Arrow social norms and social welfare functions*

An Arrow social norm was defined in Section 2.4 so that in all societies s with membership M and for all decision trees T in the complete domain $\mathfrak{J}$, social behavior $B_T(s)$ in tree T depends only on individuals' welfare behavior $B_T(\theta_i)$ $(i \in M)$ within all subtrees $T' \subset T$ of the same tree. An individual's welfare behavior $B_T(\theta_i)$ was defined to be $B_T(\theta_i 1^M)$ where $\theta_i 1^M$ denotes a society with membership M all of whom have the identical characteristic θ_i.

If an Arrow social norm is also consequentialist, and if the domain $\mathfrak{J}$ consists of all consequential decision trees, then the conclusion of Section 3.2 implies that in any society s with membership M, the behavior

$B_T(s)$ corresponds to a preference ordering $R(s)$ on the space of consequences Y. Since $B_T(\theta) = B_T(\theta 1^M)$ for all possible T and θ, it follows too that every possible individual welfare behavior corresponds to a preference ordering $R(\theta)$ determined by the individual characteristic θ.

In Section 3.6 I shall show that a consequentialist Arrow social norm with a suitably large domain of possible societies s must be dictatorial. I shall do this by showing that such a norm satisfies all the conditions of Arrow's impossibility (or "general possibility") theorem. In particular, a consequentialist Arrow social norm can be represented by a "consequence" Arrow social welfare function (or "constitution") that satisfies the controversial independence of irrelevant alternatives condition as well as the Pareto criterion.

A *consequence Arrow social welfare function* (ASWF) is a mapping f whose domain is the space $\mathcal{R}^M(\Theta^M)$ of possible profiles of individual welfare orderings on Y, whose typical member is

$$R^M(\theta^M) \equiv (R(\theta_i))_{i \in M}.$$

Here Θ^M denotes the range of possible characteristic profiles θ^M in societies with membership M.

Because consequentialism implies that both the social norm and all individual welfare norms correspond to preference orderings, and because an Arrow social norm depends only on individual welfare norms, the following is immediate.

Lemma 3 (Existence of a consequence Arrow social welfare function). Corresponding to any consequentialist Arrow social norm $B_T(s)$ defined on a set of societies S and the set of all consequential decision trees $\mathfrak{T}$, there exists a unique consequence ASWF f on the domain $\mathcal{R}^M(\Theta^M)$ such that whenever $B_T(\theta_i)$ $(T \in \mathfrak{T})$ corresponds to $R(\theta_i)$ for all $i \in M$, then $B_T(\theta^M)$ $(T \in \mathfrak{T})$ corresponds to $f(R^M(\theta^M))$.

3.4 *Independence of irrelevant alternatives*

The example of Borda's rule was discussed in Section 1.3 and was used to suggest that consequentialism implies that a consequence ASWF must satisfy independence of irrelevant alternatives. This will now be claimed formally, though the proof is sufficiently close to proofs given elsewhere that I will omit it.

Say that the consequence ASWF f satisfies *condition I* (independence of irrelevant alternatives) if, for any set of consequences $Z \subset Y$ and for any pair of individual welfare profiles $R^M, \bar{R}^M$ of the domain $\mathcal{R}^M(\Theta^M)$

that meet the requirement that for all pairs of consequences $y^1, y^2 \in Z$ and for all members $i \in M$,

$$y^1 R_i y^2 \quad \text{iff } y^1 \bar{R}_i y^2,$$

it must be true that for all pairs of consequences $y^1, y^2 \in Z$,

$$y^1 f(R^M) y^2 \quad \text{iff } y^1 f(\bar{R}^M) y^2.$$

As already shown in Hammond (1977) in effect and more especially in Hammond (1983, pp. 184–6), consequentialism actually implies that this consequence ASWF must indeed satisfy condition I. Of course, the argument of Section 1.3 also suggests this. Thus:

Lemma 4 (Independence of irrelevant alternatives). The consequence Arrow social welfare function that corresponds to a consequentialist Arrow social norm in societies with a fixed membership must satisfy condition I.

