Progress in the Theory of Social Choice and Distributive Justice

PETER J. HAMMOND

Department of Economics, Stanford University, CA 94305-6072, U.S.A.

fax: +1 (415) 725-5702;   e-mail: `hammond@leland.stanford.edu`

Abstract: By definition, "consequentialist" behaviour in finite decision trees is explicable by its consequences. Both cost-benefit tests and "consequentialist" choices of economic policy necessarily require distributional judgements. These should emerge from a social welfare objective incorporating interpersonal comparisons. To accommodate them, Arrow's IIA condition should be weakened to independence of ethically irrelevant alternatives. When consequences are risky, dynamically consistent consequentialist behaviour on an unrestricted domain of finite decision trees entails maximizing expected social welfare. Combined with an "ethical liberalism" condition, this leads to "fundamental" utilitarianism, which requires a further weakening of IIA to independence of ethically irrelevant mixed alternatives.

<center>**Social Choice and Distributive Justice**</center>

1. Introduction

The topic of distributive justice is obviously important for its own sake in a world of poverty, hunger and other manifest forms of distributive injustice. But I believe it is particularly important now because it is in danger of falling out of fashion. The political leaders of America especially and of Britain and Russia too appear to value nuclear weapons more than food for the poor in their own countries, let alone for the poor in the third world. China's new economic mechanism and responsibility system seem explicitly designed to promote inequality, although there the anticipated benefits of increased efficiency may be more widely shared. In Italy I do not know if it has been claimed that the self-employed face better incentives to be efficient if they can expect to continue evading taxes, but I should not be surprised if some of the self-employed have suggested this.[1]

2. Distributive Justice

So distributive justice is important in policy analysis, even if only because its neglect has consequences which many people find obnoxious. But what is distributive justice? Absence of poverty? Equality?

Consider each of these in turn. Absence of poverty is clearly desirable, but unattainable. "The poor are always with us." This is partly because it seems beyond our collective capacity to eliminate widespread starvation in the world, let alone real poverty. But even if we could, there are bound to be some people poorer than others, who are then *relatively deprived*, in the sense explored in the extended study by Runciman (1966). That is, they are likely to fall below the poverty line in the society, because that line would never be drawn so low as to exclude everyone. According to this view, therefore, only with complete equality would there be absence of poverty.

What of equality? As the Chinese leaders are now openly admitting, this is unlikely to be desirable even if it is possible.

---

[1] I understand now that in fact some of the self-employed came out on strike [during 1984 and 1985] against the drive to enforce the tax laws in Italy. This is a new twist to "supply-side economics"! [Added in 1995] Obviously, this introductory paragraph has become somewhat dated during the last ten years, but the issues seem no less relevant today.

<center>1</center>

So that I can draw a diagram, let me pretend that there are only two classes of individuals, A-people and B-people, whose representative real incomes are y^A and y^B respectively. The A-people happen to be more skilful and productive than the B-people, so *laissez faire* leads to L in Figure 1. Under certain assumptions which are implicit in the case for *laissez faire*, this maximizes *total income*. At L, the A-people are much better off than the B-people; indeed, as in some of the examples of Coles and Hammond (1995), the B-people may even be unable to earn enough to survive, though that is extreme.
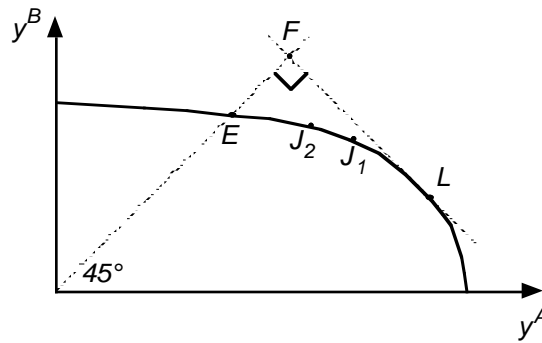

Figure 1

Defenders of *laissez faire*, if they worry about distribution at all, must also assume (though most do not realize it) that redistribution of income is costless. Then one can move along the line LF to a point F (first-best) with complete equality and efficiency. In practice, however, redistribution is only possible through "distortionary" taxes, and so moves from L can only take place along a curve as indicated by LJ_1J_2E. So, if equality is insisted on, the best that is possible is E, where total income has been greatly reduced.

To many of us (though not perhaps to the philosopher Rawls, author of *A Theory of Justice*), E is too extreme, just as L was in the other direction. A compromise at a point like J_1 or J_2 seems very desirable. But which J? And is there some *theory* which can be used to explain why a point like J is desirable while E and L are undesirable? Or must we rely solely on *moral intuitions* (which are apparently rather unreliable in this context, because so many different opinions are heard, not all of which can be dismissed as disguised self-interest).

2

3. Revelance to Project Evaluation

Such questions may seem too remote and philosophical to be of interest to practical decision makers. But they are actually rather important for them too, I believe. They evaluate projects, using techniques such as cost-benefit analysis. Take a project which produces a major scientific breakthrough, but under *laissez faire* lowers the wages of the (unskilled) B-people because their labour is in less demand. Total income increases. Is the project good? Answer — yes if there is no cross-over point (because $L'E'$ lies entirely above LE in Figure 2) or if the crossover point P lies to the left of J (as illustrated); no if P is to the right of J, generally speaking. If equality is valued enough, the extra costs of achieving the desired degree of equality, starting from L' as opposed to L, may exceed the extra benefits of L'. The necessary tax distortions may be greater.
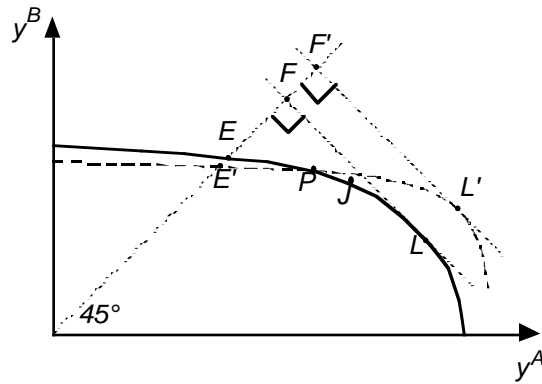

Figure 2

4. Social Choice Theory

The choice between different points in the set of feasible income distributions is one of the topics covered by *social choice theory*. This is the subject that I shall be discussing for most of the rest of this lecture. Social choice theory is about *objectives*, or *norms*, for public decision making. Hence the title of my paper "Consequentialist Social Norms for Public Decisions" (see Hammond, 1986).[2] The very term "social choice" comes from Kenneth Arrow's *Social Choice and Individual Values*, of which more shortly. The objective, or norm,

---

provides a rule for selecting one or more decisions from any feasible set. For example, it will select points from any redistribution frontier.

