1 Introduction and Outline

1.1 Interpersonal Comparisons: Some Background

Over many years, interpersonal comparisons of utility have had a significant role to play in economics. Utility began as a basic concept on which Frances Hutcheson, Cesare Beccaria, Jeremy Bentham, John Stuart Mill, and Henry Sidgwick sought to build a general ethical theory that is simple yet profound. The resulting classical utilitarian theory relied on interpersonal comparisons because it required a common unit with which to measure each person's pleasure or happiness, before adding to arrive at a measure of total happiness. According to the standard reading of Bentham, one should then proceed to subtract each person's pain or misery, also measured in the same common unit, in order to arrive at a measure of total utility.[1]

For economists, the notion of utility later became much more sophisticated. In the Benthamite tradition, consumer demand theory had been based on a cardinal notion of utility, and on the requirement that the marginal utilities of spending wealth on different commodities should be equalized. Following the ideas pioneered by Pareto, Hicks, Allen, and Samuelson, a revised demand theory was built on the more basic concept of a binary preference relation, perhaps revealed by the consumer's own behaviour. In positive economics this meant that utility became an ordinal rather than a cardinal concept. It also implied that one lacked a common unit with which to measure and compare different individuals' utilities. This allowed Robbins (1932, 1938) fo feel justified in making his widely cited claim that interpersonal comparisons of utility are unscientific.

Welfare economic theory, however, and the related discipline of social choice theory, have retained their links to ethics. In fact, without their ethical content, both theories would become empty shells, as Little (1957, pp. 79–80) for one has pointed out. For this reason, interpersonal comparisons continue to play a significant role in both these theories. But as we shall see, the utility concept has been submitted to a further twist, as for many authors it measures the value to the social planner or the ethical observer of each individual's lifetime history, rather than the personal values of their lifetime histories to ordinary individuals themselves.

[1]Actually, it seems plausible that Bentham regarded "utility" as an objective property of things, rather than as a measure of subjective pleasure minus pain. See Mongin and d'Aspremont (1998).

1.2 Outline of Chapter

This chapter will not attempt a proper survey of the large literature on interpersonal comparisons. The main reason for this is to avoid repeating what has already appeared in Hammond (1991a) or Suzumura (1996). Instead, we would like to focus attention on three specific questions which arise in connection with interpersonal comparisons.

Of these three questions, the first is why economists need these particular value judgements that Robbins deemed unscientific. In fact, what would remain of welfare economics and of social choice theory if one refused to make any interpersonal comparisons at all?

The second question relates to the first, because it asks what can be done with interpersonal comparisons. Section 2 begins by arguing that much can be achieved in welfare economics without such comparisons, at least with respect to utility. It also points out how, in welfare economics, they can be used to answer distributional questions such as what weights to place on different individuals' marginal gains and losses. In social choice theory, however, as discussed in Section 3, one has to cope with Kenneth Arrow's famous "dictatorship" theorem. Following Arrow's own reading, this result is usually interpreted as proving that a reasonable social choice procedure is impossible in general without interpersonal comparisons of utility. We shall examine the roots of this interpretation, which are closely related to what Hicks (1959) and Sen (1977) in particular have called "welfarism".

Next, Section 4 illustrates how interpersonal comparisons allow many possible escapes from Arrow's theorem, depending upon whether one can make comparisons of utility levels or of utility units. It examines various ways in which such interpersonal comparisons of utility are rendered possible by weakenings of Arrow's restrictive conditions. The section closes with an inquiry into what it means in general when one says that a social criterion "relies on interpersonal comparisons" of any kind (not necessarily of utility).

The third and last question may well strike the reader as being the most important. To the extent that interpersonal comparisons of some sort are unavoidable or at least desirable, how can they be made, and what meaning can they be given? Section 5 starts with an examination of the respective parts played by normative value judgements and factual statements in the making of interpersonal comparisons. Unsurprisingly, it turns out that value judgements are essential, although most kinds of interpersonal comparison do also require objective data about individual situations. The rest of Section 5 proceeds through a series of examples, and analyses the way in which practical interpersonal comparisons are made in each. Section 6 continues this theme with an extended discussion of examples that relate specifically to applied welfare economics. We hope that these examples will help dispel the common feeling among economists that interpersonal comparisons require obscure and contentious value judgements that are better left to political philosophers.

Then, following Hammond (1987b, 1991a) and using the standard model of expected utility described in Chapter 5, here Section 7 shows how the acceptance of basic ethical principles may constrain the mathematical structure of social preferences in such a way that interpersonal comparisons acquire straightforward interpretations in terms of social decisions. Specifically, the interpersonal comparisons embodied in the ethical decision criterion may then be seen as revealed by the choice of persons — or better, by the ethical choice of a lottery determining population size and the distribution of personal characteristics within the population.

Section 8 contains some concluding remarks.

2 Welfare Economics

2.1 Pareto Efficiency

Welfare economics is an enormous subject, touching every branch of economic science. Here, we shall not attempt more than to summarize a few of the most crucial results, while emphasizing how many of them do not rely on interpersonal comparisons at all.

Modern welfare economics, like modern social choice theory, begins with an article by Kenneth Arrow. In 1951, he presented the two fundamental theorems of the subject. Of course, there were antecedents in well known classic works by Enrico Barone, Vilfredo Pareto, Oskar Lange, Abba Lerner, Paul Samuelson, Maurice Allais, and others. But these earlier authors limited themselves to incomplete and local results based on the differential calculus. Whereas Arrow's analysis was global, exploiting the notion of convexity and the separating hyperplane theorem.

According to the first of these two theorems, each Walrasian equilibrium allocation is Pareto efficient, at least if consumers' preferences are locally nonsatiated. According to the second theorem, any Pareto efficient allocation not on the boundary of the attainable set is a Walrasian equilibrium, provided that preferences satisfy appropriate convexity and continuity assumptions. For present purposes, it is enough to recall that these two theorems relate the set of Pareto efficient allocations to the set of Walrasian equilibria with lump-sum transfers. To describe either of these two sets, there is evidently no need for interpersonal comparisons. Such comparisons serve only to choose among the elements of each set, which is really a social choice problem anyway.

2.2 Pareto Improvements

In each of the world's contemporary national economies, not to mention the global economy as a whole, there remain many imperfections which prevent the Pareto efficient allocation of resources. For example, there are public goods, external effects, distortionary taxes of the kind needed to finance public goods and to institute measures that alleviate poverty, etc. These inevitable imperfections limit the relevance to practical economics of the two fundamental efficiency theorems. In fact, these theorems are too idealistic because they characterize allocations which are perfect—or at least perfectly efficient.

For this reason, the results concerning the gains from free trade and free exchange might appear to be much more useful. Most economists think of these as belonging to the field of international economics. But there is a general third theorem of welfare economics concerning not only the gains from international trade, but also the gains from market integration, from profit maximization by a firm, from enhanced free competition between firms, from replacing a distortionary tax with lump-sum taxes raising the same revenue, from a small project that passes a cost–benefit test at suitable (producer) prices, and from technical progress that enhances the efficiency of production. All these are really instances of one general theorem, as pointed out in Hammond and Sempere (1995).

This third theorem shows that, if a new market is opened, or if existing markets are made more efficient, there is a potential Pareto improvement in the sense described originally by Barone (1908), though more commonly ascribed to Kaldor (1939) and Hicks—see the articles the latter published during the years 1939–1946 that are reprinted in Part II of Hicks (1981).[2] That is, even if some people initially lose because of adverse relative price movements caused by the new markets or by the increase in efficiency, they can always be compensated so that everybody gains in the end. Thus, an actual Pareto improvement becomes possible. But in this connection, one is always looking for a Pareto improvement, in which everybody gains and nobody loses. In this way, the need for interpersonal comparisons has still been avoided.

2.3 Private Information

These three classical theorems all rely on the assumption that lump-sum redistribution is possible without limit. Yet in reality we lack the information needed to arrange such redistribution in a suitable manner. As Vickrey (1945) and Mirrlees (1971) understood very well in their analyses of optimal income

---

[2]For an assessment of Barone's earlier contribution, see Chipman and Moore (1978) and Chipman (1987).

taxation, it is impossible to have ideal lump-sum taxes based on relevant characteristics such as workers' inherent abilities. These abilities cannot be observed. Instead, Vickrey and Mirrlees assume one sees only the incomes which workers can earn by deploying their abilities. So, instead of an ideal tax on inherent ability, one is forced to substitute a distortionary tax on income.[3]

A worker's inherent ability is merely one kind of private information. There are many other kinds—for example, a consumer's preferences and endowments, or a producer's true technology and associated cost function. Each piece of private information creates its own "incentive" constraint, limiting how that information can be used to affect the economic allocation. Guesnerie (1995) and Hammond (1990) have independently analysed general economies with very many agents who possess some private information. They have shown how welfare-improving lump-sum transfers generally depend on private information. And how incentives are preserved only by what public finance economists generally regard as "distortionary" taxes that depend on individual transactions, as well as on the distribution of privately known personal characteristics in the population. Then the two theorems linking Pareto efficient allocations to perfect markets lose virtually all their relevance. The usual Pareto frontier becomes replaced by a "second-best" Pareto frontier, which recognizes incentive constraints as well as the usual requirements of physical feasibility. Further discussion and references can be found in Hammond (1990).

Guesnerie and Hammond have also considered what would remain possible if individuals could manipulate not only by concealing or misrepresenting their private information, but if they could also combine in small groups with other individuals in order to exchange goods on the side, in a hidden economy beyond the control of the fiscal authorities. These extra manipulations imply that one can have only linear relative prices for each pair of goods whose exchange cannot be observed by the authorities. In this way, extra constraints arise and one is forced down to a "third-best" Pareto frontier. However, in the absence of externalities or public goods, all three frontiers contain whatever allocations would result from a policy of total *laisser faire*, without any attempts to redistribute wealth in order to move around the first-best frontier. See also Blackorby and Donaldson (1988), as well as Hammond (1999).

We still lack simple or intuitive economic characterizations of the constrained Pareto frontiers. There are no fundamental theorems like the two proved by Arrow. Nevertheless, it is evident that any such constrained Pareto frontier can be described without the need to make any interpersonal comparisons at

[3]If hours worked could also be observed, then skill could be inferred. Even so, incentive constraints would still prevent the economy from attaining its first-best outcome, except when the objective happens to be Rawlsian maximin. For details, see Dasgupta and Hammond (1980).

all. Both the Pareto criterion and the relevant incentive constraints can be described by making use of information only about individual preference orderings. Only the ethical social choice of a point or subset of the frontier may possibly require interpersonal comparisons.

The third theorem that was discussed in Section 2.2 is much less modified than the first two when one takes account of private information and the resulting incentive constraints. Following an idea due to Dixit and Norman (1980, 1986) that arises fairly naturally out of the work of Diamond and Mirrlees (1971), it is shown in Hammond and Sempere (1995) that Pareto improvements can still be ensured if the tax on each commodity is varied in a way that freezes the after-tax prices (and wages) faced by all consumers; this still allows prices faced by producers to vary in order to clear markets. In addition, after-tax dividends paid by firms to consumers should be frozen.[4] But any result of this kind concerns actual or potential Pareto gains, and so still avoids any need for interpersonal comparisons.

2.4 Measures of Individual Gain and Loss

So far, we have argued that the major theorems of Paretian welfare economics do not rely on interpersonal comparisons. But these major theorems cannot be applied easily to real issues of economic policy, such as how to provide affordable medical services, or lower unemployment, or reduce poverty, or provide more adequate housing, while avoiding excessive taxes or risks of high inflation. According to the familiar old proverb, "It is an ill wind that blows nobody any good." This applies even in economics. For example, a deep recession brings a lot of business for corporate lawyers, accountants and others who are responsible for winding up bankrupt firms. The reverse is: "It is a good wind that blows nobody any ill"—in other words, it is difficult to find a true Pareto improvement. In practice, real economic policy choices make some people better off, others worse off. The choice between policies then may require interpersonal comparisons.

Still, a great deal can be learned about the effects of economic policy choices even without interpersonal comparisons. This is because any economic policy reform or decision can be regarded as having effects on each separate individual. So one should be able to calculate or estimate each individual's net benefit from any policy decision. In principle, it is usually possible even to construct a money metric measure of net benefit. This is done by finding what increase or decrease in wealth would have exactly the same effect on the individual's welfare as the policy decision being contemplated, provided that private good

[4]See Hammond (2000) for further extensions to economies with a continuum of consumers, non-linear budget sets, indivisible goods, etc.

prices and public good quantities remained fixed at their *status quo* values. It is *not* done, except possibly very inaccurately, by calculating consumer surplus based on the area under an uncompensated Marshallian demand curve. For details, see Hammond (1994) or Becht (1995), amongst others. The measure that results is closely related to Hicks' equivalent variation. It tells us how much each particular individual gains or loses from a policy change, which is immensely valuable information. Yet the construction of different individuals' measures of net benefit does not require any interpersonal comparisons.

At this stage, many economists of the so-called "Chicago school", following Harberger (1971) in particular, succumb to the temptation of just adding different individuals' monetary measures. "A dollar is a dollar", they might say, regardless of how deserving is the recipient. Implicitly, they attach equal value to the extra dollar a rich man will spend on a slightly better bottle of wine and to the dollar a poor woman needs to spend on life-saving medicine for her child. Of course, any such judgement is a value judgement, even an interpersonal comparison, which lacks scientific foundation.[5] Thus, the "surplus economists" who just add monetary measures, often of consumer surplus rather than individual welfare, make their own value judgements and their own interpersonal comparisons. Moreover, their comparisons not only lack scientific content, but most people—especially non-economists—also find them totally unacceptable from an ethical point of view. Surely it is better to avoid interpersonal comparisons altogether rather than make them in such a biased way.

Many economists, including even Harberger (1978) himself (though very reluctantly), have suggested multiplying each individual's monetary measure of gain by a "welfare weight" in order to arrive at a suitable welfare-weighted total measure of benefit for society as a whole. The ratios of these welfare weights evidently represent the (constant) marginal rates of substitution between the wealth levels of the corresponding individuals in a social welfare function. These ratios reflect interpersonal comparisons between the supposed ethical worth of marginal monetary gains occurring to different individuals, even if one follows the Chicago school in equating all the welfare weights to 1. Such welfare-weighted sums can be used to identify directions in which small enough policy changes are deemed beneficial for society as a whole.

Many economists have also advocated considering welfare-weighted sums even for changes that are not small. Yet policies having a significant impact on the distribution of real wealth are also likely to change the ethically appropriate marginal rates of substitution between different consumers' incomes—the

[5]This does not mean that an analyst who adds up willingness-to-pay or surplus across people must necessarily know or assume anything about individual utility. Distinctions will be introduced later between individual and social utility, and also between interpersonal comparisons in general and interpersonal comparisons of utility specifically.

numbers which lie behind the different relative welfare weights. So one needs to be more careful. This is an issue which is discussed at greater length in Hammond (1994). Of course, interpersonal comparisons will play an inevitable role in determining any suitable set of weights for such measures of social welfare.

In addition, it is worth recalling that additive measures of monetary gain which are intended to identify potential Pareto improvements—for example, the sums of equivalent (or compensating) variation which underlie compensation tests of the kind mentioned at the end of Section 2.2—often fail to provide a consistent basis for complete social preferences. This is because different reference prices are used for different pairs of allocations to be compared. This "intransitivity" problem was actually what in part motivated Arrow's original analysis of the general social choice problem. Indeed, starting with Scitovsky (1941) and Arrow (1951), criticism of these sums-of-surplus criteria mounted during the following decades, culminating in a firm general condemnation by social choice theorists—see Blackorby and Donaldson (1990) for a synthesis. Of course, this has not prevented such criteria from being applied quite often in fields such as international economics and cost–benefit analysis ...

To summarize this section, as long as welfare economics concerns itself only with (constrained or unconstrained) Pareto efficient allocations, or with (actual or potential) Pareto improvements, there is no need for interpersonal comparisons. Even without such comparisons, one can still describe the Pareto frontier, with or without constraints of various kinds, and also look for Pareto improvements. Moreover, it is possible to construct measures of net monetary gain for each separate individual. As discussed in Hammond (1990), such individual measures already provide very useful information; much more is provided by the joint statistical distribution of these measures and of other relevant personal characteristics, such as education, family circumstances, age, or family background. In principle, this joint distribution can and should be estimated by the best possible econometric techniques. It does not depend on any interpersonal comparisons. Its interpretation depends on only one ethical value judgement—namely, the judgement that information about different individuals' reported preferences or actual behaviour can determine how those individuals' measures of benefit should be estimated. That is a serious value judgement, but one which is indispensable for the neo-classical theory of welfare economics. Without this judgement, one would have to consider issues such as how much paternalism is desirable.

In the end, then, much welfare analysis is possible without interpersonal comparisons. They would play a role only, possibly, in choosing among different Pareto efficient allocations. Or more generally, in deciding whether to institute a reform which benefits one set of individuals but harms another. Or when one wants to construct some aggregate measure of social welfare. These considerations lead us to the theory of social choice, our next topic.

We shall return to welfare economics in Section 6 especially.

3 Social Choice without Interpersonal Comparisons

3.1 Arrow's Impossibility Theorem

Like modern welfare economics, modern social choice theory starts with a 1951 publication by Kenneth Arrow—in this case, the first edition of *Social Choice and Individual Values*, based on his Ph.D. thesis submitted to Columbia University. This and the earlier article (Arrow, 1950) presented his famous "impossibility" theorem. Though this result is well known, we will present a variant of it in order to introduce some terminology which will be useful later.

Because Arrow deliberately sought to avoid interpersonal comparisons, he defined a social welfare function on a domain of individual *preference* profiles. But since we want to discuss the issue of interpersonal comparisons of *utility* in some detail, it will be convenient here to adopt a framework with utilities that was introduced by Kolm (1968) and Sen (1970a), before being adopted by d'Aspremont and Gevers (1977) and many successors.

Let X be the universal set of social states defined so that society is required to choose one social state from some feasible subset of X. Let N be a finite set of n individuals. Each individual $i \in N$ has some personal characteristics which, it is assumed, are summarized in a utility function $U_i : X \to \mathbb{R}$. For every social state $x \in X$, the number $U_i(x)$ measures individual i's utility in this state. At this stage it is not necessary to give much substantive meaning to this function. It is just an index supposed to capture all relevant features of individual situations, and to synthesize all these features in a unidimensional way. The utility function measures the individual *good*, whatever that means. For example, utility may be measured in terms of mental states such as pleasure and pain, in the Benthamite tradition. Or in terms of happiness (Sumner 1996). Or, following most economists, in terms of preference satisfaction—see Mongin and d'Aspremont (1998), for example. Also, even though the word "utility" is seldom used in such contexts, it could be an index of more objective notions such as primary goods (Rawls 1971), resources (Dworkin 1981), capabilities (Sen 1985), opportunity for welfare (Arneson 1989), access to advantage (Cohen 1989), etc. In fact, unidimensionality is the only serious ethical restriction here, because such a monistic representation of individual good is not accepted by those holding pluralist views, for whom individual good has several incommensurable dimensions.[6]

Let $\mathcal{U}(X)$ denote the set of all real-valued functions on X, and let $\mathcal{U}^N(X)$ denote the Cartesian product of n different copies of the set $\mathcal{U}(X)$. Each $U_N := \langle U_i \rangle_{i \in N} \in \mathcal{U}^N(X)$ is a *utility profile* consisting of one utility function $U_i \in \mathcal{U}(X)$ for each individual $i \in N$.

A (weak) *preference relation* R on X is a binary relation that is reflexive, but not necessarily transitive or complete. The corresponding strict preference relation, usually denoted by P, is defined so that $x \; P \; y$ iff $x \; R \; y$ and not $y \; R \; x$.

Let $\mathcal{R}(X)$ denote the set of all preference relations on X. Then a *social welfare functional* (SWFL) is a mapping $f : D \to \mathcal{R}(X)$ defined on a domain $D \subset \mathcal{U}^N(X)$ of utility profiles, whose value is some social welfare preference relation on X.

Although Arrow's original framework contained no utility functions, it is straightforward to translate the requirements he imposed on social preferences into corresponding axioms bearing on SWFLs. The central requirement disallowing interpersonal comparisons of utilities can be expressed as the condition that the SWFL f be sensitive only to the preference orderings represented by individual utility functions, and not to the utility values.

Ordinal Non-Comparability (ONC): For all $U_N, U'_N \in D$, one has $f(U_N) = f(U'_N)$ whenever for all $i \in N$ and all $x, y \in X$,

$$U_i(x) \geq U_i(y) \Longleftrightarrow U'_i(x) \geq U'_i(y).$$

The other axioms we will consider here are the following:

Unrestricted domain (U): The domain D on which f is defined is equal to the whole Cartesian product set $\mathcal{U}^N(X)$.

Transitivity (T): For all $U_N \in D$, $f(U_N)$ is transitive.