3.5 *The Pareto and Pareto indifference conditions*

In addition, the consequence ASWF that corresponds to a consequentialist Arrow social norm can be shown to satisfy not just the ordinary Pareto condition but also the extra Pareto indifference condition, which has sometimes been invoked in social choice theory. This will be an implication of the unanimity Lemma 1 of Section 2.4. First, however, a few preliminaries, followed by a statement of the two Pareto conditions.

In accordance with standard practice, let $P(\theta_i)$ and $I(\theta_i)$ denote, respectively, the strict preference and the indifference relations determined by the weak preference relation $R(\theta_i)$. Similarly, write $R(\theta^M)$ for the preference ordering $f(R^M(\theta^M))$ that corresponds to the Arrow social norm for the profile θ^M, and let $P(\theta^M)$ and $I(\theta^M)$ denote the corresponding strict social preference and indifference relations.

Now, say that the consequence ASWF f satisfies *condition P* (Pareto) if, whenever y^1, y^2 is a pair of consequences in Y and θ^M is a profile of individual characteristics in Θ^M with the property $y^1 P(\theta_i) y^2$ (all $i \in M$), then $y^1 P(\theta^M) y^2$. And say that the consequence ASWF f satisfies *condition P^0* (Pareto indifference) if, whenever instead $y^1 I(\theta_i) y^2$ (all $i \in M$), then $y^1 I(\theta^M) y^2$.

Lemma 5 (Pareto and Pareto indifference). The consequence Arrow social welfare function that corresponds to a consequentialist Arrow social norm must satisfy conditions P and P^0.

The proof is rather obvious; for any pair of consequences y^1, y^2 of Y, consider a consequential decision tree T with just one decision at the initial node and just two terminal nodes, which give rise to the pair of consequences y^1, y^2. For this tree T, the unanimity Lemma 1 of Section 2.4 gives the result immediately.

3.6 *Unrestricted domain and dictatorship*

Lemmas 2–5 have shown that a consequentialist Arrow social norm corresponds, for each fixed membership M, to a consequence ASWF f that satisfies conditions I, P, and P^0. The only remaining conditions of Arrow's theorem that remain unsatisfied are the nondictatorship condition and condition U (unrestricted domain). I want to prove that any consequentialist Arrow social norm is dictatorial. So it remains only to ensure that condition U is indeed satisfied. Yet obviously it will not be satisfied unless a rich enough domain of possible individual characteristics and of individual welfare norms is allowed in each society. Indeed, as Arrow himself pointed out, should the domain of individuals' characteristics be suitably restricted – for example, if their welfare orderings are single peaked – then there will be a nondictatorial ASWF such as majority rule (provided the number of individuals happens to be odd so that the social weak preference relation is indeed transitive). Thus, apart from consequentialism, one other assumption is required, as follows.

Say that *individual welfare norms are unrestricted* in the domain of societies S if, for any preference ordering R on the space of consequences Y, there exists an individual characteristic θ and a society $\theta 1^M$ of identical individuals such that the consequential behavior rule $B_T(\theta 1^M)$ on the domain $\mathfrak{Z}$ of consequential decision trees T corresponds to the preference ordering R. Then the following is immediate.

Lemma 6 (Unrestricted domain). The consequence Arrow social welfare function that corresponds to a consequentialist Arrow social norm must have an unrestricted domain of individual preference profiles provided that individual welfare norms are unrestricted.

Finally, say that an Arrow social norm $B_T(s)$ on a domain of societies S and of decision trees $\mathfrak{Z}$ is *dictatorial* if there exists $d \in M$ such that, given any decision tree T of $\mathfrak{Z}$ and any node n of N_T,

$$B_T(\theta^M)(n) \subset B_T(\theta_d)(n).$$

In other words, a dictatorial social norm must always prescribe behavior that is acceptable according to the individual welfare norm of the dictator, behavior that would be acceptable if all individuals were like the dictator.

Evidently, then, Arrow's theorem, in combination with Lemmas 2–6, implies Theorem 7.