The objective may be representable as what economists call a Bergson social welfare function (after a famous article by Abram Bergson in 1938), which is to be maximized over whatever decisions are feasible. But it may not be, as the first objective I shall discuss is not.

5.   Pareto Satisficing and Project Evaluation

In economics, one allocation of goods and services to individuals is said to be *Pareto superior* to another when all individuals get something they prefer — or, at least, nobody gets anything worse and somebody gets something they prefer. An allocation is *Pareto efficient* if no other allocation is Pareto superior.

The *Pareto satisficing rule* (which Sen (1970) calls the "Pareto extension rule") simply selects *all* the Pareto efficient allocations among those that are feasible — the whole of the redistribution frontier in Figure 1. It is therefore completely silent on the issue of distributive justice, and completely indifferent to the sufferings of the poor except in those cases where the rich gain as well as the poor if those sufferings are alleviated. This makes it ethically repugnant, in my view.

Nor is it very helpful. For the project considered earlier, when the curves cross the Pareto criterion cannot tell us whether the project is good or bad. It cannot determine whether society is better off with the project or without. This remains true despite the many attempts economists have made to escape from this indecisiveness by devices such as compensation tests — originated, apparently, by Enrico Barone (1908), though more commonly ascribed to Lord Kaldor (1939) and Sir John Hicks — see the articles the latter published during the years 1939–1946 that are reprinted in Part II of Hicks (1981).[3] If we start at O in Figure 3 it is true that the project lets us go to P which is Pareto superior. But Pareto satisficing after the project is introduced lets us go to R. This is Pareto inferior to having no project at all but moving from O to Q.

---

[3]   For an assessment of Barone's earlier contribution, see Chipman and Moore (1978). For some later work on this topic, see Hammond and Sempere (1995), and the references included therein.
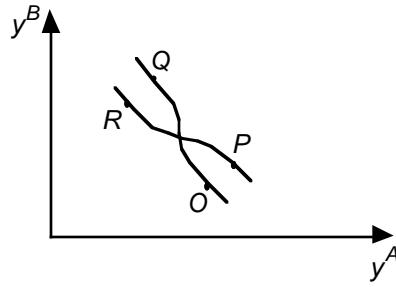
Figure 3

Economists have made several attempts to escape from such indecisiveness by devices like the Barone–Kaldor–Hicks compensation tests. A logically consistent escape is the wealth maximization approach discussed in Section 8 below. Yet even this depends on the choice of a reference price ratio, so really there can be no escape.[4]

6. Inconsequentialism of Pareto Satisficing
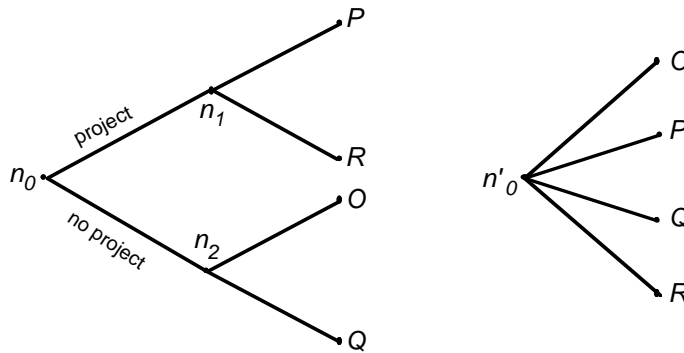
Figure 4

Another problem with Pareto satisficing is that it becomes important to consider *when* decisions are made and in what order — as well as what consequences they have for people. More precisely, the structure of the decision tree matters, as well as consequences, as illustrated by the two decision trees in Figure 4. These two trees are "consequentially equivalent" and yet Pareto satisficing behaviour produces quite different consequences.

---

[4] The last sentence has been expanded for this 1995 version.

5

Write $B(n)$ for the set of nodes immediately succeeding n that can be reached as a result of suitable behaviour or decisions. Let us consider what results from behaviour that deems any Pareto efficient outcome to be acceptable.

In the right hand tree, $B(n_0') = \{P, Q\}$ because only these two are Pareto efficient. In the left hand tree, $B(n_1) = \{P, R\}$ and $B(n_2) = \{O, Q\}$ because we are considering only Pareto efficient redistribution. This makes it hard to advise on behaviour at n_0, since the ultimate consequences are unpredictable. There are three possibilities:

 (i) $B(n_0) = \{n_1\}$ and the consequences are $\{P, R\}$;

 (ii) $B(n_0) = \{n_2\}$ and the consequences are $\{O, Q\}$;

 (iii) $B(n_0) = \{n_1, n_2\}$ and the consequences are $\{O, P, Q, R\}$.

In *no* case is the set of possible consequences $\{P, Q\}$, as in the set on the right.

7. Consequentialism and Ordinality

So the Pareto criterion makes policy recommendations depend crucially on the structure of the decision tree as well as upon consequences.
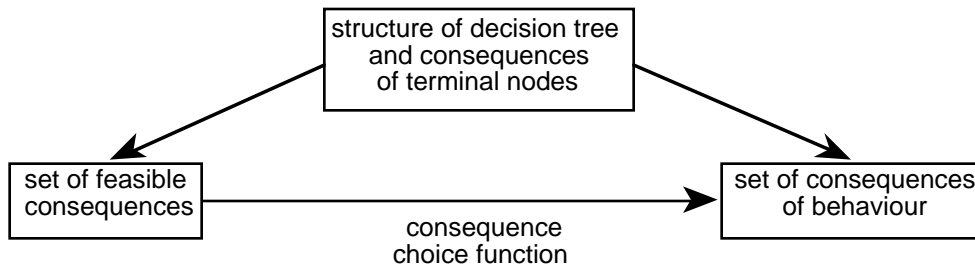
Figure 5

This flaunts a fundamental tenet of rational behaviour that actually lies behind most of the postulates which John Harsanyi was using in his discussion of game theory in the preceding lecture at this conference.[5] This tenet holds that good decisions are those which ultimately produce good consequences for people, regardless of *when* decisions have to be taken at different points of a decision tree. So in two trees which are consequentially

---

[5] That lecture was on "Rationality and Strategic Behavior." [Added in 1995] The chapter later published in the conference volume, Harsanyi (1986), was somewhat different.

equivalent, as the two shown in Figure 4 are, behaviour should also be consequentially equivalent — i.e., lead to identical consequences. With a little more formalism, there must be a *consequence choice function* making Figure 5 complete: so that behaviour in any decision tree (including any continuation subtree) must be equivalent to choosing the consequences of behaviour from the feasible set of consequences of possible behaviour. This is the tenet or "pre-axiom" that I call *consequentialism*, following the moral philosophers.