Completeness (C): For all $U_N \in D$, $f(U_N)$ is complete.

Strong Pareto (SP): Given any $U_N \in D$, let $R = f(U_N)$, and let P denote the associated strict preference relation. Then, for any pair $x, y \in X$, it must be true that $x \; R \; y$ whenever $U_i(x) \geq U_i(y)$ for all $i \in N$; and that $x \; P \; y$ if, in addition, there is an i such that $U_i(x) > U_i(y)$.

Anonymity (A): For all $U_N, U'_N \in D$, one has $f(U_N) = f(U'_N)$ whenever U'_N is derived from U_N by permuting the individuals' utility functions.

[6]Later, in Section 7, we propound a decision-theoretic approach to ethics that is intended to meet this kind of criticism.

Independence of irrelevant utilities (IIU) (usually called "independence of irrelevant alternatives"):[7]

Let A be any non-empty subset of X. Given any two functions or binary relations Q and Q' defined on X, write $Q =_A Q'$ to indicate that the two coincide on the subset A. For all $U_N, U'_N \in D$, it is required that $f(U_N) =_A f(U'_N)$ whenever $U_i =_A U'_i$ for all $i \in N$.

In the current framework, Arrow's impossibility theorem implies that, when there are at least two individuals and three social states, there is no SWFL f which satisfies all seven of the above conditions. Actually, Arrow proved a stronger result, involving the weak version of the Pareto condition according to which $x \; P \; y$ whenever $U_i(x) > U_i(y)$ for all i, and also a weaker axiom than Anonymity. As for the latter, he only required the absence of a *dictator*—that is, of an individual $d \in N$ who, given any $x, y \in X$, has the power to ensure that $x \; P \; y$ whenever $U_d(x) > U_d(y)$.

Sen (1970a) has also shown that the exclusion of interpersonal comparisons of utilities could be formulated in a more subtle way without altering the validity of the theorem, by requiring instead that social preferences be invariant only to affine rescaling of utilities. That is, the theorem remains valid if (ONC) is replaced with the following logically weaker axiom:

Cardinal Non-Comparability (CNC) : For all $U_N, U'_N \in D$, one has $f(U_N) = f(U'_N)$ whenever for all $i \in N$, there are real constants α_i and β_i, with each $\beta_i > 0$, such that $U'_i \equiv \alpha_i + \beta_i U_i$.

Such an impossibility result makes it quite tempting to conclude that social choice without interpersonal comparisons is just a *non sequitur*, or at best yields a degenerate rule such as a dictatorship. But the precise formulation of the theorem made possible by the current mathematical framework shows that this conclusion is hasty. Because the formal translation of the sentence "social choice is impossible without interpersonal comparisons" is: "there is no SWFL satisfying ONC (or CNC)", and this formal translation is logically wrong. The correct theorem involves no less than six other conditions! One first has to show that these other conditions are absolutely necessary and unexceptional before one can conclude that social choice requires interpersonal comparisons.

In the next subsection we examine what kinds of social choice are made possible by relaxing some of these other six conditions. In this way, our brief survey will cast some doubt on the claim that social choice requires interpersonal comparisons of utility.

---

[7]The axiom is taken from d'Aspremont and Gevers (1977), but was first named this way by Hammond (1987a), as far as we can tell.

3.2 Possibilities

Restricted domain

For the time being, we revert to imposing the (ONC) condition on the SWFL. As one may easily guess, what really matters then in the Unrestricted Domain axiom (U) is that all logically possible profiles of individual orderings be included in the domain. In particular, the impossibility theorem remains valid if restrictions are imposed only on the particular profile of utility functions chosen to represent each member of the unrestricted domain of preference profiles.

On the other hand, if some restrictions can be imposed on the profile of individual orderings too, then interesting new possibilities do emerge. Of these, the most interesting was mentioned as early as in Arrow (1951), following Black (1948). In our framework with utility functions it can be described easily. Suppose the elements of X can be arranged along a line in a way that makes all individual utility functions in the domain unimodal—i.e., either increasing, or decreasing, or else increasing and then decreasing. Then the domain of preference profiles is said to be "single-peaked" (even though some preferences may not have any peak at all!). In this case, *simple majority rule* is defined so that, given $R = f(U_N)$, for any $x, y \in X$ one has

$$x \; R \; y \iff \#\{i \in N \mid U_i(x) > U_i(y)\} \geq \#\{i \in N \mid U_i(x) < U_i(y)\}.$$

It is easy to show that this rule, though it obviously violates the unrestricted domain condition, nevertheless satisfies all the other conditions of Arrow's theorem.

This "possibility" result, which actually holds under somewhat weaker domain restrictions, can be applied practically in some cases, such as political contests in which right–left conflicts polarize the population's preferences. It does not help much, though, in economic issues where individual interests are widely divergent, so that any conflict between them cannot be reduced to only one dimension. The problem of income distribution, for instance, is typically of this kind when $n \geq 3$. Indeed, there are $n - 1$ irreducible dimensions when total income is fixed, but different individuals' incomes can be varied independently otherwise. It is rather obvious that this problem cannot be solved with anything like simple majority rule.

Intransitive social preferences

When the Unrestricted Domain axiom (U) is retained, simple majority rule is still well defined, but the associated relation $f(U_N)$ is generally intransitive and, indeed, produces "majority cycles"—a result famously known as Condorcet's Paradox. A theorem due to McGarvey (1953) shows that violations of transitivity can be quite severe, as essentially any complete binary relation on X can be obtained by applying simple majority rule to some utility profile. Moreover, McKelvey (1976) has shown that, in a multidimensional problem such as choosing an income distribution, it can yield totally unacceptable social preferences. This is because a sequence of majority decisions can lead from any one distribution of income to almost any other. There is also a substantial literature which studies whether majority cycles are likely to occur when some probability distribution is given on the domain of preference profiles.[8]

When X contains only two elements, however, the Transitivity axiom (T) is trivially satisfied by any SWFL, and in this particular context simple majority rule satisfies all the axioms. It can even be defended as the only reasonable SWFL satisfying all the conditions of Arrow's theorem. Indeed, it was characterized as such by May (1952), under mild additional conditions.

Relaxing the strong Pareto axiom

Having considered (U) and (T), let us go on to review the possibilities allowed by retaining these two and dropping other conditions instead. The Strong Pareto axiom (SP) deserves some comment. First, notice that among the various conditions of the theorem, this is the only one which implies that social choice must depend on individual utilities. If (SP) is dropped, any fixed social ordering will satisfy all other conditions of the theorem. Thus, (SP) can be viewed as protecting against paternalism or perfectionism, but in a rather minimal way since it focuses on situations of unanimity and does not imply any respect for personal preferences regarding private matters.[9]

Axiom (SP) has been criticized from many viewpoints. One is this idea that privacy deserves more thorough protection, and Sen (1970b) has actually shown how easily this idea can conflict with the Pareto condition. Other critics have focused on the ethical limitations of individual preferences under conditions of imperfect information, uncertainty, or lack of autonomy. And yet other critics have argued that utility ignores other ethically relevant aspects of individuals' situations.

All these lines of criticism, however, can be interpreted as mere calls for a careful selection of the relevant utility measures that are used to determine the social choice. Once the utility functions U_i have been correctly constructed to measure each individual's good, it becomes hard to oppose the "principle of

---

[8] Among many references, see DeMeyer and Plott (1970), Gehrlein and Fishburn (1976), Berg and Lepelley (1994) and Gehrlein (1997).

[9] An exception occurs if the concept of individual good is defined to be independent of all other agents' private matters. See, for example, Hammond (1982, 1995), as well as Coughlin (1986), who shows that then Pareto efficiency requires respect for individuals' own good regarding private matters.

personal good" (Broome 1991; see also Section 7.3 in this chapter) according to which more of this good for some individuals and no less for others is a social improvement. Axiom (SP) simply reflects this principle. In brief, abandoning (SP) seems to offer little hope by itself of finding interesting new possibilities, although the next subsection will study some of this matter in more detail.

At any rate, a theorem due to Wilson (1972) implies that any SWFL satisfying all of the above conditions except (SP) must be either dictatorial (there is an individual who imposes his strict preferences), or anti-dictatorial (there is an individual whose strict preferences are always contradicted by social preferences), or imposed (totally independent of the population's preferences). This definitely shows that even if (SP) could be legitimately dropped, nothing of interest would be obtained in that way.

Relaxing anonymity

Similar unappealing conclusions emerge from relaxing the Anonymity axiom (A) instead. This is because, as mentioned above, Arrow's theorem can be reformulated to say that all the other conditions together imply that there must be a dictator. With the conditions at hand here (especially Strong Pareto), one can be a little more precise. If (A) is dropped, then the other conditions jointly imply that the SWFL must be a serial (or lexicographic) dictatorship. This means that all the individuals $i \in N$ are given some ranking $i_1, \ldots, i_n$ such that, for $k = 1, \ldots, n-1$, individual i_k's strict preferences are always imposed before those of i_{k+1}. Moreover, i_k hands over to i_{k+1} only when i_k is indifferent. In particular, i_1 must be a dictator. Since serial dictatorships are so obviously ethically unattractive, this line of enquiry does not deserve to be pursued any further.

Incomplete social preferences

If instead the Completeness requirement (C) is dropped, there is a more interesting possibility result or characterization due to Weymark (1984). All the other axioms jointly imply that the SWFL must be the Pareto Rule, defined as follows. Let $R = f(U_N)$ and let P denote the associated strict preference relation. Then, given any $x, y \in X$, one has:

$$\begin{aligned} x \; R \; y \quad &\Longleftrightarrow \quad \forall i \in N, \; U_i(x) \geq U_i(y); \\ x \; P \; y \quad &\Longleftrightarrow \quad \forall i \in N, \; U_i(x) \geq U_i(y) \text{ and } \exists i \in N, \; U_i(x) > U_i(y). \end{aligned}$$

In other words, under the conditions of Weymark's theorem, the social preferences cannot say any more than what the (SP) axiom already implies. There is no way to have finer preferences than those that the Pareto criterion itself determines.[10]

This result sheds light on some aspects of welfare economics, as summarized in the previous section, which presented what can be said on the basis of the Pareto criterion alone. Nonetheless, it would be presumptuous to consider that this result really explains the failure of the various historical attempts to construct extensions of the Pareto criterion through compensation tests, aggregate surplus measures, or similar contrivances. This is because these attempts did not seek to satisfy all the conditions of the theorem. In particular, they did not seek to obey the (IIU) axiom. This result also sheds light on the common view that going beyond the Pareto criterion requires interpersonal comparisons. This is because, once again, looking at the formal translation of this bold statement exposes its fragility. Indeed, one arrives at the claim "(SP), (C) and (ONC) are incompatible", which is an error in logic, as there are four other conditions involved. In particular, simple majority rule satisfies all these three conditions, and is a quite acceptable method of aggregation in some cases. But there are still other possibilities, to which we now turn.

Dependence on "irrelevant" utilities

The Independence of Irrelevant Utilities axiom (IIU) forces social preferences over a subset (e.g., a pair) of social states to depend only on utility levels on this subset, and not at all on preferences or utilities at other alternatives of X. If one drops (IIU), a host of new possibilities arise. The most famous of these is the Borda rule, which satisfies all the other axioms. It is defined as follows, for a finite set X. Given any profile U_N, construct the "Borda utility function" of each individual $i \in N$ by

$$B_i(x) := \#\{\, y \in X \mid U_i(x) > U_i(y)\}$$

for each $x \in X$. Note that $x \; R_i \; y$ if and only if $B_i(x) \geq B_i(y)$, so this really is a utility function that represents R_i. Then define the *Borda count* by

$$B(x) := \sum_{i \in N} B_i(x)$$

[10] If the Anonymity axiom (A) is also dropped, then Gibbard was the first to prove that there is a whole class of "oligarchic" social choice rules, as discussed by Sen (1970a, 1986) and Weymark (1984). Given any non-empty subset $K \subset N$ which is the "oligarchy", the least selective such rule is defined so that $x \; P \; y$ if and only if $U_i(x) > U_i(y)$ for all $i \in K$. Obviously, P is transitive. When $K = \{d\}$, this is a dictatorship. When $K = N$, this is the weak Pareto rule. Note that there are several ways of constructing the associated weak preference relation $R = f(U_N)$. One, which satisfies (T) but not (C), involves having $x \; R \; y$ if and only if $U_i(x) \geq U_i(y)$ for all $i \in K$. Another, which satisfies (C) but not (T), involves having $x \; R \; y$ if and only if $y \; P \; x$ is false.

for all $x \in X$. Finally, define $R = f(U_N)$ as the social ordering which satisfies $x \; R \; y$ if and only if $B(x) \geq B(y)$.

It is easy to imagine other "ranking" rules similar to this one. Indeed, there is a significant literature devoted to analysing the properties of preference aggregation (or voting) rules which violate the (IIU) axiom.[11] As explained in Section 5.5 below, although it belongs to a rather different framework, the Nash bargaining solution for n persons can also be regarded as an SWFL which violates (IIU) in an essential way.

Economic domains provide an even wider scope for violating (IIU). When X describes an economic problem of resource allocation, it may be quite natural to abandon (IIU). Indeed, there is an extensive literature on fairness criteria in allocation rules, most of which does not bother with (IIU).[12] This is because individuals' preferences over counterfactual allocations commonly play a role in determining equity requirements. For instance, requiring an allocation to be rejected if some individuals in it are worse off than at the fully egalitarian split of the available resources makes sense even when the fully egalitarian split is not among the considered alternatives. Similarly, the requirement that no individual strictly prefers another's consumption bundle (the so-called "no-envy" condition) might be sensible even if permuting bundles is not considered part of the agenda for social choice. Examples of interesting social criteria derived along these lines will be presented in Section 5.

As yet another example, the requirement that the allocation be efficient over all feasible allocations makes sense even if not all feasible allocations are considered for the social choice. This last example shows that even purely Paretian considerations may actually violate (IIU): When comparing x and y, we may want to know whether they are Pareto efficient overall, not just whether one Pareto dominates the other.

Discussion

The conclusion of this brief survey of possibility results without interpersonal comparisons of utilities is clearcut. There are three directions in which interesting possibilities can be found. First, majority rule is satisfactory in some restricted domains, or when transitivity is not an issue. Second, much welfare analysis can be based on the Pareto criterion alone, as described in the previous section. Third, the idea of relaxing (IIU) and allowing social preferences over two options to depend on individual utilities or preferences at other options has been very fruitfully exploited in several important parts of the literature:

---

[11] See in particular Young (1974, 1994), Young and Levenglick (1978).

[12] A recent survey is available in Moulin and Thomson (1996).

voting rules, fairness criteria, and bargaining solutions (and one might even add welfare economics, for some Paretian applications).

This analysis refutes the broad claim that social choice is impossible without interpersonal comparisons of *utilities*, but it does not refute the weaker claim that social choice is impossible without interpersonal comparisons of *something*. After all, even the choice of a dictator must rely on the comparison of something. If all individuals were indistinguishable the dictator's preferences could not be identified and obeyed. A dictatorial rule relies on knowledge (and comparison) of the individuals' labels: if the individual labelled 1 is the dictator, it is not because of 1's utility function, or 1's preferences, or any other characteristic (because 1 remains the dictator independently of the profile). Instead, it is entirely by virtue of 1's label. Similarly, simple majority rule compares weights in the voting process by allotting "one vote" to "one man". The Borda rule is just like the utilitarian criterion applied to contrived individual scores (the Borda utilities), which implies interpersonal comparisons of score differences (see the description below of the utilitarian criterion).

These examples show that, while the weaker claim that social choice is impossible without interpersonal comparisons of something may not be very profound, it does raise an interesting question: what does it mean to say that a social ordering relies on a particular kind of interpersonal comparison? We will address this question in the next section. Before then, some additional comments on welfarism and the independence axiom are worth making.

3.3 Welfarism

The Strong Pareto axiom (SP) implies in particular a significant Pareto indifference condition, saying that two options must be socially equivalent if all individuals are indifferent between them. Then, because $U_N(x) = U_N(y)$ implies that x and y must belong to the same indifference class, this condition can be equivalently formulated as follows:

Pareto Indifference (PI): For each $U_N \in D$, there is a binary relation $R^*_{U_N}$ over $\mathbb{R}^n$ such that, for any pair $x, y \in X$, letting $R = f(U_N)$, one has

$$x \; R \; y \Longleftrightarrow (U_1(x), ..., U_n(x)) \; R^*_{U_N} \; (U_1(y), ..., U_n(y)).$$

This condition can be described as "single-profile" welfarism. The expression "welfarism", which was given this meaning by Sen (1977), refers to the view that social choice should focus on individual utilities and nothing else (such as status, rights, resources, opportunities, etc.). In the framework of this section, however, the welfarism embodied in the Pareto Indifference condition should be much less contentious because utilities can refer to any relevant measure of the individual good.

Interestingly enough, it is possible to relate (IIU) to a different kind of welfarism by rewriting it in the following, equivalent, way:

Independence of irrelevant utilities (IIU): For any pair $(x, y) \in X \times X$, there is a binary relation $R^*_{(x,y)}$ over $\mathbb{R}^n$ such that, for all $U_N \in D$, letting $R = f(U_N)$, we have

$$x \ R \ y \Longleftrightarrow (U_1(x), ..., U_n(x)) \ R^*_{(x,y)} \ (U_1(y), ..., U_n(y)).$$

This formulation describes what might be called "agenda" welfarism. This means that the preference over every ordered pair of alternatives (x, y) (or agenda) is determined by an ordering $R^*_{(x,y)}$ over the corresponding pair of utility vectors, which applies independently of all other aspects of the profile of utility functions.

It is important to note that, in a particular sense, neither (PI) nor (IIU) prevents social preferences from taking other non-utility individual characteristics into account. Imagine for a moment that the available data also contained a function $\theta_i(x)$ describing other individual characteristics (possibly influenced by x). Then both (PI) and (IIU) would allow a generalized SWFL $f(U_N, \theta_N)$ which depends in addition on the profile θ_N of these other characteristics. That is, it would be possible to have $f(U_N, \theta_N) \neq f(U_N, \theta'_N)$ for some θ_N, θ'_N. In our framework, we have justified omitting θ_i by the assumption that each U_i already contains all the relevant information about individual i. Thus, the basic welfarism which consists in excluding non-utility characteristics from the analysis has been built in, since U_i is the only individual characteristic that is given here.

Nevertheless both (PI) and (IIU) do reinforce such basic welfarism. The focus of (PI) and (IIU), however, differs in what they add. Under (PI), no other consideration about the features of *alternatives* (including consequences described by $\theta_i(x)$, in the generalized framework just described) can supplement comparisons of utility vectors. For instance, the fact that one alternative is more equitable than another (or has better consequences described by $\theta_i(x)$) does not matter if they yield the same utilities. Under (IIU), on the other hand, no consideration concerning *utility profiles* can be added to comparisons of utility *vectors* for the pair. Even if two profiles give different perspectives to individual situations, this is deemed irrelevant if they attach the same utilities to the pair under consideration.

Combining (PI) with (IIU) yields a strong "multi-profile" form of welfarism, as described in the Welfarism Lemma due to d'Aspremont and Gevers (1977).[13] This states that there is a single binary relation R^* over $\mathbb{R}^n$ such that, for any pair $x, y \in X$, for all $U_N \in D$, letting $R = f(U_N)$, one has

$$x \ R \ y \Longleftrightarrow (U_1(x), ..., U_n(x)) \ R^* \ (U_1(y), ..., U_n(y)).$$

Under (PI) and (IIU), therefore, social choice theory is reduced to the quest for just one satisfactory ordering R^* over all utility vectors—obviously an enormous simplification of the original problem.

Now, if one accepts both (PI) and (IIU), and therefore the strong version of welfarism just defined, it is easy to understand why it is difficult to combine these two axioms with (ONC). In fact, (ONC) implies the following very demanding property for R^*: for any $u, v, u', v' \in \mathbb{R}^n$ such that, for all $i \in N$, $u_i - v_i$ has the same sign as $u'_i - v'_i$, it must be true that

$$u \ R^* \ v \Longleftrightarrow u' \ R^* \ v'.$$

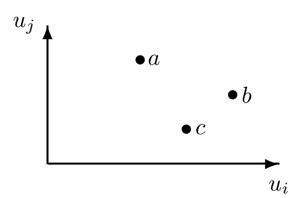

Figure 3.1 Options in the utility possibility set

In the two-agent case one can use a figure to illustrate how stringent this property is.[14] Let P^* and I^* denote the strict preference and indifference relations associated with R^*. In Fig. 3.1, suppose that $a \ I^* \ b$. For (ONC) to be satisfied, a and b must be ranked in the same way as a and c. That is, because $a \ I^* \ b$, it must be true that $a \ I^* \ c$. But then, if R^* is transitive, it follows that $b \ I^* \ c$, which contradicts (SP).