Theorem 7 (Arrow's theorem for consequentialist social norms). Any consequentialist Arrow social norm that leaves individual welfare norms unrestricted must be dictatorial. In fact, there must exist a consequence ASWF f and a dictator d with the property that, for all pairs of consequences y^1, y^2 in Y,

(a) $y^1 P(\theta_d) y^2$ implies $y^1 P(\theta^M) y^2$ and
(b) $y^1 I(\theta_i) y^2$ (all $i \in M$) implies $y^1 I(\theta^N) y^2$.

Note that conclusion (b) of Theorem 7 strengthens the usual Arrow theorem, but it is true here because of Lemma 4, which is really an implication of my definition of an Arrow social norm.

4 Conclusion: limitations and extensions

4.1 *Limitations of consequentialism*

The positive results of this essay have already been summarized in Section 1.5. In this and the next two sections, I shall discuss some of their limitations.

As I have already noted in Sections 1.1 and 3.1, many moral philosophers have produced cogent criticisms of consequentialism in general and of utilitarianism in particular. Among the criticisms, two deserve special attention. The first is the argument that is well represented (as far as I can judge) in the introduction to Sen and Williams (1982), who try to persuade us of the need to go "beyond utilitarianism" to more of a "pluralist" theory of ethics [see also Williams (1973) and Sen (1982, 1984)]. The second is the argument that consequentialism is likely to be self-defeating (see, e.g., Hodgson 1967; Parfit 1984; Scheffler 1984; Slote 1984; Harsanyi 1986).

If the consequences in any utilitarian theory of ethics are too coarsely defined, then obviously one has to go beyond these coarsely defined consequences and coarse concepts of utility in order to make consequentialism ethically acceptable. The notion of a consequence needs refining until all ethically relevant distinctions are accommodated. Sen and Williams view utilitarianism as a "monist" theory of ethics in which everything is reduced to consequences and their utilities. They argue instead for a more pluralist theory. My counterargument is that all considerations of plurality should already be taken into account when consequences are being defined and refined. An advantage of this approach is that we can replace abstruse arguments regarding the validity of consequentialism and of utilitarianism with much more practical arguments concerning what should count as a consequence in a theory of ethics.

A similar counterargument may, I believe, treat the objection that consequentialism is self-defeating. Parfit (1984) considers the self-defeating argument in its most subtle and sophisticated form, including one that recognizes the need for a refined notion of consequences (pp. 26–7). In one of its simpler forms, the objection is that if one tells lies or breaks promises in order to produce good consequences, the ultimate outcome may be bad because a liar or defaulter loses his credibility and his capacity to do good. Hodgson (1967) already recognizes that in fact such arguments are no threat to a broad enough conception of consequentialism; after all, certain acts are being condemned for their bad consequences. All such examples show is that the concept of a relevant consequence may be very subtle, including people's reputations and also the reputation of certain moral rules. Rather more interesting is the possibility that in a world full of consequentialist "pure do-gooders," everybody would be so concerned with the benefits of their actions to the world in general that very few specific benefits would ever get conferred. Perhaps that is why charity begins at home, and why we should follow Sidgwick (1907) and Adams (1976) in considering even motives as part of the relevant description of consequences. Indeed, Parfit (1984, pp. 40–3) even considers the possibility that consequentialism is *self-effacing* because it ultimately requires that we believe in some other, nonconsequentialist approach to ethics; even then, however, consequentialism is effective in determining what our nonconsequentialist ethical theory should be. In fact, Parfit's arguments show just how robust is the appeal of the consequentialist approach to ethics.

4.2 *Economic domains*

The second limitation is the assumption in Section 3.6 of an unrestricted domain of individual norms. This is unappealing when attention is concentrated on economic environments with private goods, as has been pointed out by numerous authors. The reason is that each individual usually cares only for his own private goods (in the absence of externalities) and has monotone preferences. In the extreme example with only one private good, this suffices to determine individuals' norms uniquely. Nor is the construction of an individual's welfare norm in Section 2.3 satisfactory in economic environments because, when individuals have different private concerns, they can never be all alike in the relevant sense of having identical welfare norms. Thus, following Arrow (1950, 1963) himself, in effect, my analysis is essentially restricted to domains in which all goods are public or, to use a terminology I find preferable, in which the consequence of a public decision is the "public environment." Even when no goods are private, there are natural restrictions such as monotonicity and continuity on preferences for public goods alone (see Kalai, Muller, and Satterthwaite 1979). I hope to be able to relax the unrestricted domain assumption in later work, just as Maskin (1976); Kalai and Ritz (1980); Kalai, Muller, and Satterthwaite (1979); Border (1983); and Ritz (1983, 1985) already have done to a considerable extent. Their work suggests that for fairly rich domains with at least two private goods, Arrow's theorem will remain valid under reasonably mild conditions.[4]