Hence, Pareto satisficing violates consequentialism. Under the two additional assumptions that behaviour is well defined for all finite decision trees, and that it is dynamically consistent in subtrees,[6] consequentialism is also violated by any behaviour that does not correspond to the *ordinal choice of consequences* in the following sense. First, there must exist a *binary preference relation R* on the given domain of consequences Y which is an *ordering* because:

(i) R is reflexive ($y \; R \; y$, for all $y \in Y$);

(ii) R is complete ($y \; R \; z$ or $z \; R \; y$, for all $y, z \in Y$);

(iii) R is transitive ($x \; R \; y$ and $y \; R \; z$ imply $x \; R \; z$, for all $x, y, z \in Y$).

Second, give any (finite) set $Z \subset Y$, the choice set of consequences must be

$$C(Z) = \{ \, y \in Z \mid y \; R \; z \; (\text{all } z \in Z) \, \}$$

— i.e., the set of consequences which *maximize R over Z*.

---

[6] These two extra assumptions were not made explicit in the 1985 version. As explained in Hammond (1988a, 1996), it is required first that $B(n)$ be defined and non-empty at every decision node of every finite decision tree. Second, dynamic consistency requires $B(n)$ to be the same, whether n is regarded as belonging to the whole decision tree, or to a (continuation) subtree that contains n — indeed, that is one reason for omitting the tree as an argument of $B(n)$.

8.   Wealth Maximization

Both Arrow's approach to social choice and Pareto satisficing have been popular with economists (though not necessarily with Pareto and Arrow) because they avoid *interpersonal comparisons* of the kind needed to trade off one person's gains against another's losses. Such interpersonal comparisons, as Robbins pointed out, involve "unscientific" ethical value judgements, and economists like to pretend that they are scientists, even when discussing what policies are appropriate (as opposed to analysing the effects of policies).

In Chicago, economists have another way of doing "science" which gives a social ordering. Harberger (1971) and Posner (1981) especially have advocated maximizing total wealth, measured in dollars, of course. The theoretical underpinnings of this wealth maximization criterion can be found in Chapman and Moore (1980), and Posner's book receives extensive discussion in Hammond (1982). The criterion is crude ethics indeed, valuing the extra dollar spent by a wealthy man on a better cigar as much as the extra dollar's worth of medicine needed to save the life of a sick child. It always selects the *laissez faire* point in Figures 1 and 2, and always will as long as wealth is measured relative to a fixed reference vector of prices which emerge in competitive markets. Moreover, it does not avoid making interpersonal comparisons; rather, it values people by their wealth. Plutocracy!

9.   Consequentialist Social Norms and Ethical Liberalism

I am going to present a model of a society that permits interpersonal comparisons; then I shall claim a generalized version of Arrow's famous impossibility theorem. I shall be concerned with *consequentialist norms* which, for every decision tree with consequences in a given space, specify behaviour as a function of the tree (and of its consequences).

The choice of who is to live in the society, and of how many people there should be, is actually quite a bit more complicated than has generally been realized. So I shall leave this question for later analysis,[7] and simply assume a *fixed* finite set of individuals M. Each individual $i \in M$ has a variable characteristic $\theta_i \in \Theta_i$ which determines his or her preferences and needs (including, of course, whether it is his or her!).

---

[7] [Added in 1995] Some discussion can be found in Hammond (1988b, 1992, 1996). In Hammond (1991b), I argue that interpersonal comparisons are revealed by the ethical choice of persons' characteristics.

Individual i's *personal consequences* are the pairs (x_i, θ_i), elements of the Cartesian product set $X_i \times \Theta_i$, where x_i is for economists a consumption vector (which includes access to public goods, the effects of the environment upon i, etc.).

Let $X^M := \prod_{i \in M} X_i$ and $\Theta^M := \prod_{i \in M} \Theta_i$. Then $Y := X^M \times \Theta^M$ is the relevant space of consequences. The social norm β is defined for all decision trees with consequences in Y. It is consequentialist, so there is a social ordering R on Y. In addition, it induces a norm β_i on decision trees with consequences in the set $X_i \times \Theta_i \times \{\langle \bar{x}_j, \bar{\theta}_j \rangle_{j \neq i}\}$ for every $i \in M$ and every fixed $(\bar{x}_j, \bar{\theta}_j)_{j \neq i}$. There is an associated conditional ordering $R_i(\langle \bar{x}_j, \bar{\theta}_j \rangle_{j \neq i})$ on $X_i \times \Theta_i$. A natural assumption is that X_i and Θ_i are defined broadly enough to include everything of consequence to i, so that β_i — and so the associated preference ordering R_i — are actually independent of $\langle \bar{x}_j, \bar{\theta}_j \rangle_{j \neq i}$. This is a "liberalism" assumption — the social norm coincides with the *individual norm* β_i for all *personal decision trees* with fixed consequences except those concerning i. I call it *ethical liberalism*.[8] Notice that ethical liberalism actually implies the following *strict Pareto* condition (P*):

(i) if $(x_i, \theta_i) \; R_i \; (\bar{x}_i, \bar{\theta}_i)$ (all $i \in M$) then $(x^M, \theta^M) \; R \; (\bar{x}^M, \bar{\theta}^M)$;

(ii) if (i) is true and if $(x_j, \theta_j) \; P_j \; (\bar{x}_j, \bar{\theta}_j)$ (some $j \in M$), then $(x^M, \theta^M) \; P \; (\bar{x}^M, \bar{\theta}^M)$.

It also implies that, if there is a Bergson social welfare function $W(x^M, \theta^M)$ which represents R, then $U_i(x_i, \theta_i) \equiv W(x_i, \theta_i, \langle \bar{x}_j, \bar{\theta}_j \rangle_{j \neq i})$ represents R_i (for *any* fixed $\langle \bar{x}_j, \bar{\theta}_j \rangle_{j \neq i}$) and there is a strictly increasing function $F : \Re^M \to \Re$ such that

$$W(x^M, \theta^M) \equiv F(\langle U_i(x_i, \theta_i) \rangle_{i \in M}).$$

Notice too that each R_i is an "extended" ordering on the space $X_i \times \Theta_i$ of consumption vectors and possible characteristics for i. Indeed, following Tinbergen (1957), Kolm (1972), and Arrow (1983a, ch. 11), assume that $X_i = X$, $\Theta_i = \Theta$ (all $i \in M$), and also that there exists an ordering $\tilde{R}$ on $X \times \Theta$ for which $R_i = \tilde{R}$ (all $i \in M$). Then $\tilde{R}$ is the *fundamental preference ordering*, representing the *fundamental individual norm*. As Confucius is alleged to have said, "All men are born equal; it's their habits that make them different".

---

[8] [Note added in 1996] This terminology is far from ideal. It suggests, quite misleadingly, that an individual's behaviour should determine the individual norm even when that behaviour is ethically inappropriate or morally reprehensible. As pointed out in Hammond (1996), the assumption is really one of *individualism*, requiring the social measure of ethical value to amalgamate measures of different individuals' ethical value.