Suppose instead that $a \ P^* \ b$. Then under (SP), for any pair u, v such that $u_j > v_j$, one has $u \ P^* \ v$, which means that individual j is a dictator. Under

[13] See Bossert and Weymark (ch. 20 of this volume) or Mongin and d'Aspremont (1998) for a detailed statement and proof. See also Bordes, Hammond and Le Breton (1997) for an extension to economic domains.

[14] A very useful graphical analysis is provided by Blackorby, Donaldson and Weymark (1984).

the opposite assumption $b\ P^*\ a$, individual i would be the dictator. The only reasonable alternative to dictatorship is to declare a and b non-comparable, along with all similar pairs. This implies that R^* boils down to the Pareto rule.

This simple reasoning provides the rationale for the widespread view that any social welfare criterion going beyond Paretianism must rely on interpersonal comparisons. Under welfarism, and especially because (IIU) rules out any information about individuals' utility or preference "types" (other than their utility levels for the alternatives under consideration), there is indeed no other way out.

In the literature, most critics of welfarism such as Sen (1977, 1979) have focused on its Pareto-Indifference part.[15] A variety of similar examples have been provided by Sen (1979, 1987), in which two individuals labelled A and B have constant utilities 7 and 8, respectively, in three different social states. In one such state a mild income disparity explains the small utility gap. In a second social state A is much poorer than B, but is allowed to torture B. A third social state has A poorer than B and on foot, but rejoicing at the sight of B falling off his bicycle. Sen argues that the presence of torture or malevolence in the second and third social states (combined with gross inequality) makes it plausible to rank them below the first state, even though all three have an identical vector of utility levels. Sen also refers to other non-welfarist equity requirements such as avoiding exploitation, "equal pay for equal work", ensuring respect for equal rights, etc. All such considerations are obviously intended as criticisms of the Pareto-Indifference condition.

As already mentioned above, such criticisms of welfarism can be interpreted in several different ways. One suggestion is that (PI) and therefore (SP) should be abandoned, thereby allowing some partially or even totally imposed SWFL, largely independent of individual utilities. The above standard framework of social choice seems ill-adapted to the study of imposed SWFLs because of the absence of non-utility information. It would be possible, however, to enrich the information about individuals by replacing each U_i with a multidimensional function F_i of relevant personal characteristics, constructed so that the vector $F_i(x)$ would describe all relevant functionings for individual i in any state $x \in X$. The social choice problem would then become that of defining a social ranking R on X for any profile $\langle F_i \rangle_{i \in N}$ of individual functionings. In fact, from a mathematical point of view, the only difference between this new model and the previous one is that F_i is multidimensional whereas U_i was unidimensional.

[15] Sen (1977) attacks "neutrality", but his references to non-utility features of alternatives such as liberty and exploitation clearly point to the Pareto condition. This interpretation is also defended by Bossert and Weymark (ch. 20, this volume).

Now, if the non-welfarist measure of individual good is monistic, then the original framework can be applied just as well, by simply reinterpreting U_i in a different way, incorporating the "objective" features of well-being such as not being tortured, being fairly treated, etc. In particular, the capabilities and functionings approach put forth by Sen (which is defined and examined in more detail below) seems pretty much of the monistic kind, although Sen often mentions the difficulty of aggregating the various dimensions of functionings into one index.

With this second interpretation of non-welfarist approaches as merely involving a different concept of each individual's U_i, the point of abandoning a purely subjective notion of utility is not to escape Arrow's theorem by dropping (PI) or (SP). Instead it involves the recognition that some objective aspects of different individuals' personal situations may be more easily measured and compared interpersonally than can mental states or subjective satisfaction. Referring specifically to Arrow's theorem, this suggests dropping (ONC), which we will do in the next section.

3.4 Independence of Irrelevant Alternatives

Critics of welfarism have focused on (PI) rather than (IIU). Actually, most of the literature has either defended (IIU) or similar conditions, or else appears to have taken it for granted.

Let us first mention that, with a rich enough domain of utility function profiles, combining (IIU) and (ONC) is *logically equivalent* to the following axiom, which was proposed in one piece by Arrow in his framework with individual preferences:

Independence of irrelevant alternatives (IIA): Let A be any nonempty subset of X. For all $U_N, U'_N \in D$, letting R_i (resp. R'_i) denote the preference ordering represented by U_i (resp. U'_i), it must be true that $f(U_N) =_A f(U'_N)$ whenever $R_i =_A R'_i$ for all $i \in N$.

In particular, this axiom means that the social ranking of any pair of social states must depend only on individual preferences over that pair. It may therefore be appropriate to repeat the suggestion in Hammond (1991b) that this axiom could better be called Independence of Irrelevant Personal Comparisons. In particular, this suggestion recognizes that the axiom rules out any reference to how other alternatives fare even in purely *intra*personal comparisons.

The literature hesitates somewhat over whether Arrow's Independence axiom should be read as forbidding interpersonal comparisons, or whether this prohibition is already implied by Arrow's ordinal framework involving preferences and not utilities. The above formulation of (IIA), in the current framework with utilities, does not raise any problem and clearly contains both the restrictions embodied in (IIU) and (ONC).

It is interesting to note that in his early paper Arrow (1950, p. 342) considered that both parts of (IIA) had to do with interpersonal comparisons of utility: "These conditions taken together serve to exclude interpersonal comparison of social utility either by some form of direct measurement or by comparison with other alternative social states."[16] From this broad construal of interpersonal comparisons he could derive the bold interpretation of his theorem that would later come to be widely accepted: "If we exclude the possibility of interpersonal comparisons of utility, then the only methods of passing from individual tastes to social preferences which will be satisfactory and which will be defined for a wide range of sets of individual orderings are either imposed or dictatorial." (ibid.)

But the notion of interpersonal comparisons of utility has later become narrower, so that (ONC) alone is now considered to be enough to preclude any kind of interpersonal utility comparison. Indeed, this seems more rigorous, although it remains interesting to examine the various kinds of non-utility comparisons that are forbidden by the (IIU) part of (IIA). This will be done in the following sections.

Thus, Arrow's bold interpretation of his theorem has remained largely untouched despite the fact that a narrower notion of interpersonal comparison has supplanted Arrow's own. This is due to the wide acceptance of the (IIU) part of (IIA) for reasons other than to avoid interpersonal utility comparisons. Kemp and Ng (1987), for instance, even state that "to understand independence [the (IIU) part of (IIA)] is to accept it" (p. 226).

Hammond (1977a) has provided an argument in favour of (IIU) on the basis of what he later called "consequentialism". Following the ideas expounded in Chapters 5 and 6 for the normative theory of individual behaviour, the consequentialist approach to social evaluation requires it to be based exclusively on final consequences—whose description may, however, include a great deal of detail concerning the processes leading to each final consequence. Indeed, when faced with a decision tree, consequentialism requires society's decisions to have consequences that depend on the feasible set—defined as the range of all possible consequences which could result from decisions in the tree. Society's decisions should have consequences which are otherwise independent of the structure of the decision tree. In addition, it is argued that this consequentialist hypothesis should apply equally to subtrees—specifically, it requires that, in any subtree of a decision tree, society should neglect the rest of the tree structure and focus only on the consequences obtained in the subtree under

[16] His reference to "social utility" rather than to "personal utility" is fully in accord with our own interpretation of U_i as the social measure of individual good.

consideration. How this must exclude social rules which violate (IIA) can be illustrated using the Borda rule as a prominent example.

X	a	b	c	d	e
B_i	4	3	2	1	0
B_j	2	1	0	4	3
B	6	4	2	5	3

$X(n_1)$	a	d	e
B'_i	2	1	0
B'_j	0	2	1
B'	2	3	1

Table 3.1 The Borda counts in the tree T and in the subtree $T(n_1)$

Suppose that a, b, c, d, e are five different social states in X, and that $N = \{ i, j \}$ consists of two individuals. Suppose too that the individuals' preferences are specified by

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

Then the Borda utility functions and Borda counts are given in the left-hand part of Table 3.1. Thus, a is the optimal choice from $\{ a, b, c, d, e \}$. In the decision tree T illustrated in Figure 3.2, it is optimal to move first from n_0 to n_1.

However, suppose that the Borda rule is applied once again to the subtree $T(n_1)$ after reaching node n_1. Now b and c are no longer relevant alternatives. The new feasible set is $X(n_1) = \{a, d, e\}$. The Borda utility functions and Borda counts become revised, as indicated in the right-hand part of Table 3.1. So now, according to the Borda rule, the optimal choice has become d rather than a. The final outcome of applying the Borda rule at both nodes of the decision tree is d. Yet, if the decision tree only had one decision node, forcing an immediate choice of one social state from the set $\{ a, b, c, d, e \}$, the result would be a. In brief, consequentialism is violated because the decision procedure has consequences that depend on the tree structure.

On the other hand, this reasoning has assumed that the description of consequences is coarse enough to deprive the decision-maker of any information about the individuals' characteristics—including their utility functions, and especially the values of those functions in other social states. Hammond (1987a, b) notices that enriching the description of each social state in such a way might be relevant, and would indeed entail a violation of (IIA). Similar doubts about what alternatives are relevant, or about how they should be described, had already

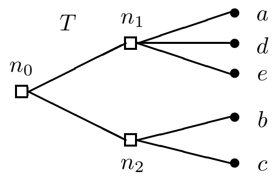

Figure 3.2 A decision tree T illustrating the Borda rule

been voiced by Bergson (1954) and Strasnick (1977). Samuelson (1987) has provided a vivid description of such unease about (IIA): "Once we agree that a [social] choice *legitimately* can depend on what "types" our persons are, and agree that defining people's types can depend on *more* than ... binary choosings, then I must agree with Bergson's contention that, operationally we are explicitly (and reasonably) deciding to *violate* the axiom of Independence of Irrelevant Alternatives. Third states of the world do seem to force themselves legitimately into our binary choices. Nor is this merely a small point connected with the details of logical implication. Most ethical systems purport to define who is the deserving one by how the contemplated individuals react to a vast panoply of possible situations." (p. 170)

If one follows Samuelson's view of ethical relevance, then the most promising escape from Arrow's "impossibility" theorem is indeed to abandon (IIU) while retaining the other axioms. If (ONC) can be abandoned as well because individual good is measurable in an interpersonally comparable way, so much the better. But this informational demand should not be regarded as a *sine qua non*.

4 Social Choice with Interpersonal Comparisons

4.1 Interpersonal Comparisons of Utility

What form of rational social decision-making is possible *with* interpersonal comparisons? This question was a major preoccupation of the search during the 1970s for satisfactory escapes from Arrow's theorem, which focused on comparisons of *utility*. This was why, following a preliminary idea due to Suppes (1966), later Sen (1970a) formulated the general concept of a *social welfare functional*, whose domain consists of profiles of utility functions rather than of preference orderings. This is the framework already used in the previous section, of course. The question at hand now is what happens when (ONC) or (CNC) is dropped or weakened so as to allow interpersonal comparisons of utility.

In order to illustrate some of the additional possibilities, suppose that the two utility function profiles U_N and U'_N are deemed equivalent if and only if there exist real constants α and β, with $\beta > 0$, such that $U'_i(x) = \alpha + \beta U_i(x)$ for all $i \in N$ and all $x \in X$. Note that such transformations preserve interpersonal comparisons of *utility levels* of the form $U_i(x) > U_j(y)$, as well as comparisons of *utility differences* of the form $U_i(x) - U_i(y) > U_j(y) - U_j(x)$. That is:

$$U_i(x) > U_j(y) \iff U'_i(x) > U'_j(y)$$

and

$$U_i(x) - U_i(y) > U_j(y) - U_j(x) \iff U'_i(x) - U'_i(y) > U'_j(y) - U'_j(x).$$

Now let $v_k(x)$ ($k = 1$ to n) denote the kth smallest individual utility level in each social state $x \in X$ — i.e., $v_k(x)$ is defined as the unique real number r satisfying

$$\#\{\, i \in N \mid U_i(x) < r \,\} < k \leq \#\{\, i \in N \mid U_i(x) \leq r \,\}$$

Then a whole class of SWFLs which are invariant under the transformations specified above are the "rank-dependent utilitarian" rules given by

$$x \; R \; y \iff \sum_{i=1}^{n} r_k \, v_k(x) \geq \sum_{i=1}^{n} r_k \, v_k(y)$$

for any collection r_k ($k = 1$ to n) of real constants. These constants should all be positive if the SWFL is to satisfy the Strong Pareto axiom.

One special case of some importance arises when $r_1 = 1$ and $r_k = 0$ for all $k > 1$. This gives the "Rawlsian" maximin rule, with[17]

$$x \; R \; y \iff \min_i \{U_i(x)\} \geq \min_i \{U_i(y)\}$$

A second special case occurs when $r_k = 1$ for all k. This gives the "utilitarian" SWFL, with

$$x \; R \; y \iff \sum_{i=1}^{n} U_i(x) \geq \sum_{i=1}^{n} U_i(y)$$

But there are many other possibilities, of course. There are also different possible degrees of interpersonal comparability. For a discussion of the various

[17] Recall that in Rawls' theory, the functions U_i to which the maximin (or leximin) criterion is applied are indices of primary goods, whereas Kolm (1972) was an early proponent of applying the maximin criterion to measures of subjective well-being instead.

possibilities, see Roberts (1980b), Blackorby, Donaldson and Weymark (1984, 1990), d'Aspremont (1985), and Bossert and Weymark (ch. 20 of this volume). Certainly, explicitly introducing interpersonal comparisons of utility allows the unpalatable conclusion of Arrow's theorem to be avoided.

Maximin does not satisfy the Strong Pareto condition. However, it can be made to satisfy (SP) by extending it lexicographically to the *leximin* SWFL, which is specified by

$$x \; P \; y \Longleftrightarrow \exists r \in \{1, 2, \ldots, n\} : \begin{array}{rcll} v_k(x) & = & v_k(y) & (k = 1, 2, \ldots, r-1) \\ \text{and} \quad v_r(x) & > & v_r(y). & \end{array}$$

Obviously, this definition implies that

$$x \; I \; y \Longleftrightarrow v_k(x) = v_k(y) \quad (k = 1, 2, \ldots, n).$$

In particular, $x \; I \; y$ if and only if the two utility vectors $\langle U_i(x) \rangle_{i \in N}$ and $\langle U_i(y) \rangle_{i \in N}$ are equal after a suitable permutation of their components.

The literature of the 1970s focused on the question of what kinds of SWFL are made possible with various kinds of interpersonal comparison. This approach is surveyed in Bossert and Weymark (ch. 20 of this volume). Since our purpose is to understand better the notion of interpersonal comparisons, it will be more fruitful here to focus on the slightly different question of what kinds of interpersonal comparison are involved in any given social criterion. We will start by concentrating on two particularly important examples—namely, the maximin and utilitarian SWFLs that were presented above.

4.2 Maximin and Comparisons of Utility Levels

The maximin SWFL evidently requires interpersonal comparisons of utility levels. The way this is captured formally is by looking at the transformations of utility profiles that would leave the criterion unaffected in all circumstances. Take any profile U_N, and consider any other profile U'_N for which there is a strictly increasing transformation $\varphi : \mathbb{R} \to \mathbb{R}$ such that U'_i is the composite function $\varphi \circ U_i$, for all $i \in N$. Then one can safely declare that the maximin SWFL will yield the same ordering over X for the two profiles, since for any pair $x, y \in X$,

$$\min_i \{U_i(x)\} \geq \min_i \{U_i(y)\} \Longleftrightarrow \min_i \{U'_i(x)\} \geq \min_i \{U'_i(y)\}.$$

Now suppose that U'_N and U_N are not related by any such transformation φ. For instance, suppose there exist two individuals $i, j \in N$ and two social states $x, y \in X$ such that $U_i(x) > U_j(y)$ whereas $U'_i(x) < U'_j(y)$. It is possible that the difference between U'_N and U_N may not affect the social preference over X because neither i nor j has the minimum level of utility in any social state. But the possibility of change is clearly there, and would be actualized if i and j indeed had the lowest utilities in the two profiles. In other words, as soon as different transformations φ_i are applied to different individuals' utility functions, one can find a profile U_N such that the SWFL generates a different social ordering when applied to the transformed profile U'_N.

Following Roberts (1980b)—see also Sen (1970a)—this example suggests the following general definition of the information required by a social criterion. Let $\mathcal{I}$ denote the set of strictly increasing functions $\varphi : \mathbb{R} \to \mathbb{R}$, and $\mathcal{I}_N$ the set of profiles $\langle \varphi_i \rangle_{i \in N}$ satisfying $\varphi_i \in \mathcal{I}$ for all $i \in N$. Say that a profile of functions φ_N is an *invariance transformation* for the SWFL f if, for all $U_N \in D$, one has

$$f(U_N) = f(\langle \varphi_i \circ U_i \rangle_{i \in N}).$$

The above argument has shown that φ_N is an invariance transformation for the maximin SWFL if and only if $\varphi_i = \varphi_j$ for all $i, j \in N$. And, more importantly, *all such transformations preserve comparisons of utility levels*. Similar reasoning applies to the leximin SWFL. Notice that such transformations preserve not only comparisons between persons within states, such as $U_i(x) > U_j(x)$, but also between pairs consisting of both persons *and* states, such as $U_i(x) > U_j(y)$. Both the maximin and leximin criteria do indeed require such comparisons.

Generalizing from this example, one can say that *an SWFL relies on the particular kind of interpersonal comparison of utilities which is preserved by the class of all invariance transformations.*

This is the approach that motivated the various weakenings of (ONC) formulated in the 1970s. For the maximin and leximin SWFLs, the following axiom was formulated, which makes clear the related set of invariance transformations:

Ordinal Level-Comparability (OLC): For all $U_N, U'_N \in D$, one has $f(U_N) = f(U'_N)$ whenever there exists $\varphi \in \mathcal{I}$, independent of i, such that $U'_i = \varphi \circ U_i$ for all $i \in N$.

An alternative formulation of exactly the same informational requirements can be made in terms of an *interpersonal ordering* $\tilde{R}$ on the Cartesian product space $X \times N$ whose members are pairs (x, i) consisting of a social state $x \in X$ combined with an individual $i \in N$. Indeed, given any profile U_N, let $\tilde{R}$ be defined by

$$(x, i)\ \tilde{R}\ (y, j) \iff U_i(x) \geq U_j(y).$$

A preference statement such as $(x, i)\ \tilde{R}\ (y, j)$ should be interpreted as indicating that it is no worse for society to have individual i be in social state x than it is to have individual j be in social state y. The interpersonal ordering is similar in spirit to the notion of "extended sympathy" discussed by Arrow (1963)—see also Arrow (1977). Two other early discussions of such level comparisons occur in Suppes (1966) and Sen (1970a). A recent synthesis can be found in Suzumura (1996).[18]

Instead of Arrow's (IIA), the maximin and leximin SWFLs both satisfy a less demanding condition. As suggested by Hammond (1991b), this may be called *independence of irrelevant interpersonal comparisons* (or IIIC). The condition requires that, if $\emptyset \neq A \subset X$ and $\tilde{R} =_{A\times N} \tilde{R}'$, where $\tilde{R}$ and $\tilde{R}'$ are the orderings on $X \times N$ derived from U_N and U'_N respectively, then $f(U_N) =_A f(U'_N)$. It is a straightforward exercise to show that (IIIC) is logically equivalent to the combination of (IIU) and (OLC).

Coming back to the traditional question of what SWFLs are allowed by (OLC) or (IIIC), apart from leximin, many other SWFLs also satisfy conditions (U), (T), (C), (SP), (A) and (IIIC). One such SWFL, for example, is the "leximax" rule defined by

$$x \; P \; y \iff \exists r \in \{1, 2, \ldots, n\} : \quad \begin{array}{rcl} v_k(x) & = & v_k(y) \quad (k = r+1, \ldots, n) \\ \text{and} \quad v_r(x) & > & v_r(y). \end{array}$$

As shown by Roberts (1980a), all the other possible rules satisfying these six conditions involve a lexicographic hierarchy of "dictatorial positions". Of all these SWFLs, only leximin satisfies the additional equity axiom formulated in Hammond (1976), following a suggestion of Sen (1973), whose effect is to give priority to the worse-off person in any "two-person situation" — see also Hammond (1979).

4.3 Utilitarianism and Comparisons of Utility Difference Ratios

What kind of interpersonal comparison does the utilitarian SWFL rely upon? A profile of functions φ_N is an invariance transformation for this SWFL if and only if there exist real constants $\alpha_1, \ldots, \alpha_n$ and β, with $\beta > 0$, such that $\varphi_i(u) = \alpha_i + \beta u$ for all $u \in \mathbb{R}$. Indeed, if f is the utilitarian SWFL, then any other kind of transformation allows one to construct a utility profile U_N such that $f(U_N)$ does not coincide with $f(\langle \varphi_i \circ U_i \rangle_{i\in N})$.