4.3 *Interpersonal comparisons*

A third serious limitation has been emphasized by Sen (1977, 1982, 1985) especially. Much of the power of Arrow's theorem derives from the paucity of information about individuals that is allowed to count. Indeed, as Arrow (1950, 1963, 1983) makes perfectly clear, his original theory of social choice was formulated precisely with the purpose of excluding interpersonal comparisons of utility, since they appear to lack any relationship to behavior. Actually, as Arrow (1977), Mirrlees (1982), and others have suggested, interpersonal comparisons of utility may be interpretable as representing preferences for alternative selves. And Yaari and Bar-Hillel (1984) showed that students with a rudimentary knowledge of economics were capable of making interpersonal comparisons in a fairly reasonable manner in several simple distribution problems.

So I believe that the lack of interpersonal comparisons needs to be tackled by considering an extended consequence space of triples (y, M, θ^M) where $y \in Y^M$ is an ordinary consequence of the kind considered here, M is the membership of a society, and θ^M is the profile of personal characteristics. A further extension, admitting cardinal utilities, would be to include lotteries over such triples. It can then be postulated that both individual and social norms apply to all decision trees with consequences in this extended space. The result is a form of utilitarianism due to Harsanyi (1955), at least when M is fixed. However, in the absence of domain restrictions, Arrow's theorem applies to this extended domain of consequences as well, and there must be a dictator. This is an implication Arrow certainly foresaw when we discussed interpersonal comparisons in 1975. The dictatorship is considerably weaker, however, than the usual one; all individuals' norms over the space of consequences Y can contribute to the social norm, in many cases, but the choice of a society (M, θ^M) and the implicit interpersonal comparisons of utility will usually have to be dictated by an "ethical" dictator (one hopes). This, at least, is my present conjecture, for which Roberts' (1980) analysis provides some support in a rather special case. [See also Pazner (1979) and Hammond (1985a).]

Even when the ethical dictator's interpersonal comparisons are admitted, however, another form of dictatorship easily arises. Consider a tree T in which the society (M, θ^M) is the same in all possible consequences so that the decision problem reduces to the choice of y in a fixed society, as has been considered throughout this chapter. Then it is in the spirit of this essay to have a consequentialist Arrow social norm depending only on individuals' behavior norms in the tree T. Thus we are back with a dictator as in Section 3.6; all the interpersonal comparisons that govern the choice of (M, θ^M) become irrelevant. In this sense, indeed, Arrow has always been right to insist that interpersonal comparisons do really violate independence of irrelevant alternatives.

So consequentialist Arrow social norms lead to a dictatorial social preference ordering, as in Arrow's general possibility theorem. Escape from dictatorship not only requires the richer information that interpersonal comparisons can provide but also the social norm within any decision tree has to depend on more than just individuals' behavior norms within the same tree, since allowing only such dependence – as in an Arrow social norm – implies independence of irrelevant alternatives in any fixed society. A weaker form of independence has to be admitted if interpersonal comparisons are to lead us away from dictatorship. Suggestions have been left for later work.[5]

NOTES

1 Attributed to St. Thomas Aquinas on p. 239 of Sayers and Reynolds' (1962) translated and annotated version of Dante's *Il Paradiso*. I am much indebted to Kenneth Arrow for this reference.

2 Actually, while Anscombe is apparently responsible for coining this particular ism, the origins of the doctrine are much older, going back to Mill in 1838 at

least and underlying the work of Moore (1903, 1912) as well as Broad (1914). See Bergström (1966). And note the quotation from St. Thomas Aquinas at the head of the chapter. For a more formal exposition of consequentialism and its implications for behavior in decision trees, see Hammond (1985b).