This extended ordering embraces *interpersonal comparisons* of the form (x, θ) $\tilde{P}$ $(\bar{x}, \bar{\theta})$ — it is better for society to have a θ-person with x than to have a $\bar{\theta}$-person with $\bar{x}$, as discussed in Hammond (1991b).[9]

10. Arrow Social Norms and Independence of Irrelevant Alternatives

So far the society described by θ^M has been variable. In Arrow's *Social Choice and Individual Values*, it is fixed. There is a social ordering $R(\theta^M)$ on X^M and there are individual orderings $R_i(\theta_i)$ on X_i (for all $i \in M$). (Actually, Arrow did not consider individual consumption allocations, but if he had done, this is how his work should be formulated). With procedures like majority voting in mind, Arrow then makes the following crucial assumption:

Independence of irrelevant alternatives (IIA): Let $Z \subset X^M$ be any subset. Write $R : Z$ for the restriction of the ordering R to Z and $R_i : Z$ for the corresponding restriction of R_i. Then IIA is stated as follows:

$$R_i(\theta_i) : Z = R_i(\bar{\theta}_i) : Z \text{ (all } i \in M) \Longrightarrow R(\theta^M) : Z = R(\bar{\theta}^M) : Z.$$

Here, of course, $R_i(\theta_i)$ is the ordering on X^M defined by

$$x^M \ R_i(\theta_i) \ \bar{x}^M \iff (x_i, \theta_i) \ \tilde{R} \ (\bar{x}_i, \theta_i).$$

When Z is the whole of X^M, this implies that there exists an *Arrow social welfare function* (ASWF) f such that, for all θ^M in Θ^M:

$$R(\theta^M) = f(\langle R_i(\theta_i)\rangle_{i \in M}).$$

What motivates this remarkable assumption? One possibility is a kind of consequentialism. Remember that consequentialism requires the social norm in any decision tree to prescribe behaviour whose consequences depend only on the possible consequences in that tree. In a fixed society θ^M, this implies that there is an ordering $R(\theta^M)$ on X^M.

But now, when θ^M changes, how should behaviour in a tree with consequences $Z \times \{\theta^M\}$ relate to behaviour in one with consequences $Z \times \{\bar{\theta}^M\}$? Arrow's contention is that all that

---

[9] In 1985, this sentence ended as follows: "it is better to be a θ-person with x than a $\bar{\theta}$-person with $\bar{x}$." I now wish to avoid such wording because it may be prolonging the common confusion between individual preferences on the one hand and, on the other, ethical measures of the welfare that individuals contribute to society.

matters are the consequences Z and the individuals' preference orderings over Z (or their behaviour norms for personal decision trees with consequences in Z). So, if $\beta_i(\theta_i) = \beta_i(\bar{\theta}_i)$ (all $i \in M$) for all personal decision trees with consequences in Z, then $\beta(\theta^M) = \beta(\bar{\theta}^M)$ in the tree with consequences Z. This, at least, is a consequentialist justification of IIA, and is certainly of considerable appeal. I shall call a social norm with this property an *Arrow social norm.*

11.   Unrestricted Domain and a Hierarchy of Dictators

The Pareto principle and IIA are two conditions needed for Arrow's impossibility theorem. Two others are:

Unrestricted Domain. For every logically possible preference ordering R_i on X_i, there exists a $\theta_i \in \Theta_i$ such that $R_i(\theta_i) = R_i$.

Nondictatorship. There is no *dictator* $d \in M$ such that, in the fixed society θ^M,

$$x_d \ P_d(\theta_d) \ \bar{x}_d \Longrightarrow x^M \ P(\theta^M) \ \bar{x}^M$$

(regardless of the preferences of others).

These four conditions cannot all be satisfied by the same ASWF. So, given that the first three are, there must be a dictator. Indeed, given (P*), which is stronger than Arrow's Pareto condition, there is a hierarchy of dictators $d_1, d_2, \ldots, d_m$ (where $m := \#M$) such that

$$x^M \ P(\theta^M) \ \bar{x}^M \iff \text{ there exists } k \in \{\, 1, 2, \ldots, m \,\}$$
$$\text{such that } x_{d_r} \ I_{d_r}(\theta_{d_r}) \ \bar{x}_{d_r} \text{ for } r = 1, 2, \ldots, k-1$$
$$\text{and } x_{d_k} \ P_{d_k}(\theta_{d_k}) \ \bar{x}_{d_k}.$$

So d_1 decides unless he is indifferent; then d_2 decides unless he is also indifferent; etc. And $x^M \ I(\theta^M) \ \bar{x}^M \iff x_i \ I_i(\theta_i) \ \bar{x}_i$ (all $i \in M$).

The unrestricted domain condition is rather different from Arrow's, because of ethical liberalism, and the proof also has to be very slightly different — for details, see Hammond (1987). To summarize:

A consequentialist Arrow social norm satisfying ethical liberalism is dictatorial in each fixed society unless the domain of possible preferences is (actually quite severely) restricted.

In particular, restrictions which are common in economics, such as monotonicity, convexity and continuity of preferences, do not let us avoid a dictatorship if there are at least two commodities in the bundle represented by each x_j. For (later) discussion of dictatorship results for such "economic domains", see especially Bordes and Le Breton (1989, 1990).

12. Independence of Ethically Irrelevant Alternatives

Let us accept consequentialism, without which concepts of distributive justice are elusive. Let us also accept ethical liberalism, without which individuals count for little. Then, unless the domain of possible preferences is severely restricted, it is Arrow's independence of irrelevant alternatives that leads to a dictatorship. How should IIA be relaxed?

An interesting requirement is that there should be more interpersonal symmetry in the set of relevant consequences. Given any $\theta^M \in \Theta^M$, define the interpersonal ordering $\bar{R}(\theta^M)$ on $X^M \times M$ by

$$(x^M, j) \ \bar{R}(\theta^M) \ (\bar{x}^M, k) \iff (x_j, \theta_j) \ \tilde{R} \ (\bar{x}_k, \theta_k).$$