Now, such invariance transformations make it possible to compare utility differences both intra- and interpersonally. Indeed, for any $i, j \in N$ and any

[18] In these works, as well as in Hammond (1976, 1979, 1991b), the primitive input is the ordering $\tilde{R}$ rather than the utility profile U_N. We have reversed the two here in order to retain a unified framework, and also in order to show the formal equivalence between the two approaches. See the concluding section for a comment on this point.

$x, y, x', y' \in X$, the comparisons

$$\begin{aligned} U_i(x) - U_i(y) &\geq U_i(x') - U_i(y') \\ U_i(x) - U_i(y) &\geq U_j(x) - U_j(y) \end{aligned}$$

are obviously unaffected by such transformations. But Bossert (1991) has pointed out that, for $n \geq 3$, the utilitarian criterion requires more information than just comparisons of utility differences, because two profiles U_N and U'_N which merely preserve comparisons of utility differences may not lead to the same ranking over X as the utilitarian SWFL. Details are provided in Bossert and Weymark (ch. 20 of this volume, sections 5 and 7). A simplified version of their example has $U_1(x) = U_2(x) = 1$, $U_1(y) = U_2(y) = 2$, and $U_3(x) = 3$, $U_3(y) = 1$. Then there exists a new utility function profile V which is identical to U, except that $V_3(x) = 4$. All utility difference comparisons are the same, but the utilitarian criterion is altered.

In fact, the key property characterizing the family of all invariance transformations is that they preserve all ratios of utility differences—i.e., all expressions of the form

$$\frac{U_i(x) - U_i(y)}{U_j(z) - U_j(w)}.$$

As a logical consequence, these invariance transformations preserve interpersonal comparisons of several utility features—most notably utility differences and their ratios, but also more exotic features such as the square roots or logarithms of the absolute values of utility differences and their ratios, etc. Having characterized the family of invariance transformations, the following axiom captures the informational requirements of utilitarianism:

Cardinal Unit-Comparability (CUC): For all $U_N, U'_N \in D$, one has $f(U_N) = f(U'_N)$ whenever there are constants $\alpha_1, \ldots, \alpha_n, \beta \in \mathbb{R}$ with $\beta > 0$ such that $U'_i = \alpha_i + \beta U_i$ for all $i \in N$.

Interestingly, these informational demands can be related to problems of social choice under uncertainty, in the extended sympathy context. Assume that social preferences must no longer rank only riskless alternatives in X, but also risky prospects or lotteries with outcomes in X. Chapter 5 discusses axioms that are sufficient to imply that behaviour in risky decision trees should maximize the expected value of a von Neumann–Morgenstern utility function. Such axioms seem equally valid (or invalid) for normative behaviour generally, regardless of whether the focus is on individual or social choice. Indeed, following the work of John Harsanyi in particular, one might well argue that higher normative standards should apply to social than to individual decision-making.

Following the notation of Chapter 5, let $\Delta(X)$ denote the set of simple probability distributions on the set X of social states. That is, each member $\lambda \in \Delta(X)$ is a mapping $\lambda : X \to [0,1]$ for which there is a finite *support* $F \subset X$ such that $\lambda(x) > 0$ if and only if $x \in F$, and also $\sum_{x \in F} \lambda(x) = 1$. Given any $\lambda \in \Delta(X)$ and any real-valued function v on X, denote the *expected value* of v w.r.t. λ by

$$\mathbb{E}_\lambda v(x) := \sum_{x \in F} \lambda(x)\, v(x).$$

Assume now that, in view of standard axioms such as those discussed in Chapter 5, there is a *von Neumann-Morgenstern* (or NM) *Bergson social welfare function* $w : X \to \mathbb{R}$ whose expected value *represents* the social ordering R on $\Delta(X)$ in the sense that, whenever $\lambda, \mu \in \Delta(X)$, then

$$\lambda \; R \; \mu \iff \mathbb{E}_\lambda w(x) \geq \mathbb{E}_\mu w(x).$$

Also, given any profile U_N, define the interpersonal ordering $\tilde{R}$ on $\Delta(X \times N)$ so that, for any $\tilde{\lambda}, \tilde{\mu} \in \Delta(X \times N)$, one has

$$\tilde{\lambda} \; \tilde{R} \; \tilde{\mu} \iff \mathbb{E}_{\tilde{\lambda}} U_i(x) \geq \mathbb{E}_{\tilde{\mu}} U_i(x).$$

Now assume that the Pareto Indifference condition (PI) is extended to probability distributions so as to require that, whenever $\lambda, \mu \in \Delta(X)$ satisfy $\mathbb{E}_\lambda U_i(x) = \mathbb{E}_\mu U_i(x)$ for all $i \in N$, then $\mathbb{E}_\lambda w(x) = \mathbb{E}_\mu w(x)$. Under this assumption and some additional domain conditions, Harsanyi (1955) showed that there must exist constant "welfare weights" ω_i ($i \in N$) such that $w(x) \equiv \sum_{i \in N} \omega_i \, U_i(x)$ on X. That is, one must have a weighted utilitarian Bergson social welfare function. Of course, if condition (SP) is supplemented by the extended (PI) rather than replaced by it, then the welfare weights ω_i must be positive for all $i \in N$. For other proofs showing that Harsanyi's result is valid even without additional domain conditions, see Border (1985), Coulhon and Mongin (1989), Broome (1990), and also Hammond (1992). A similar result also appears later in Section 7 of this chapter.

In this new framework of risky social choice, it is hardly surprising that Arrow's condition (IIA) still forces interpersonal comparisons to be ignored. Then Arrow's impossibility theorem implies that there must be a dictator $d \in N$ such that $\omega_d > 0$ and $\omega_i = 0$ for all $i \in N \setminus \{d\}$. What initially may be surprising, however, is that in the present framework involving the social choice of risky consequences or consequence lotteries in $\Delta(X)$, condition (IIIC) has exactly the same strong and unacceptable implication, provided the domain of possible utility functions is sufficiently rich. A formal result can be found in Hammond (1991b, Section 9). The basic explanation is that (IIIC) requires the social ordering R restricted to any finite set $A \subset X$ to remain invariant under any non-linear strictly increasing transformation of the function U_i. For this to be true when R is represented by $\sum_{i \in N} \omega_i U_i$ on the set A, generally the sum must collapse to the single term $\omega_d U_d$ for some $d \in N$.

The obvious remedy is to weaken the independence condition still further. The new condition, called *independence of irrelevant interpersonal comparisons of mixtures* (or IIICM), requires that, if $\emptyset \neq A \subset X$ and the two interpersonal orderings $\tilde{R}$ and $\tilde{R}'$ derived from two profiles U_N and U'_N satisfy $\tilde{R} =_{\Delta(A \times N)} \tilde{R}'$, then $f(U_N) =_A f(U'_N)$.

For any non-empty $A \subset X$ and fixed ordering $\tilde{R}$ on $\Delta(A \times N)$, the fact that $\tilde{R}$ is represented by $\mathbb{E}\, U_i(x)$ implies that $U_i(x)$, viewed as a function of (i, x), is determined uniquely on the set $A \times N$ up to a positive affine transformation—as discussed in Chapter 5, for instance. This obviously implies that the function which maps each $x \in A$ to $\sum_{i \in N} \omega_i U_i(x)$ is determined uniquely up to a positive affine transformation—in particular, the social ordering R is determined uniquely on the set A. Hence, unlike (IIIC), condition (IIICM) is weak enough to be satisfied when R is represented by $\sum_{i \in N} \omega_i U_i(x)$ with $\omega_i \neq 0$ for at least two different individuals $i \in N$. There is no need for a dictatorship or any other restriction on the constants ω_i $(i \in N)$, except the obvious requirement that they should all be positive if the Strong Pareto condition holds.[19] In particular, utilitarianism—whether weighted or unweighted—satisfies independence condition (IIICM). It even satisfies the formally stronger condition requiring that $R =_{\Delta(A)} R'$ whenever $\tilde{R} =_{\Delta(A \times N)} \tilde{R}'$.

Now, it is worthwhile noting that (IIICM) is logically equivalent to the combination of (IIU) and the following axiom, which is weaker than *both* (CUC) and (OLC).[20]

Cardinal Full Comparability (CFC): For all $U_N, U'_N \in D$, one has $f(U_N) = f(U'_N)$ whenever there are constants $\alpha, \beta \in \mathbb{R}$ with $\beta > 0$ such that $U'_i = \alpha + \beta U_i$ for all $i \in N$.

The only delicate part of showing this equivalence is in the implication from (IIU) and (CFC) to (IIICM).

Take any $A \subset X$, and assume that $\tilde{R} =_{\Delta(A \times N)} \tilde{R}'$. It follows that the corresponding von Neumann–Morgenstern utility functions U and U', defined by $U(i, x) = U_i(x)$ and $U'(i, x) = U'_i(x)$, must be cardinally equivalent on the domain $A \times N$. So there exist constants $\alpha, \beta \in \mathbb{R}$, with $\beta > 0$, such that $U' =_{A \times N} \alpha + \beta U$. In particular, $U'_i =_A \alpha + \beta U_i$ for all $i \in N$. Because of (IIU) and (CFC), it follows that $f(U_N) =_A f(U'_N)$.

[19] See Weymark (1991, 1993, 1995) and De Meyer and Mongin (1995) for discussion of this and other similar sign restrictions on the welfare weights.

[20] Notice that (IIICM), just like (IIU) and (CFC), puts restrictions on social preferences over riskless alternatives only.

Whereas Harsanyi's axiomatic derivation of utilitarianism relies on an extended version of Pareto-Indifference, there are other interesting characterizations based on the axioms (SP), (A), (IIU) and (CUC) or some weakenings of them — see d'Aspremont and Gevers (1977), Roberts (1980b), d'Aspremont (1985), Bossert and Weymark (ch. 20 of this volume).

4.4 Interpersonal Comparisons of What?

As is perhaps already clear from the previous subsection, the economic literature has often been somewhat imprecise about what kind of interpersonal comparisons are really needed to construct particular social preferences. This is especially the case when comparisons of objects other than utility are involved.

Here is an illustration of the difficulty. The utilitarian criterion is based on the sum of individual utilities, and is widely considered to involve interpersonal comparisons of utility—in particular, of utility differences and their ratios. But, as a first example, consider Borda utilities computed from individual preferences over X. The utilitarian criterion applied to such utilities is the Borda SWFL, which does satisfy (ONC) in the relevant framework, and therefore does not need any interpersonal comparisons of utility. The most reasonable conclusion in this case is probably that the Borda SWFL does not involve comparisons of "utility", but does compare (differences of) *Borda* utilities.

As a second example, assume that the cardinally comparable utilities included in the utilitarian sum are supplied by an ethical observer who carefully forms an opinion about individual preference intensities. Then it seems only reasonable to say that this approach does rely on interpersonal comparisons of utility.

But this second case starts to look more complicated if one learns that the preference intensities have actually been evaluated by the ethical observer on the basis of suitably normalized NM utility functions defined on X (for instance, with utility levels 0 and 1 assigned to the worst and the best alternatives, respectively). Therefore, just as with Borda utilities, the cardinally comparable utilities can actually be given an interpretation in terms of preferences. Can we then maintain that this approach really does rely on interpersonal comparisons of utility?

In order to adjudicate such borderline cases, we propose the following simple method. First, find individual indices of well-being U_i and an SWFL f such that social preferences are well described by the social ranking $f(U_N)$. Then declare that these social preferences rely on interpersonal comparisons of those *particular* indices U_i if f does so according to the concept of invariance transformation described in sections 4.2 and 4.3 above.

This simple method may provide multiple, but always compatible, answers, because there may be several ways to identify indices U_i and a SWFL f. In the example of the Borda rule, this method reaches the appropriate conclusion that there is no interpersonal comparison of ordinary utilities, which is indeed the case when the U_i are ordinary utilities and f is the Borda rule. But the method also demonstrates that there is some comparison of Borda utilities, which is indeed the case when the U_i are Borda utilities and f is the utilitarian SWFL. In the second example above, on the other hand, this method will simply say that there is some comparison of utilities, because the U_i are unambiguously provided from outside the model by the ethical observer, and f is the utilitarian criterion. In the last version of this example, however, one can add that the relevant NM utilities are derived from preferences over a broader set than X.

The simple method we propose here applies to most cases, but will have to be extended in order to accommodate some more exotic kinds of comparison. For instance, consider the simple majority rule restricted to adult citizens only. What kind of comparison does it involve? It seems to require the ability to compare dates of birth and citizenship status to threshold or benchmark values (implying indirect interpersonal comparisons of such variables), and also, in order to apply the rule of "one (wo)man, one vote" (but zero votes for aliens and children), the ability to make common ratio-scale comparisons of voting weights—that is, comparisons which are not altered when all weights are multiplied by a common positive constant. In a sense, this latter kind of comparison can be cast into the above mould by declaring that U_i measures i's voting weight, and that social preferences want all positive weights to be equal. But that is not a full representation of the social preference relation, which also compares social states, and relies on counting ballots—a procedure which, as our method says, does not involve any comparisons of ordinary utility.

Similarly, a dictatorial SWFL would rely on society's ability to identify the dictator, which implies indirectly an interpersonal comparison of labels. (If two individuals were both identified as the dictator, the SWFL would be in trouble if their preferences disagreed!) But the full social preferences over social states rely not only on identifying the dictator, but also on the precise content of the dictator's preferences.

We end this section with two remarks. The first is that when one says that such-and-such social preferences "rely on" or "involve" interpersonal comparisons of such-and-such a kind, one should not be misled into believing that the comparisons are absolutely indispensable in forming those social preferences. In fact, social preferences often need surprisingly little information.

This can be shown by first noticing that the concept of invariance transformations is rather restrictive, because it involves a particular form of transformation—namely, a vector of functions $\langle \varphi_i \rangle_{i \in N} \in \mathcal{I}_N$. As suggested by Sen

(1970a) and explored in Roberts (1980b), invariance transformations partition the domain D into equivalence classes made up of profiles which can be mutually transformed into one another by such transformations.[21] But the information actually used by social preferences is much better represented by directly partitioning the domain D into equivalence classes such that $f(U_N) = f(U'_N)$ for any two profiles U_N and U'_N in the same class. Then one can rigorously assert that f relies on a particular piece of data *provided that piece of data has the same value for all profiles belonging to the same equivalence class.*

U_N	i	j
x	1	6
y	3	4

U'_N	i	j
x	12	2
y	10	3

Table 4.1 Two utility profiles

Viewed this way, it is easy to check that very few interpersonal comparisons can be made with the data on which the most usual SWFLs really rely. For example, consider the maximin SWFL applied to the case when $X = \{x, y\}$, and N consists of the two individuals i and j. The two profiles indicated in Table 4.1 belong to the same equivalence class because the maximin rule ranks y above x in both cases. But all intra- and interpersonal comparisons of utility levels are overturned. This should not be so troubling. The maximin criterion, after all, does not really depend on whether i is better off than j in state x or not. It only depends on comparing minimum utilities between states.

Therefore, and this is our point here, in saying that the maximin rule relies on interpersonal comparisons of utility levels, one is really only referring to the computations made by any *algorithm* which implements the criterion, rather than to the actual information that is eventually relevant. If the social planner were directly told that the minimum utility is greater in social state y, there would be no (further) need to make an interpersonal comparison of any kind.

Our second remark is that there is a difficulty with the usual classification which puts (ONC) and (CNC) together into one class of no interpersonal com-

[21] Rigorously, this holds true only when the set of invariance transformations, together with the operation of function composition, form an algebraic group. This is indeed the case in the examples given above. See Roberts (1980b, p. 424) and Bossert and Weymark (ch. 20 in this volume)

parisons, while all other conditions such as (OLC) and (CUC) are recognized as involving interpersonal comparisons.[22]

The case of (ONC) is not problematic since it requires the SWFL effectively to ignore utility functions entirely and rely only on individual preferences. If utility functions are ignored, there is certainly no interpersonal comparison of anything related to utility.

As for (CNC), Sen (1970a) has shown that under (IIU), it is logically equivalent to (ONC). Intuitively, "the notion of preference intensity requires the comparison of at least three situations" (Roberts 1980b, pp. 432–433). The easy proof of this equivalence involves noting that when only two alternatives are considered, even a non-affine change of utility levels at these two alternatives which does not alter preferences could be obtained with affine transformations, one for each individual. Indeed, finding these transformations simply amounts to solving the pair of equations

$$U_i'(x) = \alpha_i + \beta_i U_i(x) \quad \text{and} \quad U_i'(y) = \alpha_i + \beta_i U_i(y)$$

for every $i \in N$. Therefore, a "welfarist" who adopts (IIU) may indeed blame (CNC) for preventing the SWFL from incorporating any interpersonal comparisons.

On the other hand, when (IIU) is not assumed, it is no longer obvious that (CNC) precludes interpersonal comparisons of everything related to utility. Of course, (CNC) does make it impossible to compare utility levels or utility differences. But this is also the case for non-comparable ratio scale (or NRS) measurability, defined as follows:

Non-comparable Ratio Scale (NRS): For all $U_N, U_N' \in D$, one has $f(U_N) = f(U_N')$ whenever there are positive constants $\beta_i \in \mathbb{R}$ such that $U_i' = \beta_i U_i$ for all $i \in N$.

Yet, as recalled by Bossert and Weymark (ch. 20 in this volume), under (NRS) measurability the ratios $U_i(x)/U_i(y)$ are interpersonally comparable. This reasoning can be extended to (CNC), because with this informational basis, any ratio of utility differences

$$\frac{U_i(x) - U_i(y)}{U_i(x') - U_i(y')}$$

has a definite value that is invariant under affine transformations. Accordingly these ratios, or any function of them, can be compared interpersonally.[23]

[22] As defined in section 3.1, (CNC) is based on independent affine transformations of individual utility functions.

[23] Even under (ONC), one can also compare something related to utility—namely, the signs of utility differences. But this is really degenerate, since it involves no more information than individual preferences.

This discussion therefore suggests a picture that is rather simpler than the one usually presented. Indeed, suppose interpersonal comparisons of utility are understood to include comparisons of utility expressions of any kind, possibly going beyond a purely ordinal representation of preferences. Then (trivially) only (ONC) does not involve interpersonal comparisons, while all other conditions are associated with some kind of comparison of more or less complex utility expressions.

Two considerations may have motivated the usual view of (CNC) as excluding any interpersonal comparisons, just like (ONC).

First, one can interpret cardinal utilities in terms of pure individual preferences. For example, assuming one is willing to focus on NM utility functions, cardinal utilities may be derived from preferences over lotteries, when these satisfy the expected utility hypothesis. Similarly, general additively separable intertemporal preferences can be represented by cardinal utilities. Under any such interpretation, the ratios of utility differences are just features of preferences. For instance, in the case when each U_i is an NM utility, the ratio

$$\frac{U_i(y) - U_i(x)}{U_i(y') - U_i(y')}$$

is the constant marginal rate of substitution between probability shifts from y to x and probability shifts from y' to x'—see, for example, Section 2.3 of Chapter 5 of this *Handbook*.

But, as we have seen above, the fact that some utility functions may be derived from preferences over some more or less exotic set of options just alters the empirical content of the utilities U_i, but does not alter the mathematical form of the SWFL f, or the class of interpersonal comparisons it involves. Recall that even cardinally comparable utilities can be derived from preferences over lotteries, with some normalization. Now, consider the case when cardinal utilities of the (CNC) kind are the arguments of an SWFL, without being derived in any way from individual preferences over anything (except, of course, that they must represent preferences over X). Then such utility functions do permit comparisons of ratios of utility differences, even though these ratios cannot be interpreted as marginal rates of substitution.

There is, however, another difference between the invariance transformations involved in (ONC), (CNC), (NRS) on the one hand, and (OLC), (CUC), (CFC), (CRS) on the other — where (CRS) is defined by:

Common Ratio Scale (CRS): For all $U_N, U'_N \in D$, one has $f(U_N) = f(U'_N)$ whenever there is a unique positive constant $\beta \in \mathbb{R}$ such that $U'_i = \beta U_i$ for all $i \in N$.

The latter class of four different kinds of invariance transformation all contain functions φ_i which are related to each other. In fact (OLC), (CFC) and (CRS) require these all to be the same function, whereas in (CUC) they must all involve the same multiplicative constant. This interdependence clearly suggests that some interpersonal information is at stake. By contrast, no such interdependence prevails in the other class of invariance transformations, and every individual function φ_i can be chosen independently of the others within some appropriate class—general increasing functions for (ONC), increasing affine functions for (CNC), and increasing linear functions for (NRS).

The problem with this criterion of interdependence is that it puts (NRS) on the side of non-comparability. Similarly, the translation-scale measurability condition (TSM)—which admits transformations $\langle \varphi_i \rangle_{i \in N}$ such that $\varphi_i(U_i) = U_i + \alpha_i$ for all $i \in N$—implies no interdependence between functions. It should therefore involve no interpersonal comparisons in this sense, even though such translations do preserve utility differences and accordingly permit comparisons of those differences.