3 As cited in Bergström (1966, p. 63).

4 In fact, in Hammond (in press) the unrestricted domain assumption is replaced by an ethical liberalism assumption that is compatible with economic domains.

5 This contradicts what I have previously suggested in Hammond (1976b). For possible weakenings of Arrow's independence condition, see Hammond (1985a; in press).

REFERENCES

Adams, R. M. (1976), "Motive utilitarianism," *The Journal of Philosophy,* 73: 467–81.

Anscombe, G. E. M. (1958), "Modern moral philosophy," *Philosophy* 33: 1–19.

Arrow, K. J. (1950), "A difficulty in the concept of social welfare," *Journal of Political Economy,* 58: 328–46.

Arrow, K. J. (1963), *Social choice and individual values,* 2nd ed., New Haven: Yale University Press.

Arrow, K. J. (1971), *Essays in the theory of risk-bearing,* Chicago: Markham, and Amsterdam: North-Holland.

Arrow, K. J. (1977), "Extended sympathy and the possibility of social choice," *American Economic Review* (Papers and Proceedings), 67: 219–25.

Arrow, K. J. (1983), *Collected papers, Vol. I, Social choice and justice,* Cambridge, Mass.: Harvard University Press.

Bergström, L. (1966), *The alternatives and consequences of actions: An essay on certain fundamental notions in teleological ethics,* Stockholm: Almquist and Wiksell.

Border, K. C. (1983), "Social welfare functions for economic environments with and without the Pareto principle," *Journal of Economic Theory,* 29: 205–16.

Broad, C. D. (1914), "The doctrine of consequences in ethics," *The International Journal of Ethics,* 24: 293–320.

Broome, J. (1984), "Rationality and the sure-thing principle," University of Bristol, Economics Discussion Paper No. 84/158.

Dubey, P. and M. Kaneko (1982), "Information about moves in extensive games: I and II," Yale University, Cowles Foundation Discussion Papers Nos. 625 and 629.

Eister, J. (1979), *Ulysses and the sirens: Studies in rationality and irrationality,* Cambridge: Cambridge University Press.

Fine, B. and K. Fine (1974), "Social choice and individual ranking I and II," *Review of Economic Studies,* 41: 303–22 and 459–75.

Fishburn, P. C. (1973), *The theory of social choice,* Princeton: Princeton University Press.

Gärdenfors, P. (1973), "Positionalist voting functions," *Theory and Decision,* 4: 1–24.

Hammond, P. J. (1976a), "Changing tastes and coherent dynamic choice," *Review of Economic Studies,* 43: 159–73.
Hammond, P. J. (1976b), "Equity, Arrow's conditions, and Rawls' difference principle," *Econometrica,* 44: 793–804.
Hammond, P. J. (1977), "Dynamic restrictions on metastatic choice," *Economica,* 44: 337–50.
Hammond, P. J. (1983), "Ex-post optimality as a dynamically consistent objective for collective choice under uncertainty," in P. K. Pattanaik and M. Salles (Eds.), *Social choice and welfare,* Amsterdam: North-Holland.
Hammond, P. J. (1985a), "Independence of irrelevant personal consequences," for presentation at the 5th World Congress of the Econometric Society.
Hammond, P. J. (1985b), "Consequentialist behaviour in decision trees and expected utility," Stanford University, Institute for Mathematical Studies in the Social Sciences, Economics Working Paper No. 112.
Hammond, P. J. (in press), "On reconciling Arrow's theory of social choice with Harsanyi's fundamental utilitarianism," in G. Feiwel (ed.), *Arrow and the foundations of the theory of economic policy,* New York: Macmillan.
Harsanyi, J. C. (1955), "Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility," *Journal of Political Economy,* 63: 309–21.
Harsanyi, J. C. (1986), "Utilitarian morality in a world of very half-hearted altruists," this volume.
Hodgson, D. H. (1967), *Consequences of utilitarianism,* Oxford: Clarendon Press.
Kalai, E., E. Muller, and M. Satterthwaite (1979), "Social welfare functions when preferences are convex, strictly monotonic, and continuous," *Public Choice,* 34: 87–97.
Kalai, E. and Z. Ritz (1980), "Characterization of the private alternative domains admitting Arrow social welfare functions," *Journal of Economic Theory,* 22: 23–36.
Kuhn H. W. (1953), "Extensive games and the problem of information," in H. W. Kuhn and A. W. Tucker (Eds.), *Contributions to the theory of games II,* Princeton: Princeton University Press.
Maskin, E. S. (1976), "Social welfare functions on restricted domains," and "Social welfare functions for economics," mimeo.
Mirrlees, J. A. (1982), "The economic uses of utilitarianism," in A. K. Sen and B. A. O. Williams (Eds.), *Utilitarianism and beyond,* Cambridge: Cambridge University Press.
Moore, G. E. (1903), *Principia Ethica,* Cambridge: Cambridge University Press.
Moore, G. E. (1912), *Ethics,* London: Home University Library, Williams and Norgate.
Moore, G. E. (1942), "A reply to my critics," in P. A. Schlipp (Ed.), *The philosophy of G. E. Moore,* Evanston and Chicago: Northwestern University.
Parfit, D. (1984), *Reasons and persons,* Oxford: Oxford University Press.
Pazner, E. A. (1979), "Equity, nonfeasible alternatives and social choice: A reconsideration of the Concept of Social Welfare," in J.-J. Laffont (Ed.) *Aggregation and revelation of preferences,* Amsterdam: North-Holland.
Rawls, J. (1971), *A theory of justice,* Cambridge, Mass.: Harvard University Press.