Say that R satisfies *independence of ethically irrelevant alternatives* if, for any $Z \subset X$, the restriction $R(\theta^M) : Z$ depends only upon the restriction $\bar{R}(\theta^M) : Z \times M$. Let

$$\zeta(Z, M) := \{ (x, \theta) \in X \times M \mid \exists x^M \in Z; \exists i \in M : (x_i, \theta_i) = (x, \theta) \}$$

be the set of *individual consequences*, all of which are possible given Z and the fixed society θ^M. These are the *ethically relevant alternatives* — and they count insofar as $\bar{R}(\theta^M) : Z \times M$ depends on $\tilde{R} : \zeta(Z, M)$. For example, in choosing between the two income distributions (y^A, y^B) and $(\bar{y}^A, \bar{y}^B)$ in a society with characteristics (θ^A, θ^B), all four possible individual consequences (y^A, θ^A), $(\bar{y}^A, \theta^A)$, (y^B, θ^B), $(\bar{y}^B, \theta^B)$ are relevant — and the fundamental preferences over all four are allowed to count insofar as the restriction of $\tilde{R}$ to these four individual consequences affects the ordering $\bar{R}(\theta^A, \theta^B)$ over (y^A, y^B, A), (y^A, y^B, B), $(\bar{y}^A, \bar{y}^B, A)$, $(\bar{y}^A, \bar{y}^B, B)$ given θ^A and θ^B. This lies right at the heart of ethical arguments, which must be valid even after we put ourselves in the position of the poor and *vice versa.*

This is a new, more restricted, kind of independence condition whose implications have since been more fully explored in Hammond (1991a). The condition *is* consistent, however,

with "Rawlsian maximin", where $R(\theta^M)$ is represented by $\min_{i \in M} u(x_i, \theta_i)$, with u as any "level comparable" interpersonal utility function representing $\tilde{R}$ on $X \times \Theta$. Indeed, I claim:[10]

A consequentialist social norm satisfying ethical liberalism for some fundamental individual norm, as well as independence of ethically irrelevant consequences, and which always prefers individuals worse off than any gainer to gain something as well ("equity") must be Sen's (1970) lexicographic extension of Rawls' maximin rule in any fixed society $\bar{\theta}^M$ unless the domain of preferences on $X^M \times M$ is (actually rather severely) restricted as θ^M varies.

This claim can be justified by first noting that a consequentialist social norm must maximize a social welfare ordering. Furthermore, ethical liberalism implies the strict Pareto condition (P*). Note further that independence of ethically irrelevant consequences is the same as the condition called IIPC in Hammond (1991a), which implies "independence of irrelevant alternatives" for the generalized social welfare functions considered in Hammond (1976a, 1979). The results of the last two cited papers then imply that the social norm must be lexicographic maximin.[11]

This suggests that perfect equality is optimal. But it neglects questions of risk.

13. Risky Consequences and Fundamental Utilitarianism

Suppose that decision trees with risk are allowed, and that behaviour satisfies a weak continuity property as probabilities vary. Then, under the same two assumptions as those mentioned in the last paragraph of Section 7,[12] consequentialism can be shown to imply behaviour that maximizes the expected value of a von Neumann–Morgenstern (or NM) utility function (see Hammond, 1988a).

Thus, there is an NM social welfare function w on $X^M \times \Theta^M$ such that the social norm maximizes the expectation $\mathbb{E}w(x^M, \theta^M)$. And there is an NM individual welfare function v on $X \times \Theta$ such that the fundamental individual norm maximizes $\mathbb{E}v(x, \theta)$. Ethical liberalism implies that there is a function $F : \Re^M \to \Re$, strictly increasing in all its arguments, such that

$$\mathbb{E}w(x^M, \theta^M) \equiv F(\langle \mathbb{E}v(x_i, \theta_i) \rangle_{i \in M})$$

---

[10] For the sake of clarity, this paragraph and the previous one have been slightly expanded from the 1985 version.

[11] This entire paragraph is a 1995 addition to the 1985 version.

[12] This important qualification was not made explicit in the 1985 version.

13

and in particular

$$w(x^M, \theta^M) \equiv F(\langle v(x_i, \theta_i)\rangle_{i \in M}).$$

As John Harsanyi first showed in his famous 1955 paper, this implies that F is linear, so that (ignoring irrelevant additive constants)

$$w(x^M, \theta^M) \equiv \sum_{i \in M} \omega_i \, v(x_i, \theta_i)$$

where each $\omega_i > 0$ — see also Hammond (1992, 1996), as well as the works cited therein. Equal treatment of individuals then requires that ω_i is the same for all i, so we can take $\omega_i = 1$ (all $i \in M$). Thus,

$$w(x^M, \theta^M) \equiv \sum_{i \in M} v(x_i, \theta_i)$$

which I shall call *fundamental utilitarianism.*

What happened to independence of ethically irrelevant alternatives, implying the lexicographic extension of maximin? Well, it appears that, with risky consequences, independence of ethically irrelevant alternatives (together with the other assumptions of our last result) excludes weak continuity of behaviour as probabilities vary. As shown formally in Hammond (1991a, p. 14), it certainly excludes expected social welfare maximization. So:

Independence of ethically irrelevant mixed alternatives. Let $\mathcal{M}_0(\cdot)$ denote the set of simple probability measures (finite probability distributions) on a given space. The new independence condition is:

For any subset Z of $\mathcal{M}_0(X^M)$, the restriction $R(\theta^M) : Z$ of $R(\theta^M)$ to Z depends only on $\bar{R}(\theta^M) : \mathcal{M}_0(Z \times M)$ for every $\theta^M \in \Theta^M$, where $\bar{R}(\theta^M)$ on $\mathcal{M}_0(X^M \times M)$ is the ordering represented by the expected value of $U(x^M, i) := v(x_i, i)$.

This independence condition *is* satisfied by fundamental utilitarianism, which was motivated by other, perhaps more appealing conditions.

14.   Differences from Harsanyi's Utilitarianism

Although fundamental utilitarianism is formally very similar to a version of utilitarianism due to Vickrey (1945, 1960, 1961) and Harsanyi (1953, 1955, 1976, 1978), there are some differences. For one thing, there is no presumption of an original position in which each individual has to decide from X^M after imagining that he is unsure who he will be (what characteristic he has). Individuals must forget themselves in an original position, and this may not be plausible.

Instead, there is a fundamental utility function on $\mathcal{M}_0(X \times \Theta)$. Consequences in this set really can occur, and although in any fixed society θ^M, no individual actually has to choose between, say (x, θ) and $(\bar{x}, \bar{\theta})$ with $\bar{\theta} \neq \theta$, nevertheless I claim that preferences between these two (and probability mixtures of them) are *ethically relevant.* As explained at the end of Section 9, such preferences should correspond to ethical opinions concerning what characteristics it would be better for individuals to have.[13]

Given ethical liberalism — how else are individuals norms to count? — and anonymity, one is then led inexorably to fundamental utilitarianism.

15.   Conclusions

Fundamental utilitarianism orders personal characteristics as well as economic allocations, etc. It determines interpersonal (or inter-characteristic) comparisons of utility.

It does appear to imply basing concepts of distributive justice on peoples' attitudes to risk, which Arrow (1963, p. 10; 1983, p. 48), for one, has often questioned. But v is an *ethical* von Neumann–Morgenstern utility function, representing the *social norm* for personal issues. It need not coincide with any function representing the individual's actual behaviour. Consequentialism allows paternalism in our objectives, although paternalism has costs which we should not ignore when policies are being chosen.