It follows that, in order to distinguish (CNC) from the other conditions—namely, (ONC) and (NRS)—the interdependence criterion has to be combined with a condition on the class of admissible transformations. This must be the entire class of increasing affine functions, nothing less. Yet even the restriction to increasing affine functions seems somewhat arbitrary. This is because, just as the translations involved in (TSM) preserve units, whilst the increasing linear functions involved in (NRS) preserve zeros, the increasing affine functions preserve the usual economic measures of curvature.

For example, suppose each individual i has a twice continuously differentiable utility function $u_i(w_i)$ whose only argument is personal wealth w_i, and which satisfies $u_i'(w_i) > 0$, $u_i''(w_i) < 0$ for all positive w_i. Then a standard "absolute" measure of economic curvature is $-u_i''(w_i)/u_i'(w_i)$, and a "relative" measure is $-w_i\, u_i''(w_i)/u_i'(w_i)$ (for $w_i > 0$). Both measures are limits of ratios of utility differences. For example, l'Hôpital's rule can be applied to show that

$$-\frac{u_i''(w)}{u_i'(w)} = \lim_{\epsilon \to 0} \left[\lim_{\eta \to 0} \frac{1}{\eta} \left(\frac{u(w+\eta) - u(w)}{u(w+\epsilon) - u(w)} - \frac{u(w+\epsilon+\eta) - u(w+\epsilon)}{u(w+\epsilon) - u(w)} \right) \right].$$

It follows that both measures of curvature are invariant to separate increasing affine transformations of the form $\tilde{u}_i(w_i) = \alpha_i + \delta_i u_i(w_i)$ for all i and w_i, where $\delta_i > 0$ for all i—as is also easily shown by a direct argument.

In the case when u_i is a von Neumann–Morgenstern function representing preferences over lotteries, these two measures of curvature are the well-known degrees of absolute and relative risk aversion, respectively. Indeed, the measures can then be used to make interpersonal comparisons of risk aversion, following the results of Pratt (1964) in particular. They can also be regarded as measuring "fluctuation aversion" in the intertemporal context, when preferences are represented by an additively separable function of the form $\sum_{t=1}^{T} \beta_{it} u_i(c_{it})$, where β_{it} is i's "utility discount factor" for period t, c_{it} denotes i's consumption in period t, and the utility function u_i is assumed to be independent of t.

But no matter how the cardinal utility functions are constructed, we have seen how different individuals' ratios of utility differences and utility curvature are preserved by separate increasing affine transformations of utility. It follows that, even when (CNC) invariance is imposed, "economic" measures of curvature can be meaningfully compared between persons regardless of how they are interpreted empirically. The fact that welfare economists may have become more accustomed to manipulating zeros and units should not induce us to neglect the possibility of considering curvature, too, as a relevant and interpersonally comparable feature of any cardinal utility function.

In conclusion, though (CNC) does exclude many important kinds of interpersonal comparison, it certainly does not exclude *all* kinds.

5 The Basis of Interpersonal Comparisons

5.1 Descriptive or Normative?

Early authors in welfare economics, such as Pigou or Lerner, appear to have been strongly influenced by Benthamism and seemed to consider interpersonal comparisons of utility as essentially factual. Many of these authors, however, assumed that the human ability to produce subjective utility out of economic resources was evenly spread through the population. And it was not entirely clear whether this was a factual assumption or a convenient ethical principle designed to obtain egalitarian conclusions.

Robbins (1932, 1938) declared that interpersonal comparisons of utility were necessarily normative, mostly on the ground that there is no empirical way to confirm or disconfirm them, mental states being inscrutable. This had the effect of leading most economists to shy away from such sloppy foundations. Robbins himself, however, was quite clear that normative prescriptions and value judgments were part and parcel of "Political Economy", and that it was perfectly legitimate and even desirable to introduce them in analyses of economic policy, provided their non-scientific status was explicitly acknowledged.[24]

The debate about the normative character of interpersonal comparisons is not closed, and this is understandable because a statement like "individual i's utility is greater than j's" looks more like a factual proposition ("is") than a prescriptive one ("ought"). Various recent contributions include Cooter and Rappoport (1984, 1985), Little (1985), Hennipman (1988), Blaug (1992), Davis (1990, 1991), Rosenbaum (1995).

---

[24]See Robbins (1981) in particular, and also Hennipman (1992).

An example may illustrate why this matter is complicated. Consider the following statements, in relation to some policy reform under consideration:

Social Criterion (SC): One ought to adopt the reform if it increases the sum of utilities.

Individual Preferences (IP): Only i and j are affected by the reform, and according to their own preferences, i gains while j loses.

Interpersonal Comparison (IC): Individual i's gain of utility exceeds j's loss.

Social Decision (SD): The reform ought to be put into effect.

Among these propositions, [SC] and [SD] are evidently normative, while [IP] is purely factual. What about [IC]? If our concept of utility allowed it to be easily measured and also allowed utility gains and losses to be compared between individuals, then there would be no question that [IC] is also a purely factual observation. But if the measurement and comparison of utility are regarded as obscure and even contentious operations, it would seem inevitable that [IC] should rely more on compassionate moral intuition than on dispassionate observation.

This impression is strongly reinforced by the fact that, *assuming* [SC] *and* [IP], *then* [IC] *is logically equivalent to* [SD]. Since [SD] is clearly normative, and in view of Hume's Law that an "ought" cannot be deduced from an "is", it seems that any statement equivalent to it must also have some normative content. In effect, to say that i's gain is greater seems equivalent to saying that i is given priority (a normative decision). This argument, however, is quite misleading because it forgets that [SC] lies in the background.

An alternative, and probably more rigorous, view is that [IC] is just the factual element which enables the social planner to jump from the abstract principle [SC] to the concrete decision [SD]. If [SD] is adopted by the social planner and [IC] is derived from it by an observer who witnesses the decision, then both are simply wrong if the fact of the matter is that i's gain is actually smaller. Unless there is *no* fact of the matter, which then means that [SC] describes a totally inapplicable principle. This possible equivalence between interpersonal comparisons and social decisions, under some ethical assumptions, will be exploited in Section 7 of this chapter.

A similar example could be designed for the maximin criterion. We leave it to the reader to check that the above analysis would apply equally well to this criterion provided that [SC] and [IC] are replaced with:

Social Criterion (SC′): One ought to adopt the reform if it increases the minimum level of utility.

Interpersonal Comparison (IC$'$): Individual i's utility before the reform is less than all other individuals' utilities both before and after the reform.

The two examples above suggest that social decision-making usually has to follow a four-step procedure of the following general form:

(SWFL) Specify an SWFL (f).

(CIG) Formulate a concept of individual good (U_i) that is appropriate as an argument to f.

(OP) Choose an observable proxy for each individual's U_i that is rich enough for these proxies to determine $f(U_N)$.

(D) Collect the necessary data about the proxy for each individual's U_i in order to determine f.

The first step in this list is normative, and involves basic ethical principles of social choice. The second step (CIG) is also purely normative, and involves discussions about what kind of individual preferences ought to be taken into account, etc. The third (OP) is also commonly normative, although this may be less obvious. The reason it is usually normative is that when the concept of individual good is not easily observed, one has to choose some kind of observable proxy or indicator that is more or less related to the purely ethical concept. This choice is likely to rely more on rules of thumb than anything else, but it is definitely a normative step to decide whether a given indicator is acceptable or not. (Notice that, as is implicit in the formulation of (OP), the observable proxy need not be a direct measure of individual good U_i. Instead, as illustrated in the examples below, each proxy may involve a much less demanding kind of data. It is sufficient that decisions based on f can be made with the relevant information.) Only the last step (D) involves exclusively factual propositions about the real world.

Where are interpersonal comparisons needed in this list? Usually at the last step, when they cannot be made without any empirical basis. But one must admit that they are value-laden because they are performed in relation to some concept (CIG) and on the basis of some proxies (OP). Simon (1974) has argued that the distinction between a conceptual definition (CIG) and an empirical measure (OP) is useless and that individual good can be *defined* by an observable proxy directly. In some sense, this is just a particular (extreme) case of the above method, and one can view Simon's approach as just a particular attempt to get at the right concept.

Early welfare economists such as Pigou and Lerner tried to skip the difficult step (OP), and even (D), by relying on particular assumptions enabling them to draw practical conclusions without empirical worries. Their central argument was that a given amount of total income has to be shared equally under utilitarianism if individuals have the same strictly concave utility function of income. Therefore one could directly seek to equalize income without bothering to measure utility or even define it. Of course, when individuals have different tastes, they cannot have identical indirect utility functions of prices and income. Even when individuals have identical tastes, however, the assumption that they all have the same utility function of income, viewed as a factual assertion, cannot be assessed without a much more precise definition of utility. It may be true for some notions of utility and false for others. It may be true, in particular, for a conception of individual good for which U_i is just a measure of the social utility that the social planner derives from income made available to i (as opposed to an empirical notion of utility that would be related to the causal effect of income on i herself). In fact, this approach allows a normative conception of the individual good which builds in the above assumption that total income should be shared equally.

This example is quite important in that it points to the possibility that some definitions of individual good may require very little empirical measurement. In such cases (OP) may be vacuous, or at least create no normative worries about how best to approximate the ideal measure.

In the rest of this section and the next, we shall proceed to examine several examples illustrating how economists and other social scientists have tackled the problem of making social decisions that require interpersonal comparisons. For each example, we will try to identify the ethical choices (SWFL), (CIG), etc., as well as the kinds of interpersonal comparison involved in the specified SWFL, and the provisions for empirical application which are made.

The rest of this section discusses some general social choice problems, whereas the next section focuses on some specific economic environments that have received considerable attention in applied welfare economics. The list of examples is by no means exhaustive, of course. Instead, our purpose is chiefly to illustrate the above concepts by showing some of the many different ways in which interpersonal comparisons can be made and then incorporated into the social choice rule. We shall also clarify some difficulties in interpreting the form of the interpersonal comparisons that lie behind some of these approaches.

5.2 Capabilities

Following the work of Rawls (1971) and Sen (1973, 1980) in particular, egalitarianism has regained its dignified position among philosophers and economists. This has triggered a lively debate about the ethically appropriate equalisandum. Equality, yes, but equality of what? We will not try to cover the great variety of answers that have been proposed in the literature.[25] Instead we will focus on just one idea due to Sen (1985, 1987, 1992), as an interesting but not necessarily representative example.

Sen's proposal is to equalize "capabilities", which are to be understood as sets of "functionings" to which individuals have access. A functioning is just any kind of action or state that an individual may experience. A very similar proposal has been articulated by Cohen (1989, 1993). Such views can be roughly described as based on the following steps.

(SWFL) The greater is $\min_i U_i$, the better.

(CIG) U_i is an index of capability, which synthesizes all relevant dimensions of functionings.

The complexity of the concept proposed in (CIG) explains why the last steps (OP) and (D) have remained rather vague so far, although Nussbaum (1993) and others have attempted to compile a list of relevant items.

The content of this concept of individual good is predominantly normative. Its construction seems to require that the following questions be successively addressed. First, one has to define the list of relevant functionings. This involves difficult issues about what is important in contributing to the success of a life. It also requires coping with debates about whether a given item such as freedom of association has intrinsic or only instrumental importance. Second, some aggregation method must weigh the relative importance of the various functionings and produce a unidimensional index. Again, the conflicts between alternative views of the good life will hinder progress toward consensus regarding the appropriate index. Third, one must determine whether an individual has access to a bundle of functionings, or is barred from it. This issue is related to complex definitions of freedom and free will. Fourth, in view of the definition of access, an index of capability must synthesize the various levels of an index of functionings which are accessible, possibly with various degrees of accessibility. The second and the fourth issues may be reversed in this sequence, depending on which seems more tractable.

What is striking about Sen's proposal is that, apart from its egalitarian flavour, its general definition covers a host of different approaches because the richness of the concept of capabilities makes it possible to describe most notions of individual good under its guise. Even a pure Benthamite could accept the above (CIG), while deciding that the only functioning that matters is pleasure minus pain, and that a functioning is accessible only when it is actually obtained. There is a sense in which the proposal by Sen and Cohen is useful in

[25] Apart from the other work cited in this subsection, one may also mention Arneson (1989, 1990), Dworkin (1981, 2000), Fleurbaey (1995), Roemer (1993, 1998), van Parijs (1995).

revealing that all approaches have to address these difficult issues, even when they hide behind the illusion that subjective satisfaction is "evidently" what matters in life.

Of course, the Benthamite conception is not one that Sen or Cohen would approve. Indeed, they devote a lot of argument to advocating a notion of individual good that contains objective features, not only subjective happiness or satisfaction, and in which access (capability) differs in principle from actualization (functioning).

In view of the likely difficulty in reaching a consensus over how to measure capabilities, Sen has insisted that one should recognize the possibility of constructing a partial ordering based on domination over all dimensions, or similar clearcut comparisons.

5.3 Social Indicators of Happiness

There are several approaches which try to ground the measurement of subjective satisfaction or happiness on behavioural data. These are surveyed in Hammond (1991a), so we will mention here only one of these approaches, which seems to be gaining ground in some economic circles.

In many countries, over the course of several decades, representative samples of individuals have completed various questionnaires asking whether, regarding their own life in general, they feel "very happy", "pretty happy", "not too happy", etc. Such data make it possible to estimate directly the percentage of the population who are willing to report feeling these different qualitative degrees of happiness. One can also compare different countries, or the same country at different times, as well as studying the influence of various personal characteristics or macroeconomic variables upon people's reported degrees of happiness. Contributions to this literature include Easterlin (1974, 1995), Simon (1974), Myers (1993), Veenhoven (1993), Clark and Oswald (1994), Ng (1996), Frank (1997), Kahneman, Wakker and Sarin (1997), Oswald (1997), Winkelmann and Winkelmann (1998), Di Tella, MacCulloch and Oswald (2001), among many others.[26]

In this approach the definition of an SWFL is not exactly the primary issue. Instead, the view seems to be that something like the percentage of people who report being happy or very happy is worth maximizing for its own sake, while at the same time minimizing symptoms of severe distress such as rates of attempted suicide also deserves attention. How the various indicators should be combined into a consistent social goal is usually left in the dark.

---

[26] Myers (1993) and Oswald (1997) offer comprehensive surveys. The journal *Social Indicators Research* contains many publications related to the so-called "social indicators" movement.

What deserves attention for present purposes is the use made of answers to direct self-evaluative questions about happiness. These studies assume that a similar answer such as "I feel pretty happy" made by any pair of individuals can legitimate the conclusion that they are in a similar situation, in some ethically relevant sense. Some problems, however, are mentioned with respect to cross-country comparisons, because of language differences and cultural idiosyncrasies.[27] But one may also worry whether culture gaps *within* a country could make answers similarly hard to compare across different social groups. Manipulating the data as if answers were directly interpersonally comparable reflects a contentious value judgement, therefore—albeit one which is not necessarily unreasonable. After all, if the whole population were ready to declare itself very happy, that would certainly reflect a good feature of the society.

Assuming for instance that it is the percentage of "pretty happy" and "very happy" answers which should be maximized, one could simplistically describe this approach as follows, in the framework proposed above:

(SWFL) The greater the sum $\sum_i U_i$, the better.

(CIG) U_i measures happiness through people's verbal declarations in surveys.

(OP) Normalize by setting $U_i = 1$ if the answer is "pretty happy" or "very happy", and 0 otherwise.

Notice that with such a binary utility function, the leximin criterion coincides with utilitarianism.

In one of the papers cited above (Di Tella *et al.* 2001), a measure of the average life satisfaction in a given country for a given year is computed from answers to a question of the form "On the whole, are you: (i) not at all satisfied; (ii) not very satisfied; (iii) fairly satisfied; or (iv) very satisfied with the life you lead?" The "utility" numbers 1 to 4 respectively are associated with the four different answers. Since the authors focus on the average life satisfaction over the population, this is definitely a utilitarian approach to the social aggregation problem. The combination of the method of aggregation and this measure of individual good reflects the implicit judgement that it is as beneficial to make someone jump from "not at all satisfied" to "not very satisfied" as it is to propel someone from "fairly satisfied" to "very satisfied". Or that it is a matter of social indifference if one individual climbs from "not at all satisfied" to "fairly satisfied" while two others fall from "very satisfied" to "fairly satisfied".

Clark and Oswald (2002) have recently proposed using such sample surveys in order to estimate the monetary value, for the average person in the

[27] During the 1980s, the percentage of "very happy" people was as low as 13% in Italy and as high as 62% in Denmark! (Oswald 1997, p. 1819).

population, of life events such as sickness, marriage or unemployment. The idea is to collect data for a large sample of individuals, and to use linear regression techniques in order to estimate a "happiness equation". Then the coefficient for each different life event is divided by the coefficient for income in order to estimate the marginal rate of substitution between that life event and income. Accordingly, the computed measures are based entirely on ordinal non-comparable preferences for individuals, though these are then averaged over the population. This averaging process does require an assumption about the comparability of individual answers to happiness questionnaires.

5.4 Bargaining

The axiomatic theory of bargaining, like the theory of social choice, started with a 1950 paper, this time by John Nash. It has developed into an extensive literature since then, for which excellent surveys have been written by Roth (1985), Peters (1992), and Thomson (1994). Here, however, we will focus on the particular solution proposed by Nash himself, extended to the case of n persons. This is because the Nash solution has remained the most prominent due to its many interesting properties in cooperative as well as in non-cooperative settings.

Bargaining theory in general does not really fit into the traditional framework of social choice covered in this chapter, mainly because it does not seek to provide a complete ranking of all conceivable alternatives, but only a selection of one good alternative from every subset in a small domain of possible subsets of alternatives. Nonetheless, the Nash solution does maximize a preference ordering generated by a particular social welfare functional. It can therefore be analysed within the social choice context as well, as was done by Kaneko and Nakamura (1979) and Roberts (1980b). Both characterize the Nash SWFL with the help of some weaker variants of the independence axiom. In Roberts' result, (IIU) is replaced by an axiom which says that $f(U_N) =_A f(U'_N)$ whenever $U'_i =_{A\cup\{d\}} U_i$ for all $i \in N$, where $d \in X$ is a particular "disagreement" alternative which is supposed to come about in case the bargaining process fails to reach an agreement. Kaneko and Nakamura rely on a similar weakening of (IIU). They formulate it quite differently, however, because their framework is different, as explained below.

Let us try to fit the n-person Nash bargaining solution within our four-step procedure. One possibility for the first two steps is as follows:

(SWFL) The greater the product $\prod_{i\in N} [U_i(x) - U_i(d)]$, the better.

(CIG) U_i is a cardinal utility function.

In the bargaining context, the alternative d is whatever would obtain in case the negotiation process fails. The social ranking is meaningfully defined only for alternatives which are unanimously preferred to d. It is therefore convenient to assume, following the authors cited above, that d is the worst alternative for all agents. In the social choice framework, this can be viewed as a restriction on the domain of admissible preferences.

In the case when U_i is chosen from among the cardinally equivalent NM representations of a given preference ordering over lotteries, the data required to measure U_i are easily accessible in principle from observable choice behaviour, and this need not be explored further here. The ethical relevance of such preferences over lotteries for riskless social decision-making has been hotly debated ever since Arrow (1951), and it remains highly contentious. For references and a brief summary, see Hammond (1991a, section 5.2).

Actually, in Nash (1950) as well as in Kaneko and Nakamura (1979), X is supposed to be a set of lotteries, so that individual preferences over this set are sufficient information, and one can drop the requirement that cardinal utilities be provided from outside. For Kaneko's and Nakamura's approach in particular, a better description of the first two steps would probably be:

(SWFL) The greater the product $\prod_i [V_i(x) - V_i(d)]$, where V_i is any NM function representing the same preferences as U_i, the better.

(CIG) U_i is *any* utility function (not necessarily NM) representing i's preferences over X.

Let us examine what interpersonal comparisons are involved in these two slightly different approaches. In the first one above, one may say that the SWFL, which satisfies the (CNC) condition, involves the ordering R^* over $\mathbb{R}^n$ which is represented by $\prod_i u_i$, where $u_i(x) := U_i(x) - U_i(d)$. Since this particular R^* does require interpersonal comparisons of ratios, we may conclude that the Nash SWFL, in this first approach, relies on interpersonal comparisons of ratios of u_i, that is, of utility gains from d.[28]

In the second approach, the SWFL relies only on information about preferences, and therefore satisfies (ONC). It involves no interpersonal comparisons of utilities in the ordinary sense. However, our method also allows us to say that it does involve comparisons of ratios of NM gains from d. As can be deduced from some analysis of Chapter 5, recalled in Section 4.4 above, such ratios are actually (constant) marginal rates of substitutions between probability shifts between alternatives and d.