Ritz, Z. (1983), "Restricted domains, Arrow social welfare functions and noncorruptible and nonmanipulable social choice corresondences: The case of private alternatives," *Mathematical Social Sciences,* 4: 155–79.

Ritz, Z. (1985), "Restricted domains, Arrow social welfare functions and noncorruptible and nonmanipulable social choice correspondences: The case of private and public alternatives," *Journal of Economic Theory,* 35: 1–18.

Roberts, K. W. S. (1980), "Possibility theorems with interpersonally comparable welfare levels," *Review of Economic Studies,* 47: 409–20.

Sayers, D. L. and B. Reynolds (translators and editors) (1962), *The divine comedy of Dante Alighieri the Florentine: Cantica III, Paradise,* Harmondsworth: Penguin Books.

Scheffler, S. (1984), *The rejection of consequentialism,* Oxford: Clarendon Press.

Sen, A. K. (1970), *Collective choice and social welfare,* San Francisco: Holden Day.

Sen, A. K. (1977), "On weights and measures: Informational constraints in social welfare analysis," *Econometrica,* 45: 1539–72.

Sen, A. K. (1982), *Choice, welfare and measurement,* Oxford: Basil Blackwell, and Cambridge, Mass.: M.I.T. Press.

Sen, A. K. (1984), "Consistency," Presidential Address to the Econometric Society.

Sen, A. K. (1985), "Information and invariance in normative choice," this volume.

Sen, A. K. and B. A. O. Williams (Eds.) (1982), *Utilitarianism and beyond,* Cambridge: Cambridge University Press.

Sidgwick, H. (1907), *The methods of ethics,* 7th. ed., London: Macmillan.

Slote, M. (1984), *Common sense morality and consequentialism,* London: Routledge and Kegan Paul.

Strotz, R. H. (1956), "Myopia and inconsistency in dynamic utility maximization," *Review of Economic Studies,* 23: 165–80.

von Neumann, J. and O. Morgenstern (1953), *Theory of games and economic behavior,* 3rd ed., Princeton: Princeton University Press.

Warnock, M. (Ed.) (1962), *Utilitarianism: Selections from the writings of Jeremy Bentham, John Stuart Mill, and John Austin,* London: Collins, the Fontana Library in Philosophy.

Williams, B. A. O. (1973), "A critique of utilitarianism," in J. J. C. Smart and B. A. O. Williams, *Utilitarianism for and against,* Cambridge: Cambridge University Press.

Yaari, M. E. and M. Bar-Hillel (1984), "On dividing justly," *Social Choice and Welfare,* 1: 1–24.