Another difficulty is that when there are disagreements over the fundamental norm, there is no satisfactory way of resolving them.

If my arguments are accepted, then all we know is the general form of the social welfare objective — but this is still too imprecise to be useful. We still do not know which point

---

[13]   The last sentence was added in 1995.

J of Figure 1 to select, nor how to evaluate projects with any great degree of confidence. That awaits developments in ethics as much as in economics. Neither E nor L are likely to be right, however.

In conclusion, do I have any specific recommendation for policy makers? Well, only these two. First, in cost-benefit analysis, using different weights for different income (or wealth) classes is certainly justified. Moreover, most interesting projects are likely to have a significant impact on the distribution of income, even if only at the level of a village in which a "small" project is undertaken — e.g., drilling a deeper well. Such projects, then, themselves affect the welfare weights, and careful calculations are needed to allow for this. Elsewhere, I have suggested some approximations that may be useful — see Hammond (1988c, 1991c) and, for more recent exact measures, Hammond (1994) and also Becht (1995).

The second recommendation concerns taxation. It seems to have become fashionable to regard all taxes as distortionary. Except, of course, the theoretically ideal but hopelessly impractical lump-sum taxes of first-best welfare theory. As far as I am aware, the closest approximation to such taxes that anybody has dared to institute in recent history is the British government's failed experiment in the late 1980s with "community charges," popularly known as the "poll tax". Not surprisingly, such taxes were very difficult to enforce in a modern mobile society, especially one in which gaps have been developing in the welfare safety net. In any case, the point is that the only viable alternatives to complete *laissez faire* necessarily involve financing any public goods or welfare benefits with the proceeds of distortionary taxes. So suitably designed distortionary public finance systems are likely to be better than any *laissez faire* system with no distortions. For this reason, measures of distortion such as deadweight loss are largely illusory. Generally, one cannot avoid measuring directly the welfare gains and losses from changes in one or more governments' tax and expenditure policies.[14]

---

[14] The entire last paragraph was added in 1996.

1995 Postscript: Some Responses to Petretto's Questions

Along with some interesting remarks, Petretto (1986) asked several questions which deserve a belated response, so far as I am able to give one. I pass over questions that relate more closely to the companion paper Hammond (1986), or which I have already tried to answer in the revised version of the lecture.

Question 1 [pp. 173–4] *Should the static model, which is typical of the Arrow–Sen tradition in social choice theory, be considered inadequate for representing practical issues of public policy? Or should one consider the static model, represented by a tree with only a single decision node, as only a special case of the dynamic model?*[15]

In the first place, it is quite true that the formal models discussed in Arrow (1963) and Sen (1970, 1986), for example, are all static. That is, they are trees that contain only a single decision node, in effect, and so are in this sense special cases of the "dynamic model" which allows a general finite decision tree with many decision nodes. Of course, the last sentences of Arrow's book do defend collective rationality on the grounds that a choice process should have certain stability properties. However, there is no mention of decision trees in Arrow's work or in the subsequent attempts by Plott (1973), Campbell (1978) and others to model "path independence".

Actually, the consequentialist axiom discussed in the paper and elsewhere is precisely a condition for the static model to be adequate. If it is not satisfied, then the outcome of the choice process is likely to depend on features of the decision tree that cannot be represented within any static model. In fact, one of the best defences of consequentialism in risky decision trees is to notice that, if it is not satisfied, a project that is started after passing an initial cost-benefit test risks being abandoned later after a revised cost-benefit test. There is a clear analogy here to the "potential addict" example of Hammond (1976b). For further details, see also Hammond (1988d).

So, where consequentialism is violated, a decision-maker has to be concerned with possible later changes of taste.

---

[15] The italicized English wording of the first two questions reflects both the typescript that Petretto had prepared at the time of the conference in 1985, as well as my own rather free translation of the somewhat revised proceedings in Italian.

Question 2 [p. 175] *It is legitimate to ask whether it is really true that in practical decision problems the choice should be based on consequences, independently of the structure of the decision tree or other considerations. (Can one say that the consequentialist ends justify the means?) ... Here I have in mind, for example, a prominent question in modern public finance theory, namely the conflict between horizontal equity, defined with reference to [*status quo relative] utility levels as in Feldstein (1976), and welfare maximization in a second best context — see Atkinson (1980), Stiglitz (1982), Balcer and Sadka (1982), Boadway and Bruce (1984).*"*

The concept of horizontal equity discussed in this question is violated if and only if a policy reform has the effect of changing the ordering of different individuals' utility levels. Now, not all horizontally inequitable reforms are bad: for example, where compliance with tax laws has been less than perfect in the past, there is much to be said for a tax reform that allows those who have previously been more honest to advance in the ranking of different individuals' real living standards at the expense of those who have been able to cheat. But I admit that there may be many other cases where horizontal equity is ethically desirable. In principle, however, considerations of horizontal equity can be accommodated fairly easily within a consequentialist framework. All that one needs to do is make every individual's welfare a function not only of their economic circumstances after reform, but also of what real standard of living they could reasonably have expected without any reform, as well as of the rankings before and after the reform of different individuals' real standards of living. Naturally, this considerably complicates the description of each relevant consequence in the domain on which each individual's welfare function is defined. Nevertheless, no fundamental principle is at stake.

A deeper issue is whether consequences can always be defined independently of the structure of the decision tree. In particular, can consequences be defined in a way that does not impose natural restrictions on the domain of decision trees? In fact, there can be difficulties. For example, suppose that each consequence includes within itself a complete description of the collapsing sets of possible consequences that the agent has faced at each successive decision node. This may be important in discussions of rights, to the extent that they can be represented by the agent's range of choice. Then the problem is that the set of possible consequences corresponds uniquely to the structure of decision tree, except

that parts of the tree and the range of their possible consequences may be duplicated. So the important assumption of an unrestricted domain — or at least of a domain sufficiently unrestricted to allow the arguments set out in Hammond (1988a) to retain their validity — is not entirely innocuous. See also Munier (1996) and Hammond (forthcoming).

Question 3   In Section 3 of his comments, Petretto asks about the significance and use of individual welfare norms. As was explained in Section 9, these are meant to signify appropriate *social* choices in decision trees whose consequences are relevant only to one individual. By considering different characteristics which this one individual might have, the preference ordering revealed by a consequentialist individual welfare norm will involve inter-characteristic comparisons between different types of individual.

However, I am now ready to admit that such inter-characteristic comparisons will fail to generate comparisons between persons whose domains of possible characteristics are entirely disjoint — e.g., people of different race or gender, insofar as race or gender are welfare relevant characteristics. This is one reason why, in later work such as Hammond (1991b, 1992), I have chosen instead to base interpersonal comparisons on ethical preferences regarding the frequency distribution of different types of people in society. Even then, such interpersonal comparisons can create difficulties. Many people might think that some reasonable numerical balance between men and women in the population is desirable. In which case the relevant individual welfare function will depend not only upon a person's own gender, but also on the proportions of the two genders.