This distinction between the two approaches may look like a mere subtlety, but it is crucial for deciding what empirical basis underlies the required in-

[28]Notice that, although this SWFL satisfies (CNC) and is therefore able to compare any expression based on ratios of utility differences, it only compares a restricted set of such ratios.

terpersonal comparisons. If X is simply a set of riskless alternatives, then it is impossible to derive cardinally measurable utility functions from individual preferences over this set alone. By requiring U_i to be a cardinal function such as an NM utility, the first version of [CIG] above needs more information than individual preferences over X; it needs preferences over lotteries, for instance. Indeed, requiring information about preferences over a *broader* set is related to the need for some kind of interpersonal comparisons of utility (if only of the limited (CNC) kind). In the work of Kaneko and Nakamura, by contrast, nothing more than individual preferences over X is needed, but X has enough structure, and U_i satisfies enough restrictions (namely, the underlying preferences satisfy the expected utility hypothesis), so that cardinal indices of well-being can be derived from the parallel linear indifference curves over X, and then aggregated into a Nash SWFL.

5.5 Relative Utilitarianism

Several authors such as Isbell (1959), Kaplan (quoted by Arrow 1963), and Schick (1971) have proposed a variant of utilitarianism based on normalizing NM individual utility functions so that, for instance, their range coincides with the interval $[0,1]$. Dhillon and Mertens (1999) have recently provided an axiomatic characterization of this variant, in which they drop (IIU) while adopting a framework similar to Kaneko and Nakamura (1979).

Assume that X consists of the set of lotteries over a finite number of riskless alternatives. For each $i \in N$, let each V_i be any NM function representing the same preferences as U_i, and then define $V_i^0 := \min_{x\in X} V_i(x)$ as well as $V_i^1 := \max_{x\in X} V_i(x)$. Also, let us agree to disregard individuals i for whom $V_i^1 = V_i^0$. Then Dhillon and Mertens' approach can be described as follows:

(SWFL) The greater the sum

$$\sum_i \frac{V_i - V_i^0}{V_i^1 - V_i^0}$$

the better.

(CIG) U_i is any function representing i's preferences over X.

Here again the empirical basis of this approach is not problematic. An interesting question is how to describe the kind of interpersonal comparisons on which this SWFL relies. It does satisfy (ONC) and therefore avoids any comparison of utility. But it does compare something relative to NM functions. Recall that the (CNC) informational basis enables one to compare ratios of utility differences, or *any formula* based on them. It is not true, for instance, that the above SWFL directly compares the ratios

$$\frac{V_i - V_i^0}{V_i^1 - V_i^0}$$

any more than classical utilitarianism $\sum_i U_i$ compares utility levels U_i, in spite of the fact that the informational basis would allow it. But one can say that relative utilitarianism does compare *differences* of such ratios interpersonally, because it involves the utilitarian ordering R^* over $\mathbb{R}^n$ applied to the transformed vectors $\langle u_i \rangle_{i \in N}$ defined by

$$u_i := \frac{V_i(x) - V_i^0}{V_i^1 - V_i^0}.$$

The works by Kaneko and Nakamura (1979) and by Dhillon and Mertens (1999) demonstrate that interesting results in social choice can be obtained on the basis of data derived entirely from individual choice behaviour—or, put more bluntly, that non-dictatorial social choice is possible on the basis of individual preferences only. Their success is due mainly to relaxing (IIU), but is also due to the richer structure that is assumed for X, and to the corresponding restrictions that are imposed on individual preferences. A similar recipe for success can be found in the literature on fairness, which is our next topic.

5.6 Fairness

The literature on fair allocation is concerned with how to distribute private goods between several consumers. Unfortunately, it focuses almost exclusively on the selection of first-best allocations, and does not provide a fine-grained ranking of allocations.

Of course, one could simply divide the feasible set of allocations into two classes, fair and unfair, with each fair allocation strictly preferred to each unfair allocation, but the set of all fair allocations and the set of all unfair allocations both treated as incomparable. Yet for many preference profiles a condition such as Strong Pareto requires some unfair allocations to be preferred to others. This makes it impossible to satisfy condition (SP) for many preference profiles. Accordingly, the literature usually replaces the (SP) axiom by the weaker requirement that the selected allocations be Pareto efficient.

Nevertheless, there are at least two examples of SWFLs based on fairness. One proposed by Pazner and Schmeidler (1978) and by Pazner (1979) is defined as follows. First, an economic environment must be specified, and it is natural to choose a particular version of the general framework exploited throughout this chapter. In order to do so, assume that X consists of the set of feasible allocations of ℓ private goods when a total endowment vector $\Omega \in \mathbb{R}^{\ell}_{++}$ has to be shared among the individuals. An allocation denoted by $x = \langle x_i \rangle_{i \in N}$ specifies a consumption bundle $x_i \in \mathbb{R}^{\ell}_{+}$ for each $i \in N$. Feasibility requires that $\sum_i x_i \leqq \Omega$. Each individual utility function U_i is assumed to be self-centred, meaning that it is affected only by x_i, and not by x_j for any $j \neq i$. This assumes away consumption externalities—an assumption which may be justified either on the factual grounds that individuals are purely selfish in the problem at hand, or else on the ethical grounds that neither feelings of benevolence or malevolence, nor the desire to emulate one's neighbours, should affect the distribution of resources. To remind ourselves of this assumption, we use the notation $U_i(x_i)$ instead of $U_i(x)$. In addition, we impose the usual economic assumption that each individual's utility function is continuous, monotone, and quasi-concave.

Given the continuity and monotonicity assumptions on U_i, a particular ordinally equivalent utility function $v_i(x_i)$ can be defined as the share of Ω which i needs to achieve a consumption bundle indifferent to x_i. That is, $v_i(x_i)$ is the scalar multiple of Ω which satisfies the equality

$$U_i(x_i) = U_i(v_i(x_i)\,\Omega).$$

This definition obviously implies that $v_i(0) = 0$ and $v_i(\Omega) = 1$, just as with relative utilitarianism.

Finally, the *egalitarian-equivalent* SWFL that Pazner and Schmeidler proposed is defined by simply applying the leximin criterion to the utility vector $\langle v_i(x_i) \rangle_{i \in N}$. In other words, one has:

(SWFL) The leximin criterion applied to $\langle v_i(x_i) \rangle_{i \in N}$.

(CIG) U_i is any continuous and monotonic function representing self-centred preferences over x_i.

The empirical basis of such self-centred preferences is usually considered not to be a problem, as long as one disregards incentives for preference revelation. But this is mainly because economists readily assume that preferences over consumption bundles are actually selfish, and are also unaffected by other individuals' consumption. When consumption externalities do occur, it is not always obvious how to construct appropriate self-centred preferences. Philosophers such as Goodin (1986) and Harsanyi (1982) have examined how to "launder" preferences so as to remove undesirable features. Another proposal by Kolm (1995) consists in looking at preferences over allocations in which all consumers share an identical consumption bundle.

This SWFL evidently relies on interpersonal comparisons of the levels of the normalized utility functions v_i. However, like relative utilitarianism, it satisfies

(ONC) and therefore does not involve any comparisons of utility. Notice that the normalized functions v_i can be constructed thanks to the topological structure of X and the economic restrictions on the utility profile. But the main feature of this SWFL is that it does not satisfy (IIU). It does, however, satisfy a weaker independence condition originally due to Pazner (1979)—namely, that the social ranking of two allocations depends only on the individual indifference curves going through the bundles at the two allocations. See also Samuelson (1977), Mayston (1974, 1980), and Simmons (1980) for related concepts.

The above egalitarian-equivalent SWFL, however, is not the most prominent solution in this model. Instead it is the Walrasian equilibrium with equalized income, for which axiomatic justifications have been provided by Gevers (1986) and Thomson (1988). This can be described in an intuitively appealing way as the procedure which consists of giving every individual an equal endowment of Ω/n, then setting up a perfectly competitive market. By itself, this is not a fine-grained SWFL.

Nevertheless, following Fleurbaey and Maniquet (forthcoming), one can construct an SWFL which always selects such egalitarian Walrasian allocations as the first-best outcomes. Indeed, define the usual closed upper contour set $UC_i(x_i) := \{x'_i \in \mathbb{R}^\ell_+ \mid U_i(x'_i) \geq U_i(x_i)\}$. Then the proposed SWFL generates the social ordering that is represented by the "Bergson" social welfare function defined by

$$\Lambda(x) := \min \{\lambda \mid \lambda\Omega \in \text{ co } [\cup_{i \in N} UC_i(x_i)]\},$$

In words, $\Lambda(x)$ is equal to the minimum share of Ω which belongs to the convex hull of the union of the sets UC_i.

In summary, one has:

(SWFL) The greater $\Lambda(x)$, the better.

(CIG) U_i is any economic utility function representing self-centred preferences over x_i.

Again, this SWFL does satisfy (ONC). It also violates (IIU) but satisfies the weaker axiom mentioned above. What kind of interpersonal comparisons does it rely upon? It does not seem easy to uncover any preference ordering R^* over real vectors that would serve to determine $\Lambda(x)$, beyond the trivial fact that in virtue of Anonymity, all individuals count equally in the computation. An alternative presentation of this definition may help to understand why.

Denote the individual expenditure function as

$$e_i(p, u) := \min\{\, p\, x_i \mid x_i \in \mathbb{R}^\ell_+,\ U_i(x_i) \geq u\,\}.$$

Next, given any fixed price vector $p > 0$, define the "money-metric share" utility function $v_i(p, x_i)$ by

$$v_i(p, x_i) := \frac{e_i(p, U_i(x_i))}{p\,\Omega}.$$

In other words, v_i measures the share of Ω that, as an initial endowment when the price vector is p, would give i the same utility as x_i. Now, one can easily check that

$$\Lambda(x) = \max_p \min_i v_i(p, x_i).$$

Therefore, one can define this SWFL by applying the maximin criterion to the utility levels $v_i(p^m(x), x_i)$, where $p^m(x)$ solves the above maximization problem. But in this formulation, the evaluation of individual i's situation depends on $p^m(x)$, and thereby on the whole profile of preferences.

In other words, $v_i(p^m(x), x_i)$ is not a description of i's situation of the separable form $U_i(x)$, which was stipulated in the general definition of interpersonal comparisons previously proposed in Section 4.4. It is not clear whether this definition can be extended in any meaningful way to include such non-separable criteria.

5.7 The Condorcet Criterion

In the context of voting, the set X is not given any structure. No significant domain restriction is therefore available apart from the traditional case of single-peaked preferences (or unimodal utility functions) mentioned earlier. This implies that the only way to obtain positive results is to relax (IIU). The literature on voting rules also focuses on selecting a subset of alternatives instead of constructing fine-grained social preferences. Accordingly, it does not retain (SP) and so does not yield many results having direct interpretations within the usual framework of social choice theory.

Nevertheless, starting with Young and Levenglick (1978), in various interesting contributions Peyton Young has used voting models to analyse a variant of the social choice problem—see also Young (1988, 1994, 1995). The modification requires finding a correspondence which, for each preference or utility profile, determines a set of orderings over X, not necessarily the unique ordering generated by any SWFL. This approach has two particular features which deserve to be described here.

First, it relies on restricting the scope of (IIU) to pairs $x, y \in X$ of alternatives which, given the profile U_N, are *adjacent* in the social ranking P corresponding to $f(U_N)$—meaning that $x\ P\ y$, but there is no z such that $x\ P\ z\ P\ y$. Only then does (IIU) apply and require there to be a selected ordering R' with $x\ P'\ y$ for any profile of individual utilities U'_N which coincides with U_N on $\{x, y\}$.[29]

Second, the Condorcet criterion (also called the Kemeny-Young rule) that is characterized on the basis of this weaker (IIU) condition has a remarkable way of avoiding interpersonal comparisons. It is usually defined for profiles of *strict* individual preferences which allow no indifference; for these, it relies exclusively on pairwise applications of majority rule. Indeed, define

$$m(x, y) := \#\{i \in N \mid U_i(x) > U_i(y)\}$$

and

$$M(R) := \{(x, y) \in X \times X \mid x\ P\ y\}.$$

Then the Condorcet criterion selects all strict orderings R which maximize the expression

$$\sum_{(x,y)\in M(R)} m(x, y)$$

over the set of all logically possible strict orderings of X. In other words, the rankings deemed best are those which would be approved by the largest aggregate number of voters if voting took place separately over each pair.

What kind of interpersonal comparisons does this criterion rely on? The only point at which individuals are compared is when each is given a unit weight in the computation of $m(x, y)$. Letting $I(a) = 1$ if $a > 0$ and $I(a) = 0$ otherwise, one can rewrite

$$m(x, y) = \sum_i w_i I\left(U_i(x) - U_i(y)\right)$$

with $w_i = 1$ for all $i \in N$. With other weights w_i one could easily bias this Condorcet criterion in order to fit some individuals' preferences more closely.[30] But, just like majority rule, the criterion is invariant to any common rescaling of the weights, with $w'_i = \beta w_i$ for $\beta > 0$. In summary, the Condorcet criterion relies on exactly the same interpersonal comparisons as majority rule.

[29] Actually, this is not exactly Young's condition because he relies on a pairwise consistency condition which involves variable sets of alternatives. We ignore this complication here.

[30] In the limit, letting $w_i = 0$ for all i except one individual d yields dictatorship. This does not imply that dictatorship relies on the same sort of comparisons as majority rule. For instance, we noticed in Section 4.3 above that dictatorship is a limiting case of weighted utilitarianism; yet one cannot infer that dictatorship involves comparisons of utility differences.

6 Interpersonally Comparable Measures of Economic Welfare

6.1 Optimal Income Taxation

Following Vickrey (1945) and Mirrlees (1971), the approach usually adopted in the literature on optimal income taxation is rather akin to that of the early welfare economists. More specifically, it relies on the factual assumption that all households have identical preferences over consumption and labour supply, which can be represented by a concave utility function. This assumption is recognized to be unrealistic, obviously. Nevertheless, as will become clear in due course, it greatly simplifies the problem of specifying a normatively appropriate social criterion, as well as making the ensuing optimal control problem much more tractable.

In one recent work due to Atkinson (1995), the four-step procedure set out above is carried out as follows:

(SWFL) The greater the sum $\sum_i U_i$, the better.

(CIG) Each individual's cardinal utility function U_i is the same strictly increasing concave transform $\psi(U^*)$ of the least concave representation U^* of the population's common preferences. The concave transform is chosen by the social planner.

(OP) Some flexible functional form is chosen, allowing the least concave representation U^* to depend on a parameter vector θ. This parameter vector also determines the consumption demand and labour supply functions. These functions, which are observable in principle, then serve as proxies for the chosen parameter vector θ^* and so for the utility function U^*.

(D) What needs to be estimated by appropriate econometric techniques are the parameters θ^* determining the common household preference relation over bundles of consumption and leisure or labour supply.

For some particular assumptions about the shape of household preferences it is even enough to know only the elasticity of labour supply, which can be more easily estimated than global features of preferences. The assumption that all households have the same preferences is quite indispensable with this approach, because it makes it acceptable, and even appealing, to apply the same concave transform to the least concave representation of their common preferences. This amounts to considering that two households with the same labour–consumption bundle have the same social priority. If heterogenous preferences for households were allowed, it would still be possible to estimate the least concave representation of the various preferences, but it would be less easy to justify the choice of particular (different or identical) concave transforms for different preferences.[31]

Of course, the function U^* has no empirical basis except that derived from the observation of common household preferences—or rather, of their consumption demand and labour supply behaviour. Yet such observations determine utility only up to arbitrary increasing transformations. Thus, the additional information needed to determine ψ cannot come from observations of individual behaviour. Instead, the degree of concavity of the transform ψ and the associated interpersonal comparisons of utility differences are entirely determined by the social planner's normative attitude toward inequality. Alternatively, they can be seen as expressing society's collective degree of aversion to inequality. Once a particular function ψ has been chosen, the transformed utilities $U_i = \psi(U^*)$ can be incorporated in the additive social criterion specified in (SWFL), as well as used to perform all the related interpersonal comparisons.

There is an alternative approach to optimal income taxation, more in the spirit of Harsanyi, which is reflected in the original work of Vickrey (1945) as well as Mirrlees (1982). After the same (SWFL) as above, it features:

(CIG) U_i is any NM utility function (viewed as a function of x and i) whose expected value a typical individual would maximize in an "original position" when faced with an equal chance of becoming anyone in the population.

More generally, it is likely that many authors in this field, unlike Atkinson, think of U_i as measuring an individual's characteristic and not merely the social planner's preferences. Mirrlees (1982) did not provide a very detailed practical recipe for (OP) and (D), and one may surmise that with such a concept of individual good, the computational exercises which determine the optimal income tax schedule have a purely hypothetical status of the following kind: "If the individuals' common utility function happened to coincide with U_i, then the optimal tax schedule would be ..." In fact, since the social planner's preferences are no better known (or more stable) than the individuals' "true" utility, Atkinson's approach leads to similar hypothetical conclusions. In both approaches, a multiplicity of optimal tax computations is warranted in order to allow for the indeterminacy of the ultimate criterion.

[31] Note, however, that even when preferences are the same, the criticisms raised in Hammond (1991a, Section 5.4) concerning "isomorphy" still seem to retain their force. For instance, if a particular high-skilled job does make people better off in itself—perhaps because it confers high social status—then one might defend the view that a household with the same labor-consumption bundle as another but a better job should have lower social priority. See also Broome (1998) for closely related criticism, and also the discussion by Schüssler (1998).

6.2 *Isomorphic Cardinal Utility Functions*

Equivalence Scales

The previous discussion of optimal taxation assumed an economy in which consumers have identical cardinal utility functions depending on just two goods—consumption and labour/leisure. Here, we shall consider a more general economy in which consumer units $i \in N$ (which could be households rather than individuals) have heterogeneous but closely related cardinal utility functions which depend on quantities of ℓ consumption goods.

Indeed, for each $i \in N$ and for $g = 1, 2, \ldots, \ell$, let x_g^i denote consumer i's consumption of good g.[32] Using a framework pioneered by Barten (1964), it is assumed that there exists one common cardinal utility function $U^R : \mathbb{R}_+^\ell \to \mathbb{R}$ applying to everybody such that each consumer i's preferences over commodity vectors $x^i = (x_1^i, x_2^i, \ldots, x_\ell^i) \in \mathbb{R}_+^\ell$ are represented by the relevant member of the parametric family of utility functions

$$U_i(x^i; m^i) = U^R\left(\frac{x_1^i}{m_1^i}, \frac{x_2^i}{m_2^i}, \ldots, \frac{x_\ell^i}{m_\ell^i}\right) \tag{6.1}$$

for a vector $m^i = (m_1^i, m_2^i, \ldots, m_\ell^i) \in \mathbb{R}_{++}^\ell$ of positive *household equivalence scales* m_g^i that are specific to each household and each commodity. Note that U^R is actually the utility function of a *reference household* for whom each $m_g^R = 1$. In principle, this reference household may be purely hypothetical, but is often taken to be a household of some specific type—for example, a single adult with no special needs. Under this assumption, household preferences and utilities are said to be *isomorphic* because each can be derived from any other by rescaling the quantities of each good in an obvious way.

Given this parametric family of isomorphic utility functions, when faced with the price vector $p = (p_1, p_2, \ldots, p_\ell) \in \mathbb{R}_+^\ell$ and the income level y_i, the demands of each consuming unit i are derived by maximizing (6.1) w.r.t. x^i subject to the budget constraint $\sum_{g=1}^\ell p_g\, x_g^i = y_i$. Equivalently, defining the reference consumer's rescaled quantities $x_g^R := x_g^i / m_g^i$, the associated vector $x^R := (x_1^R, x_2^R, \ldots, x_\ell^R)$ is chosen to maximize $U^R(x^R)$ subject to $\sum_{g=1}^\ell p_g\, m_g^i\, x_g^R = y_i$. Under the usual assumptions of strictly convex, continuous and monotone preferences, this implies that different consumers' demands are related by the equations

$$x_g^i = h_g^i(p_1, p_2, \ldots, p_\ell, y_i) = m_g^i\, h_g^R(m_1^i\, p_1, m_2^i\, p_2, \ldots, m_\ell^i\, p_\ell, y_i) \tag{6.2}$$

---

[32] Our notation is chosen so that a superscript i denotes a vector pertaining to consumer i, whose components are indicated by subscripts. On the other hand, a subscript i denotes a scalar pertaining to consumer i.

where $h_g^R(p_1, p_2, \ldots, p_\ell, y_i)$ is the reference household's demand function for commodity g. The corresponding indirect utility functions obviously satisfy

$$\begin{aligned} V_i(p, y_i) &\equiv U_i(h_1^i(p, y_i), \ldots, h_\ell^i(p, y_i)) \\ &\equiv U^R(h_1^R(m_1^i\, p_1, \ldots, m_\ell^i\, p_\ell, y_i), \ldots, h_\ell^R(m_1^i\, p_1, \ldots, m_\ell^i\, p_\ell, y_i)) \\ &\equiv V^R(m_1^i\, p_1, \ldots, m_\ell^i\, p_\ell, y_i) \end{aligned} \tag{6.3}$$

Household Attributes

Equivalence scales represent a special case of the more general framework considered by Deaton and Muellbauer (1980, pp. 222–227). Suppose that each household i has a parametric utility function of the form

$$U_i(x^i) = U(x^i; a^i)$$

for some common utility function U and some finite-dimensional "attribute" vector a^i. We assume that a^i is sufficient to determine not only the household's demand behaviour but also, ultimately, how its welfare level should be compared with that of other households, as well as how its utility function should be aggregated into the social welfare functional. Obviously the demand functions and indirect utility functions of households with different attributes are related by the equations

$$h_g^i(p, y_i) = h_g(p, y_i; a^i) \quad (g = 1, \ldots, \ell); \qquad V_i(p, y_i) = V(p, y_i; a^i).$$

for suitable common demand functions $h_g(\cdot; a^i)$ and indirect utility functions $V(\cdot; a^i)$. The associated expenditure functions, which measure the cost of achieving a given utility level, satisfy

$$e_i(p, u_i) = e(p, u_i; a^i)$$

Note that, for each fixed p and a, the two functions V and e must be inverses of each other—that is, they must satisfy

$$e(p, V(p, y; a); a) \equiv y; \quad V(p, e(p, u; a); a) \equiv u \tag{6.4}$$

for all y, u.