Question 4   Section 5 of Petretto's comments concern the contrast between consequentialism and Rawls — or, more exactly, the leximin criterion. In particular, Petretto asks why leximin is excluded by consequentialist decision theory.

The reason for this exclusion relates to Harsanyi's (1975a, b, 1978) objection to "Rawlsian" maximin — that it violates the standard axioms for decision making under risk. The consequentialist axioms are somewhat different, but they imply some of the key standard axioms. In particular, maximin also violates consequentialism.

Of course, leximin can emerge as a limit of fundamental utilitarianism in the case of extreme aversion to inequality in different people's ordinal utility levels — cf. Hammond (1975). Indeed, my understanding is that recently Rawls himself has become less insistent

on his "difference principle," seeing it now as a limiting case of extreme risk aversion in the original position that is a feature of both his and Harsanyi's approaches to ethical decision-making.

Acknowledgements

The first version of this paper appeared in July 1985 as a somewhat revised version of notes used for a January 1985 lecture in Milan that was presented to the Istituto per la Comunicazione Scientifica (ICOS) meeting on "Decision-Making in Administration and in Complex Organizations: Theories, Models and Experiences". An Italian translation was published as one chapter of a supplement to the tri-annual periodical *Il nuovo governo locale*. See Sacconi (1986), ch. 3, pp. 89–106.

I am most grateful to Lorenzo Sacconi of ICOS for inviting me to present the earlier lecture and for organizing an excellent meeting. Also to Alessandro Petretto for the helpful comments he made at the conference in Milan, some of which were published in Italian as Petretto (1986) in the conference volume cited above. And finally to an anonymous referee for noticing in particular where some additional explanation would be useful.

For this later English version, I have added an additional paragraph in the conclusions and also the postscript. I have also made several minor changes to the original English typescript, the most important of which are indicated in notes at the end. Also, having been warned that the U.S. publishers in 1996 cannot match the Italians' ability in 1985 to include footnotes, several of the original footnotes have been incorporated into the main text. Finally, there are also several additional references to some more technical papers that have appeared since 1985.

References

K.J. ARROW (1951, 1963) *Social Choice and Individual Values (2nd edn.)* (New Haven: Yale University Press).

K.J. ARROW (1983) *Collected Papers of Kenneth J. Arrow, Vol. 1: Social Choice and Justice* (Cambridge, Mass.: Harvard University Press).

A.B. ATKINSON (1980) "Horizontal Equity and the Distribution of the Tax Burden," in H.J. Aaron and M.J. Boskin (eds.) *The Economics of Taxation* (Washington, D.C.: The Brookings Institution), pp. 3–18.

Y. Balcer and E. Sadka (1982) "Horizontal Equity, Income Taxation and Self-Selection with an Application to Income Tax Credits", *Journal of Public Economics*, vol. 19, pp. 291–309.

E. Barone (1908) "Il ministero della produzione nello stato colletivista", *Giornale degli Economisti*, vol. 37, pp. 267–293 and 391–414; translated as "The Ministry of Production in the Collectivist State" in F.A. Hayek (ed.) (1935), *Collectivist Economic Planning* (London: Routledge & Kegan Paul), pp. 310–334; reprinted in P. Newman (ed.), *Readings in Mathematical Economics, Vol. I: Value Theory* (Baltimore: The Johns Hopkins Press).

M. Becht (1995) "The Theory and Estimation of Individual and Social Welfare Measures", *Journal of Economic Surveys*, vol. 9, pp. 53–87.

A. Bergson (1938) "A Reformulation of Certain Aspects of Welfare Economics", *Quarterly Journal of Economics*, vol. 52, pp. 310–334.

R.W. Boadway and N. Bruce (1984) *Welfare Economics* (Oxford: Basil Blackwell).

G. Bordes and M. Le Breton (1989) "Arrovian Theorems with Private Aternatives Domains and Selfish Individuals", *Journal of Economic Theory*, vol. 47, pp. 257–281.

G. Bordes and M. Le Breton (1990) "Arrovian Theorems for Economic Domains. The Case where there are Simultaneously Private and Public Goods", *Social Choice and Welfare*, vol. 7, pp. 1–17.

D.E. Campbell (1978) "Realization of Choice Functions," *Econometrica*, **46**, pp. 171–180.

J.S. Chipman and J.C. Moore (1978) "The New Welfare Economics, 1939–1974", *International Economic Review*, vol. 19, pp. 547–584.

J.S. Chipman and J.C. Moore (1980) "Compensating Variation, Consumer's Surplus, and Welfare", *American Economic Review*, vol. 70, pp. 933–949.

J.L. Coles and P.J. Hammond (1995) "Walrasian Equilibrium without Survival: Equilibrium, Efficiency, and Remedial Policy", in K. Basu, P.K. Pattanaik, and K. Suzumura (ed.) *Choice, Welfare and Development: A Festschrift in Honour of Amartya K. Sen* (Oxford: Oxford University Press) ch. 3, pp. 32–64.

M.S. FELDSTEIN (1976) "On the Theory of Tax Reform", *Journal of Public Economics*, vol. 6, pp. 77–104.

P.J. HAMMOND (1975) "A Note on Extreme Inequality Aversion", *Journal of Economic Theory*, vol. 11, pp. 465–467.

P.J. HAMMOND (1976a) "Equity, Arrow's Conditions, and Rawls' Difference Principle", *Econometrica*, vol. 44, pp. 793–804.

P.J. HAMMOND (1976b) "Changing Tastes and Coherent Dynamic Choice", *Review of Economic Studies*, vol. 43, pp. 159–173.

P.J. HAMMOND (1979) "Equity in Two-Person Situations: Some Consequences", *Econometrica*, vol. 47, pp. 1127–35.

P.J. HAMMOND (1982) "*The Economics of Justice* and the Criterion of Wealth Maximization", *Yale Law Journal*, vol. 91, pp. 1493–1507.

P.J. HAMMOND (1986) "Consequentialist Social Norms for Public Decisions," in W.P. Heller, R.M. Starr and D.A. Starrett (eds.), *Social Choice and Public Decision Making: Essays in Honor of Kenneth J. Arrow, Vol. I* (Cambridge: Cambridge University Press), ch. 1, pp. 3–27.

P.J. HAMMOND (1987) "On Reconciling Arrow's Theory of Social Choice with Harsanyi's Fundamental Utilitarianism," in G.R. Feiwel (ed.) *Arrow and the Foundations of the Theory of Economic Policy* (Macmillan and New York University Press), ch. 4, pp. 179–221.