A convenient way to measure the different households' utilities in this framework is through their *money metric* indirect utility functions $V^M(p, y; a)$. These transformed utility functions earn their name by being constructed, separately for each attribute vector a, so that $V^M(p^R, y; a) \equiv y$ for a particular *reference price vector* p^R, which for simplicity is assumed to be the same for all households. In this way, at the reference price vector, each household's utility becomes identified with its money income.

In order to determine $V^M(p,y;a)$ for any triple (p,y,a) with $p \neq p^R$, let us replace u by $V(p,y;a)$ and p by p^R in the second part of (6.4). The result is

$$V\left(p^R, e(p^R, V(p,y;a);a);a\right) = V(p,y;a).$$

Since $V^M(\cdot;a)$ is an ordinal transformation of $V(\cdot;a)$, it follows that

$$V^M(p,y;a) = V^M\left(p^R, e(p^R, V(p,y;a);a);a\right),$$

Using the identity $V^M(p^R, y';a) \equiv y'$ to simplify the right-hand side, we obtain

$$V^M(p,y;a) = e(p^R, V(p,y;a);a). \tag{6.5}$$

This equation serves to define the money metric utility function $V^M(\cdot;a)$ for each household attribute a.

Note carefully that attributes are defined so that two households with identical attributes and identical incomes must have the same utility number when confronted with the same price vector. Nevertheless, because interpersonal comparisons between households with different attribute vectors a have not been introduced so far, the different money metric indirect utility functions $V^M(\cdot;a)$ remain ordinally non-comparable. That is, one can apply any simultaneous attribute-dependent transformations of the form $\tilde{V}(\cdot;a) = \psi(V^M(\cdot;a),a)$ to these utility functions, where $\psi(u,a)$ is any increasing function of u, for each fixed a. In particular, the equality $V^M(p',y';a') = V^M(p'',y'';a'')$ has no ethical significance when $a' \neq a''$—indeed, it will be true by construction whenever $p' = p'' = p^R$ and $y' = y''$ no matter how different the two attribute vectors a' and a'' may be, or how much greater the needs of either household may be relative to the other.

At this stage we do introduce comparisons of utility levels between households with different attributes. Specifically, we choose particular attribute-dependent transformations $\psi(u,a)$ so that the new utility functions $\tilde{V}(\cdot;a) = \psi(V^M(\cdot;a),a)$ have the property that the equality $\tilde{V}(p',y';a') = \tilde{V}(p'',y'';a'')$ does signify an ethical judgement that the two households in their respective situations have the same utility level. In this way, we recalibrate different households' money metric utility functions onto one common scale in accord with these level comparisons. Let $\tilde{e}(p,u;a)$ denote the appropriately transformed expenditure function, defined so that it and $\tilde{V}(p,y;a)$ are inverses of each other for each fixed pair (p,a)—i.e., they satisfy (6.4).

Given the interpersonally comparable utility function $\tilde{V}(p,y;a)$, applying yet another suitable increasing transformation ϕ to all households' utility functions simultaneously yields a *level comparable money metric utility function* $V^*(p,y;a) = \phi(\tilde{V}(p,y;a))$ with the property that, for a particular *reference* *household* with attribute vector a^R, and at the reference price vector p^R, one has $V^*(p^R, y; a^R) = V^M(p^R, y; a^R) = y$ for all y. This requires that

$$y = V^*(p^R, y; a^R) = \phi(\tilde{V}(p^R, y; a^R)) = \phi(\psi(V^M(p^R, y; a^R), a^R)) = \phi(\psi(y, a^R))$$

for all income levels y. So ϕ must be the inverse of the particular transformation $\psi(\cdot, a^R)$ that has been used to convert $V^M(\cdot; a^R)$ into the level comparable utility function $\tilde{V}(\cdot; a^R)$ of the reference household.

Arguing as in the derivation of (6.5), given any triple (p, y, a), the unique appropriate value of $V^*(p, y; a)$ can also be found from the chain

$$\begin{aligned} V^*(p, y; a) &= \phi\left(\tilde{V}(p, y; a)\right) = \phi\left(\tilde{V}\left(p^R, \tilde{e}(p^R, \tilde{V}(p, y; a); a^R); a^R\right)\right) \\ &= V^*\left(p^R, \tilde{e}(p^R, \tilde{V}(p, y; a); a^R); a^R\right) \end{aligned}$$

Here, the second equality follows from applying the second part of (6.4) with u replaced by $\tilde{V}(p, y; a)$, p by p^R, and a by a^R. Because of the identity $V^*(p^R, y'; a^R) \equiv y'$, it follows that

$$V^*(p, y; a) = \tilde{e}(p^R, \tilde{V}(p, y; a); a^R) \tag{6.6}$$

That is, $V^*(p, y; a)$ must be the amount of income that the reference household needs at the reference price vector p^R in order to reach the same utility level as a household with attribute a and income y.

Let $e^*(p, u; a)$ denote the expenditure function associated with V^*. Obviously $e^*(p^R, u; a^R) = u$, but $e^*(p^R, u; a)$ will usually differ from u for the typical attribute vector $a \neq a^R$. Note that the related forms of the functions $V^*(p, y; a)$ and $e^*(p, u; a)$, like those of $\tilde{V}$ and $\tilde{e}$, depend upon whatever ethical values underlie the interpersonal comparisons needed to construct $\tilde{V}$. In particular, these functions cannot be inferred from demand behaviour alone.

Consider now the scalar function $\mu(u, a)$ defined by

$$\mu(u, a) := e^*(p^R, u; a)/e^*(p^R, u; a^R) \tag{6.7}$$

This ratio, which Lewbel (1989) calls a *cost of characteristics index*, can be interpreted as the equivalence scale representing the proportionate cost of living (or cost of achieving utility level u) at the reference price vector p^R for a household with attributes a, as compared with the reference household. Because of (6.4), note that $y/\mu(u, a) = e^*(p^R, u; a^R)$ when $y = e^*(p^R, u; a)$. It follows that $\mu(u, a)$ must satisfy the equation

$$V^*(p^R, y; a) = V^*\left(p^R, \frac{y}{\mu(u, a)}; a^R\right)$$

As Deaton and Muellbauer (1980, p. 224) point out, equations (6.4) and (6.6) together imply that the indirect utility function is determined implicitly by

$$u = V^*(p, y; a) = y \, \frac{e^*(p^R, u; a^R)}{e^*(p^R, u; a)} \, \frac{e^*(p^R, u; a)}{e^*(p, u; a)} = \frac{y}{\mu(u, a) \, P(p, u; a)} \tag{6.8}$$

where the function $P(p, u; a)$ is defined by

$$P(p, u; a) := e^*(p, u; a)/e^*(p^R, u; a)$$

This suggests that $P(p, u; a)$ can be interpreted as the true cost-of-living index for a household with attribute a facing price vector p, relative to the cost-of-living at the reference price vector p^R. Unlike $\mu(u, a)$, however, the function $P(p, u; a)$ is invariant under increasing attribute dependent transformations of the form $V \mapsto \psi(V; a)$, thus allowing $P(p, u; a)$ to be inferred from demand behaviour.

Inequality Aversion and Social Welfare

Suppose that the indirect utility function V^* representing interpersonal comparisons of utility levels really is cardinal. In this case, several authors have argued that differences in the values of $V^*(\cdot; a)$ for different attribute vectors a should also be interpersonally comparable—see the work cited in Blackorby and Donaldson (1991) and in Section 5.4 on pp. 221–4 of Hammond (1991a), as well as Jorgenson (1990, 1997b). Alternatively, Deaton and Muellbauer (1980, p. 225) suggest that, even if the level comparable money metric functions $V^*(p, y_i; a^i)$ specified in (6.6) are not regarded as interpersonally comparable welfare indicators, nevertheless they do allow comparisons "of the objective circumstances—the constraints—faced by each individual." This, they claim, makes the functions suitable for measuring inequality. If this is accepted, and if the measure of inequality corresponds to the loss of social welfare as it does in the approach pioneered by Kolm (1968) and Atkinson (1970), then it seems natural to postulate that the indirect SWFL should depend on the household indirect utility functions $V^*(\cdot; a^i)$ in a way that satisfies (CFC) invariance.

In fact, when the level comparable money metric indirect utility functions $V^*(p, y_i; a^i)$ are being used, there is a natural zero level of money metric utility also—namely, the reference household's utility level $V^*(p^R, 0; a^R) = 0$ when it faces the reference price vector with zero (unearned) income. Then it makes sense to impose the stronger requirement of (CRS) invariance. This allows one to consider indirect Bergson social welfare functions, whose arguments are the price vector p and the income distribution $y^N = (y_i)_{i \in N}$, taking the specific form

$$W_\rho(p, y^N) = \frac{1}{1-\rho} \sum_{i \in N} [V^*(p, y_i; a^i)]^{1-\rho} \tag{6.9}$$

for some parameter $\rho \geq 0$ satisfying $\rho \neq 1$. This is a natural extension to many goods of the one-parameter family of social welfare functions considered by Atkinson (1970). By analogy with the well-known Arrow–Pratt measure of relative risk aversion, ρ is called the constant "relative rate of inequality aversion".

When $\rho = 1$, one uses instead the alternative logarithmic form

$$W_1(p, y^N) = \sum_{i \in N} \ln[V^*(p, y_i; a^i)] \tag{6.10}$$

When $\rho = 0$ there is no inequality aversion at all; only mean income is relevant to welfare. When $\rho = \infty$ there is extreme inequality aversion, with $W(p, y^N) \equiv \min_{i \in N} V^*(p, y_i; a^i)$ as in the maximin criterion.[33]

It should be noted that the procedure specified above only defines an indirect social welfare function, and so really only applies to allocations that can be decentralized by facing each consumer with the same commodity price vector. These can be Walrasian or competitive equilibrium allocations, possibly in markets affected by lump-sum redistribution of wealth. More generally, the procedure can also be applied to the allocations which are demanded by consumers when they all face the same linear prices for each good that differ from producer prices—perhaps as a result of linear commodity taxation, as in Diamond and Mirrlees (1971). In principle, in order to extend the social welfare ordering to other allocations, one could construct the level comparable direct utility function $U^*(x; a)$ that corresponds to the level comparable money metric indirect utility function $V^*(p, y; a)$, and then use this to construct a direct social welfare function W_ρ^D of the form

$$W_\rho^D(x_N) = \frac{1}{1-\rho} \sum_{i \in N} [U^*(x^i; a^i)]^{1-\rho}$$

Generally, however, it is not possible to derive an explicit analytical expression for U^*.

In the framework considered here, when indirect utility functions are being aggregated, our four-step procedure could be carried out as follows:

[33] A different way of taking limits, involving a version of the "overtaking" criterion used in optimal growth theory, leads to "leximin"—the lexicographic extension of maximin. For details, see Hammond (1975).

(SWFL) The greater the value of the indirect social welfare function $W_\rho(p, y^N)$, the better.

(CIG) The concept of the individual good is each consumer's welfare, as measured by the value of the level comparable money metric indirect utility function $V^*(p, y_i; a^i)$.

(OP) Assuming that the form of the money metric indirect utility function $V^*(p, y; a; \theta)$ is known up to a parameter θ, and that $V^*(\cdot; a; \theta)$ is a different function of (p, y) for each attribute vector a, the observable proxy which determines each household's money metric utility is the corresponding vector demand function $h^*(p, y_i; a^i; \theta)$ whose components can be determined from Roy's identity

$$h_g^*(p, y_i; a^i; \theta) = -\frac{\partial V^*/\partial p_g}{\partial V^*/\partial y_i}$$

for $g = 1, 2, \ldots, \ell$.

(D) Appropriate econometric techniques are needed to estimate any unknown common parameter vector θ which determines the vector demand function $h^*(p, y; a; \theta)$—or alternatively, if non-parametic techniques can be used, to estimate the function $h^*(p, y; a)$. Additional estimation may be needed to infer the attribute parameters a^i which determine each household's vector demand function $h^i(p, y_i) = h^*(p, y_i; a^i)$.

Notice in particular how, given the functional form specified under (OP), step (CIG) builds in special interpersonal comparisons based on what Deaton and Muellbauer call "objective circumstances", as discussed above. There seems to be no good ethical reason for maintaining these comparisons when constructing an SWFL. For example, some kind of welfare-weighted sum, such as

$$W_\rho(p, y^N) = \frac{1}{1-\rho} \sum_{i \in N} \omega_i [V^*(p, y_i; a^i)]^{1-\rho} \tag{6.11}$$

might be much more appropriate than the unweighted sum (6.9). In other words, this approach to constructing an SWFL, like any other involving interpersonal comparisons, cannot rely on demand behavior as the sole basis for those comparisons. The analyst must resist being seduced by simple functional forms which may surreptitiously convey dubious ethical value judgments.

Finally, whether the indirect social welfare function takes the unweighted form (6.9) or the weighted form (6.11), it may be worth recalling that only in very special cases will an optimal distribution of a fixed total income equate the levels of the function $V^*(p, y_i; a^i)$ for different individuals i. Indeed, as shown in Hammond (1977b), for such utility level equalization to be optimal, different households' demand functions need to be closely related—see also Sections 6.3 and 6.4 below.

Lewbel's Independence of Base Condition

An important special case occurs when the cost of characteristics index defined by (6.7) happens to be independent of u, no matter how the reference price vector p^R is chosen. This is the *independence of base* (or IB) property whose implications Lewbel (1989, 1991, 1993) in particular has analysed. The property is satisfied if and only if there exists a cost of characteristics function $c(p, a)$, independent of u, such that

$$c(p, a) := e^*(p, u; a)/e^*(p, u; a^R) \tag{6.12}$$

for each price vector p and attribute vector a. In this way, the proportional cost of each attribute vector a, relative to that of the reference attribute vector a^R, is expressed as a function of the price vector p, as one would expect. Because e^* is homogeneous of degree 1 in p, it follows that c must be homogeneous of degree 0.

Equation (6.12) evidently implies that

$$e^*(p, u; a) \equiv c(p, a)\, e^R(p, u)$$

where $e^R(p, u) := e^*(p, u; a^R)$ is the expenditure function of the reference consumer with attribute vector a^R. The existence of such a multiplicative decomposition for $e^*(p, u; a)$ appears as Lemma 1 in both Lewbel (1989) and Lewbel (1991).

Next, put $y = e^*(p, u; a)$. Then $y/c(p, a) = e^*(p, u; a^R)$. Applying (6.4) twice, it follows that

$$u = V^*(p, y; a) = V^*(p, y/c(p, a); a^R) = V^R(p, y/c(p, a))$$

where $V^R(p, y) := V^*(p, y; a^R)$ is the indirect utility function of the reference consumer. Provided that different consumers' levels of welfare are compared at the same price vector p, one obviously has

$$V^*(p, y; a) \geq V^*(p, y'; a') \Longleftrightarrow y/c(p, a) \geq y'/c(p, a') \tag{6.13}$$

This property leads Lewbel (1989, p. 382) to describe $y/c(p, a)$ as the *scaled income* of a consumer with income y and attribute vector a.

Note how, because of (6.13), the (IB) property implies that

$$\begin{aligned} V^*(p, y; a) \geq V^*(p, y'; a') \quad &\Longleftrightarrow \quad y/c(p, a) \geq y'/c(p, a') \\ \Longleftrightarrow \quad \lambda y/c(p, a) \geq \lambda y'/c(p, a') \quad &\Longleftrightarrow \quad V^*(p, \lambda y; a) \geq V^*(p, \lambda y'; a') \end{aligned}$$

for all $\lambda > 0$ and all p, p', a, a'. Indeed, as Blackorby and Donaldson (1993a) point out, the resulting property

$$V^*(p, y; a) \geq V^*(p, y'; a') \Longleftrightarrow V^*(p, \lambda y; a) \geq V^*(p, \lambda y'; a'), \tag{6.14}$$

which they describe as *income-ratio comparability* (IRC), is actually equivalent to the (IB) property. This is because $e^*(p, u; a)/e^*(p, u; a^R) = y/\bar{y}$ where y and $\bar{y}$ satisfy $u = V^*(p, y; a) = V^*(p, \bar{y}; a^R)$. But then putting $\lambda = 1/\bar{y}$ and replacing y' by $\bar{y}$ in (6.14) implies that $V^*(p, y/\bar{y}; a) = V^*(p, 1; a^R)$. It follows that

$$e^*(p, u; a)/e^*(p, u; a^R) = y/\bar{y} = e^*(p, V^*(p, 1; a^R); a)$$

This is indeed independent of u, as property (IB) requires.

When property (IB) is satisfied, an obvious slight variation of (6.8) allows the money metric utility function to be determined implicitly from the equation

$$u = \tilde{V}^*(p, y; a) = y \, \frac{e^*(p^R, u; a^R)}{e^*(p, u; a)} \, \frac{e^*(p, u; a)}{e^*(p, u; a)} = \frac{y}{c(p, a)\, P(p, u)}$$

Here

$$P(p, u) := e^*(p, u; a^R)/e^*(p^R, u; a^R) \tag{6.15}$$

is a cost-of-living index for the representative household when its standard of living is fixed at the utility level u. Thus, Lewbel's measure of scaled income has to be deflated by this cost-of-living index.

The four-step procedure set out above is somewhat simplified in this special case. A much greater simplification arises when the reference households' preferences are homothetic, because then (6.15) implies that $P(p, u)$ is independent of u. In this special case the equation $u = \tilde{V}^*(p, y; a) = y/c(p, a)\, P(p)$ determines money metric utility explicitly.

Note that when the (IB) property is satisfied, multiplying $c(p, a)$ by any positive-valued scalar function $m(a)$ of the attribute vector a would make no difference to any household's demand behaviour. This reflects the fact that the function $c(p, a)$ embodies whatever ethical values lie behind the utility level comparisons involved in constructing the functions V^* and e^*, in addition to observable differences in households' demand behaviour.

6.3 Exact Aggregation: Parallel Linear Engel Curves

Gorman's Aggregation Condition

A problem with the above approach is the need in step [D] to estimate the common parameter vector θ of each household's demand function using micro data at the level of the individual household. In many cases, such data will either not be available at all, or at best be less reliable than data concerning aggregate demand. This has sparked some interest in conditions under which the common parameter vector θ, at least, can be estimated from data concerning aggregate demand together with a few statistics regarding the distribution of income.

Of particular interest here are the conditions given by Gorman (1953) to ensure that the aggregate demand $\sum_{i\in N} h^i_g(p, y_i)$ for each commodity $g = 1, 2, \ldots, \ell$ can be expressed as a function $H_g(p, Y)$ of the price vector p and of aggregate income $Y = \sum_{i\in N} y_i$, independent of how this aggregate income is distributed between different consumers. In other words, for each fixed price vector p, the functional equation

$$H_g\left(p, \sum_{i\in N} y_i\right) \equiv \sum_{i\in N} h^i_g(p, y_i) \tag{6.16}$$

must hold globally for all income distributions y^N.