P.J. HAMMOND (1988a) "Consequentialist Foundations for Expected Utility", *Theory and Decision*, vol. 25, pp. 25–78.

P.J. HAMMOND (1988b) "Consequentialist Demographic Norms and Parenting Rights", *Social Choice and Welfare*, vol. 5, pp. 127–145.

P.J. HAMMOND (1988c) "Principles for Evaluating Public Sector Projects," in P. Hare (ed.) *Surveys in Public Sector Economics* (Oxford: Basil Backwell), ch. 2, pp. 15–44.

P.J. HAMMOND (1988d) "Orderly Decision Theory: A Comment on Professor Seidenfeld," *Economics and Philosophy*, **4**, pp. 292–297.

P.J. HAMMOND (1991a) "Independence of Irrelevant Interpersonal Comparisons," *Social Choice and Welfare*, **8**, pp. 1–19.

P.J. HAMMOND (1991b) "Interpersonal Comparisons of Utility: Why and How They Are and Should Be Made," in J. Elster and J.E. Roemer (eds.), *Interpersonal Comparisons of Well-Being* (Cambridge: Cambridge University Press, 1991), ch. 7, pp. 200–254.

P.J. HAMMOND (1991c) "Approximate Measures of the Social Welfare Benefits of Large Projects," Stanford University, Institute of Mathematical Studies in the Social Sciences, Economics Technical Report No. 410 (1983); in S. Dahiya (ed.) *Theoretical Foundations of Development Planning, Vol. V* (New Delhi: Vedams Books International).

P.J. HAMMOND (1992) "Harsanyi's Utilitarian Theorem: A Simpler Proof and Some Ethical Connotations", in R. Selten (ed.) *Rational Interaction: Essays in Honor of John Harsanyi* (Berlin: Springer-Verlag, 1992), pp. 305–319.

P.J. HAMMOND (1994) "Money Metric Measures of Individual and Social Welfare Allowing for Environmental Externalities," in W. Eichhorn (ed.) *Models and Measurement of Welfare and Inequality* (Berlin: Springer-Verlag) pp. 694–724.

P.J. HAMMOND (1996) "Consequentialist Decision Theory and Utilitarian Ethics," in F. Farina, F. Hahn, and S. Vannucci (eds.) *Ethics, Rationality, and Economic Behaviour* (Oxford: Clarendon Press), pp. 92–118.

P.J. HAMMOND (in preparation) "Disorderly Behaviour with Opportunistic Consequences".

P.J. HAMMOND AND J. SEMPERE (1995) "Limits to the Potential Gains from Economic Integration and Other Supply Side Policies," *Economic Journal* **105**, pp. 1180–1204.

A.C. HARBERGER (1971) "Three Basic Postulates for Applied Welfare Analysis", *Journal of Economic Literature*, vol. 9, pp. 785–797.

J.C. HARSANYI (1953) "Cardinal Utility in Welfare Economics and in the Theory of Risk-Taking", *Journal of Political Economy*, vol. 61, pp. 434–5; reprinted in Harsanyi (1976).

J.C. HARSANYI (1955) "Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility", *Journal of Political Economy*, vol. 63, pp. 309–321; reprinted in Phelps (1973) and Harsanyi (1976).

J.C. HARSANYI (1975a) "Nonlinear Social Welfare Functions: Do Welfare Economists Have a Special Exemption from Bayesian Rationality?", *Theory and Decision*, vol. 6, pp. 311–32; reprinted in Harsanyi (1976).

J.C. HARSANYI (1975b) "Can the Maximin Principle Serve as a Basis for Morality? A Critique of John Rawls's Theory", *American Political Science Review*, vol. 69, pp. 594–606; reprinted in Harsanyi (1976).

J.C. HARSANYI (1976) *Essays on Ethics, Social Behavior, and Scientific Explanation* (Dordrecht: D. Reidel).

J.C. HARSANYI (1978) "Bayesian Decision Theory and Utilitarian Ethics", *American Economic Review (Papers and Proceedings)*, vol. 68, pp. 223–8.

J.C. HARSANYI (1986) "Un nuovo concetto generale di soluzione per i giochi cooperativi e non cooperativi", in Sacconi (1986), ch. 1, pp. 35–59.

J.R. HICKS (1981) *Wealth and Welfare: Collected Essays on Economic Theory, Vol. 1* (Cambridge, Mass: Harvard University Press).

N. KALDOR (1939) "Welfare Propositions of Economics and Interpersonal Comparisons of Utility," *Economic Journal*, **49**, pp. 549–552.

S.-C. KOLM (1972) *Justice et Equité* (Paris: Editions du Centre National de la Recherche Scientifique).

B.R. MUNIER (1996) "Comment" in K.J. Arrow, E. Colombatto, M. Perlman, and C. Schmidt (eds.) *The Rational Foundations of Economic Behaviour* (London: Macmillan), pp. 43–47.

A. PETRETTO (1986) "Commento", in Sacconi (ed.), ch. 9, pp. 172–180.

E.S. PHELPS (ED.) (1973) *Economic Justice: Selected Readings* (Harmondsworth: Penguin Books).

C.R. PLOTT (1973) "Path Independence, Rationality, and Social Choice," *Econometrica*, **41**, pp. 1075–1091.

R.A. POSNER (1981) *The Economics of Justice* (Cambridge, Mass.: Harvard University Press).

W.G. RUNCIMAN (1966) *Relative Deprivation and Social Justice: A Study of Attitudes to Social Inequality in Twentieth Century England* (London: Routledge & Kegan Paul; Harmondsworth: Penguin).

L. SACCONI (ED.) (1986) *La Decisione: Razionalità collectiva e strategia nell'amministrazione e nelle organizzazioni* (Milan: Franco Angeli).

A.K. SEN (1970) *Collective Choice and Social Welfare* (San Francisco: Holden Day).

A.K. SEN (1986) "Social Choice Theory," in K.J. Arrow and M.D. Intriligator (eds.) *Handbook of Mathematical Economics, Vol. III* (Amsterdam: North-Holland), ch. 22, pp. 1073–1181.

J.E. STIGLITZ (1982) "Utilitarianism and Horizontal Equity: The Case for Random Taxation", *Journal of Public Economics*, vol. 18, pp. 1–34.

J. TINBERGEN (1957) "Welfare Economics and Income Distribution", *American Economic Review (Papers and Proceedings)*, vol. 47, pp. 490–503.

W.S. VICKREY (1945) "Measuring Marginal Utility by Reactions to Risk", *Econometrica*, vol. 13, pp. 319–33.

W.S. VICKREY (1960) "Utility, Strategy and Social Decision Rules", *Quarterly Journal of Economics*, vol. 74, pp. 507–35.

W.S. VICKREY (1961) "Risk, Utility and Social Policy", *Social Research*, vol. 28, pp. 205–217; reprinted in Phelps (1973).