Suppose for simplicity that each household's demand function h^i_g for good g is differentiable w.r.t. y_i. Then one can differentiate each side of (6.16) partially w.r.t. y_i in order to obtain the well-known result that, for each fixed p and each individual $i \in N$, one should have

$$\frac{\partial H_g}{\partial Y} = \frac{\partial h^i_g}{\partial y_i}, \tag{6.17}$$

independent of i. Of course, as p varies the equations (6.17) imply that, for each good g, one has $\partial H_g/\partial Y = \partial h^i_g/\partial y_i = b_g(p)$ for some common function $b_g(p)$, independent of i. Hence, after allowing for different constants of integration $c^i_g(p)$ for each individual $i \in N$, good g, and price vector p, there must exist $\mathbb{R}^\ell$-valued functions $c^i(p)$ and $b(p)$ of the price vector p such that

$$h^i(p, y_i) = c^i(p) + b(p)\, y_i \text{ (all } i \in N); \quad H(p, Y) = C(p) + b(p)\, Y \tag{6.18}$$

where $C(p) := \sum_{i\in N} c^i(p)$.[34]

[34]The same result can be proved under much weaker assumptions—in particular, without explicitly assuming differentiability. Indeed, for each $i \in N$, let $\underline{y}_i$ be a minimum income level at which each h^i_g is defined. Let $\underline{Y} := \sum_{i\in N} \underline{y}_i$. Given any fixed price vector p, define the function $f(\xi) := H_g(p, \underline{Y} + \xi) - H_g(p, \underline{Y})$ for all $\xi \geq 0$. Then the identity (6.16) implies: first, $f(\xi) = h^i_g(p, \underline{y}_i + \xi) - h^i_g(p, \underline{y}_i)$ for each $i \in N$; second, $f(\xi + \eta) = f(\xi) + f(\eta)$. The latter is a famous functional equation due to Cauchy. Now it is not hard to show successively that $f(n\xi) = nf(\xi)$ for any positive integer n, then that $f(r\xi) = rf(\xi)$ for any positive rational number r. Assuming that f is merely continuous, not necessarily differentiable, it follows that $f(\alpha\xi) = \alpha f(\xi)$ for any positive scalar α. But then $f(\xi) = \xi f(1)$ for any positive scalar ξ. Now it is easy to derive the aggregation conditions (6.17) and (6.18).

Accordingly all the different consumers' *Engel curves*, which for each fixed p and g, graph the expenditure $p_g\, x_g^i$ on good g against income y_i, must be parallel straight lines with common slope $p_g\, b_g(p)$. Moreover, to ensure budget balance, the functions $c^i(p)$ and $b(p)$ must obviously satisfy the relations $p\, c^i(p) = 0$ and $p\, b(p) = 1$. Finally, to ensure that the vector demand function $h^i(p, y_i)$ is homogeneous of degree 0, each $c^i(p)$ should be homogeneous of degree 0, while $b(p)$ should be homogeneous of degree -1.

Gorman's main contribution was to give conditions under which demand functions with these aggregation properties are consistent with individual utility maximization. In fact, sufficient conditions are that there should exist a particular ordinal measure of utility u_i, which is restricted to take non-negative values, such that different consumers' expenditure functions can all be expressed in the common linear form

$$e_i(p, u_i) = \gamma_i(p) + \beta(p)\, u_i \tag{6.19}$$

Here, the scalar functions $\gamma_i(p)$ and $\beta(p)$ should both be concave and homogeneous of degree 1, thus ensuring that $e_i(p, u_i)$ has the same properties for each fixed non-negative u_i. In addition, $\beta(p)$ should be positive valued, thus ensuring that e_i is always strictly increasing in u_i.

The corresponding ordinal indirect utility functions can be found by inverting (6.19) for each fixed p to obtain

$$V_i(p, y_i) = \frac{y_i - \gamma_i(p)}{\beta(p)} \tag{6.20}$$

Consumer i's indirect utility function V_i is defined on the domain of all pairs (p, y_i) such that income y_i exceeds the "subsistence level" $\gamma_i(p)$ associated with a zero level of utility. From now on, we assume that $\gamma_i(p)$ is the objectively specified least cost of achieving some basic minimum standard of living when the price vector is p.

Differentiate (6.19) partially w.r.t. each p_g in turn, and use the notation $\gamma_g^i(p)$, $\beta_g'(p)$ to denote the partial derivatives $\partial\gamma_i/\partial p_g$ and $\partial\beta/\partial p_g$. In this way, one obtains the "Hicksian" compensated demands

$$x_g^i(p, u_i) = \gamma_g^i(p) + \beta_g'(p)\, u_i \tag{6.21}$$

as functions of p, for each fixed utility level u_i. Then, using (6.20) to substitute for $u_i = V_i(p, y_i)$ yields the ordinary uncompensated demands

$$h_g^i(p, y_i) = \gamma_g^i(p) + \frac{\beta_g'(p)}{\beta(p)}\,[y_i - \gamma_i(p)] \tag{6.22}$$

Comparing (6.22) with (6.18), obviously one must have

$$b_g(p) = \beta'_g(p)/\beta(p) = \partial \ln \beta(p)/\partial p_g; \quad c^i_g(p) = \gamma^i_g(p) - b_g(p)\,\gamma_i(p) \tag{6.23}$$

The corresponding aggregate demand functions must satisfy

$$H_g(p, Y) = \Gamma'_g(p) + \frac{\beta'_g(p)}{\beta(p)}\,[Y - \Gamma(p)] \tag{6.24}$$

where $\Gamma(p) := \sum_{i \in N} \gamma_i(p)$ and $\Gamma'_g(p) := \partial\Gamma/\partial p_g$. Notice that (6.24) is the demand function of an aggregate "representative consumer" whose indirect utility function is $V(p, Y) = [Y - \Gamma(p)]/\beta(p)$.

Special Cases

There are several important special cases. One is the more commonly cited aggregation condition which Samuelson (1956) derived by imposing the additional requirement that all demand quantities are defined and non-negative whenever income is non-negative. Under this extra restriction, all the parallel linear Engel curves must pass through the origin. Then (6.18) requires that $c^i(p) \equiv 0$ for all $i \in N$, so all consumers must have identical homothetic preferences which generate identical demand functions $h^i_g(p, y) \equiv b_g(p)y$. This property suggests that the more general preferences corresponding to (6.20) should be described as *quasi-homothetic*, following Gorman (1961, 1976).

A second special case is the *linear expenditure system* originally formulated by Klein and Rubin (1947–48)—see also Geary (1949–50) and Stone (1954). This occurs when $\beta(p)$ is the multivariable "Cobb–Douglas" function $\prod_{g=1}^{\ell} p_g^{\beta_g}$, where the parameters β_g are non-negative real numbers which sum to one, while $\gamma_i(p)$ is the linear function $p\,\underline{x}^i$ for some fixed "subsistence" consumption vector $\underline{x}^i \in \mathbb{R}^\ell$. In this case, equation (6.22) implies that each household's expenditure on each good g is given by the expression

$$p_g\, h^i_g(p, y_i) = p_g\, \underline{x}^i_g + \beta_g\,(y_i - p\,\underline{x}^i) \tag{6.25}$$

Because the right-hand side of (6.25) is linear in the observable prices p_g and income y_i, the unknown parameters β_g and $\underline{x}^i_g$ can be estimated using a form of linear regression—see, for example, Deaton (1975, ch. 4). Moreover, each household has a direct utility function which can be expressed in the explicit form $U_i(x^i) \equiv \prod_{g=1}^{\ell} (x^i_g - \underline{x}^i_g)^{\beta_g}$—a functional form for which Gorman credibly claims priority, though its first appearance in print seems to have been in Samuelson (1947–8).

Yet another special case generalizes this linear expenditure system to the *constant elasticity of substitution* (or CES) system with the same $\gamma_i(p)$ as above, but with $\beta(p) \equiv \left[\sum_{g=1}^{\ell} \beta_g\, p_g^{(\epsilon-1)/\epsilon}\right]^{\epsilon/(\epsilon-1)}$ for some $\epsilon > 0$ with $\epsilon \neq 1$.[35] The corresponding expenditures satisfy

$$p_g\, h_g^i(p, y_i) = p_g\, \underline{x}_g^i + \frac{\beta_g\, p_g^{(\epsilon-1)/\epsilon}}{\sum_{f=1}^{\ell} \beta_f\, p_f^{(\epsilon-1)/\epsilon}}(y_i - p\, \underline{x}^i) \tag{6.26}$$

Note that (6.26) reduces to the linear expenditure system (6.25) in the limiting case when $\epsilon = 1$. For $\epsilon \neq 1$ each household $i \in N$ has a positive-valued direct utility function $U_i(x^i)$ which satisfies the equation

$$\frac{[U_i(x^i)]^{1-\epsilon}}{1-\epsilon} = \frac{1}{1-\epsilon} \sum_{g=1}^{\ell} \beta_g^{\epsilon} (x_g^i - \underline{x}_g^i)^{1-\epsilon}$$

For the rest of this subsection, we revert to the general case with demands given by (6.18) and indirect utilities by (6.20).

An Equity-Regarding SWFL

So far this subsection has concentrated on individual demand behaviour. Accordingly, the indirect utility functions $V_i(p, y_i)$ have been treated as ordinal non-comparable representations of preferences. Assume now that one can make interpersonal comparisons of the levels, differences, and any other relevant aspects of these utility functions. It then seems natural to impose the restriction that the SWFL W be *equity regarding* in the sense that it is increased by sufficiently small *progressive transfers* from consumers with higher utility levels to those with lower utility levels, as long as total income is preserved. In other words, the marginal utility of income should be higher for consumers with lower utility levels.

Then, provided that W is also additively separable, Paretian, and differentiable, it must take the form $\sum_{i \in N} \phi(V_i)$ for some increasing function ϕ independent of i such that the derivative ϕ' is decreasing. This is the case of *complete dual comparability* discussed in Hammond (1977b). The function ϕ may be interpreted as reflecting interpersonal comparisons of *social* utility. Specifically, ϕ must be chosen so that, for each pair of consumers $i, j \in N$, the ratio

$$\frac{\phi'(V_i)}{\phi'(V_j)} \frac{\partial V_i/\partial y_i}{\partial V_j/\partial y_j} = \phi'\left(\frac{y_i - \gamma_i(p)}{\beta(p)}\right) / \phi'\left(\frac{y_j - \gamma_j(p)}{\beta(p)}\right)$$

represents the social marginal rate of substitution between the incomes y_i, y_j of these two consumer units.

[35] This demand system owes its name to the formal mathematical similarity with the constant elasticity of substitution (or CES) production function due to Arrow *et al.* (1961).

A somewhat more general formulation introduces positive scalar welfare weights m_i reflecting the size or "normative significance" of consumer unit i—for example, the number of adult equivalents in a household, as discussed by Blackorby and Donaldson (1993a, b), for example. Then

$$V_i(p, y_i) := [y_i - \gamma_i(p)]/m_i\beta(p) \tag{6.27}$$

is an ordinal measure of welfare for each household member. A level interpersonal comparison of the form $V_i(p, y_i) > V_j(p, y_j)$ can be interpreted as signifying that, when the consumer price vector is p, household i with income y_i is better off as a whole than household j with income y_j. Instead of the unweighted sum $\sum_{i \in N} \phi(V_i)$, however, a more appropriate measure of social welfare would seem to be

$$W := \sum_{i \in N} m_i \phi(V_i). \tag{6.28}$$

Indeed, because (6.27) and (6.28) together imply that $\partial W/\partial y_i = \phi'(V_i)/\beta(p)$, this weighted sum does have the property that $V_i > V_j$ implies $\partial W/\partial y_i < \partial W/\partial y_j$. Thus, the weighted sum favours progressive transfers from households whose members have higher utility to those whose members have lower utility, as required for W to be equity regarding. This is the case of *intermediate dual comparability* discussed in Hammond (1977b, 1980). Note that an optimal distribution of a given total income Y takes the form $y_i = \gamma_i(p) + m_i[Y - \Gamma(p)]/M$ where $M := \sum_{i \in N} m_i$. Hence each relative weight m_i/M is equal to i's share of any incremental income.

An appealing special case occurs when $\phi(V) \equiv \frac{1}{1-\rho} V^{1-\rho}$ for some $\rho \geq 0$ with $\rho \neq 1$. In the case of identical homothetic preferences, the corresponding social welfare function is

$$W_\rho(p, y^N) \equiv \frac{1}{1-\rho} \sum_{i \in N} m_i \left[\frac{y_i}{\beta(p)}\right]^{1-\rho}$$

Obviously, when $m_i = 1$ for all $i \in N$, this reduces to a special case of the more general function defined by (6.9).

To summarize, in this case with parallel linear Engel curves, our four-step procedure can be carried out as follows:

(SWFL) The greater the welfare function $W \equiv \sum_{i \in N} m_i\, \phi(V_i)$, the better.

(CIG) Each consumer's cardinal indirect utility function is given by $V_i(p, y_i) = [y_i - \gamma_i(p)]/m_i\beta(p)$, as in (6.20).

(OP) Given the function $\beta(p)$, the observable proxies for each V_i are the scalar m_i, as well as the subsistence cost-of-living function $\gamma_i(p)$.

(D) Given the known subsistence cost-of-living functions $\gamma_i(p)$ and so the aggregate $\Gamma(p) = \sum_{i \in N} \gamma_i(p)$, data on aggregate demands, prices and aggregate incomes can be used to estimate $\Gamma(p)$ and $\beta(p)$ from (6.24). Then the joint distribution of consumers' incomes y_i, equivalence scales m_i and subsistence cost-of-living functions $\gamma_i(p)$ will determine aggregate social welfare.

Note that, although we are treating each household's welfare weight m_i as an "observable proxy", it is actually entirely independent of demand behaviour. Thus, it is really a purely ethical parameter, though it will obviously depend on observable features of household i.

Finally, we note that there has been some recent controversy over whether an SWFL of the form given in (6.28) is compatible with the standard individualistic approach. Specifically, it has become rather usual to presume that the ratio m_i/n_i of m_i, household i's welfare weight or number of equivalent adults, to n_i, the number of people in the household, should decrease as n_i increases—perhaps reflecting the idea that there should be economies of scale within a household. Then, if V_i is somehow constructed to measure the utility of a "representative" individual in this household, social welfare should presumably be calculated as $\sum_{i \in N} n_i \phi(V_i)$, with all individuals being given the same weight regardless of how large a household they belong to.

In contrast, when social welfare is measured by formula (6.28), with weights m_i applying to different households, then as Ebert (1997) and Shorrocks (1995) have noticed, it is possible to reallocate individuals, along with their equivalent shares of household income, between households in a way that increases social welfare. As an example, assume that the equivalence scale for a couple is 1.5, and $\gamma_i(p) = 0$ for all i. Now suppose two single-person households who each have income \$10,000 unite to form one couple with income \$15,000. Our assumptions imply that their individual utility is unchanged. Yet according to formula (6.28) the contribution of the couple to social welfare changes from $2\phi(10,000)$ to $1.5\phi(15,000/1.5)$—in other words, social welfare decreases by $0.5\phi(10,000)$. Generally, provided that m_i/n_i is decreasing in n_i, an increase of social welfare results whenever any large household whose members all have a positive utility level $\phi(V_i)$ is divided into several smaller ones, with all individuals receiving an equivalent income so that their utility levels are preserved. This seems a blatant violation of Pareto Indifference.

On the other hand, replacing (6.28) with an SWFL like $W \equiv \sum_{i \in N} n_i \phi(V_i)$ with $V_i = y_i/m_i\beta(p)$ implies that $\partial W/\partial y_i$ is proportional to $n_i \phi'(V_i)/m_i$. As Glewwe (1991) in particular points out, because ϕ' is strictly decreasing, this new form of W favours regressive income transfers from households i with slightly lower utility levels V_i but a high ratio m_i/n_i to other households j with slightly higher utility levels V_j but a low ratio m_j/n_j.

Really, this is yet another instance of the general problem that arises when an additive social welfare function is applied to a population with heterogeneous characteristics—a problem noticed in Arrow (1971) and discussed in Sen (1973) as well as Hammond (1977b). The problem disappears with infinite inequality aversion, in which case (6.28) is replaced with the maximin (or leximin) criterion.

6.4 Exact Aggregation: The Translog Model

Lau's Aggregation Condition

A different form of exact aggregation is due to Lau (1982). Different consumer units are distinguished by values of a finite-dimensional attribute vector $a = \langle a_j \rangle_{j=1}^J$. Lau's aggregation result forms the basis of the framework that has appeared extensively in the theoretical and empirical work by Jorgenson (1990, 1997) and various co-authors—see also Slesnick (2001). Following Christensen, Jorgenson and Lau (1975), this framework uses an indirect utility function $V^*(p, y; a)$ whose logarithm can be expressed in the *transcendental logarithmic* (or "translog") form

$$\begin{aligned} \ln V^*(p, y; a) \quad = \quad & \sum_{g=1}^{\ell} \alpha_g \ln\left(\frac{y}{p_g}\right) + \frac{1}{2} \sum_{g=1}^{\ell} \sum_{k=1}^{\ell} \beta_{gk} \ln\left(\frac{y}{p_g}\right) \ln\left(\frac{y}{p_k}\right) \\ & + \sum_{g=1}^{\ell} \sum_{j=1}^{J} \gamma_{gj} \ln\left(\frac{y}{p_g}\right) a_j \end{aligned} \tag{6.29}$$

Clearly, this function is explicitly constructed to be homogeneous of degree zero in (p, y). It loses no generality to impose the useful normalization $\sum_{g=1}^{\ell} \alpha_g = 1$. Also, as usual with a quadratic form, it loses no generality to replace both β_{gk} and β_{kg} with $\frac{1}{2}(\beta_{gk}+\beta_{kg})$, thus ensuring that the symmetry condition $\beta_{gk} = \beta_{kg}$ is satisfied for all g, k.

Recall Roy's identity, which states that $h_g^*(p, y; a) = -\frac{\partial V^*/\partial p_g}{\partial V^*/\partial y}$. It follows that the associated expenditure shares $w_g^*(p, y; a)$ devoted to each good g must satisfy

$$w_g^*(p, y; a) = \frac{p_g\, h_g^*(p, y; a)}{y} = -\frac{\partial V^*/\partial \ln p_g}{\partial V^*/\partial \ln y} = -\frac{\partial \ln V^*/\partial \ln p_g}{\partial \ln V^*/\partial \ln y} \tag{6.30}$$

and so

$$w_g^*(p, y; a) = \frac{1}{D(p, y; a)} \left[\alpha_g + \sum_{k=1}^{\ell} \beta_{gk} \ln\left(\frac{y}{p_k}\right) + \sum_{j=1}^{J} \gamma_{gj}\, a_j \right] \tag{6.31}$$

where the denominator is the positive scalar defined by

$$D(p, y; a) := \frac{\partial \ln V^*}{\partial \ln y} = 1 + \sum_{g=1}^{\ell} \sum_{k=1}^{\ell} \beta_{gk} \ln \left(\frac{y}{p_k} \right) + \sum_{k=1}^{\ell} \sum_{j=1}^{J} \gamma_{kj} \, a_j \tag{6.32}$$

The corresponding aggregate demands H_g must satisfy

$$p_g \, H_g = \sum_{i \in N} w_g^*(p, y_i; a^i) \, y_i \tag{6.33}$$

In order to make each function H_g depend on incomes only via the $\#J + 2$ aggregates $Y = \sum_{i \in N} y_i$, $\sum_{i \in N} y_i \ln y_i$, and $\sum_{i \in N} a_j^i \, y_i$ $(j = 1, 2, \ldots, J)$, one can impose the aggregation conditions due to Lau (1982) and to Jorgenson, Lau and Stoker (1982). As explained in the latter paper and in Jorgenson and Slesnick (1983), these conditions require that

$$\sum_{g=1}^{\ell} \sum_{k=1}^{\ell} \beta_{gk} = 0 \quad \text{and} \quad \sum_{g=1}^{\ell} \gamma_{gj} = 0 \quad (j = 1, 2, \ldots, J) \tag{6.34}$$

Then the denominator (6.32) takes the simpler form

$$D(p) = 1 - \sum_{g=1}^{\ell} \sum_{k=1}^{\ell} \beta_{gk} \ln p_k \tag{6.35}$$

which is independent of y and a, as well as homogeneous of degree 0.

From (6.31) and (6.33), the corresponding aggregate demands H_g satisfy

$$\begin{aligned} D(p) \, p_g \, H_g \quad = \quad & \left(\alpha_g - \sum_{k=1}^{\ell} \beta_{gk} \ln p_k \right) \sum_{i \in N} y_i + \sum_{k=1}^{\ell} \beta_{gk} \sum_{i \in N} y_i \ln y_i \\ & + \sum_{j=1}^{J} \gamma_{gj} \sum_{i \in N} a_j^i \, y_i \end{aligned} \tag{6.36}$$

This form allows the various unknown parameters α_g, β_{gk}, and γ_{gj} to be estimated, at least in principle, from enough dispersed observations of prices p_g and of the aggregates H_g, $\sum_{i \in N} y_i$, $\sum_{i \in N} y_i \ln y_i$, and $\sum_{i \in N} a_j^i \, y_i$ $(j = 1, 2, \ldots, J)$.

Under the aggregation conditions (6.34), which imply (6.35), the indirect utility function (6.29) simplifies to

$$\ln V^*(p, y; a) = D(p) \ln[y/m(p; a) P(p)] \tag{6.37}$$

where $P(p)$ is the *income deflator* defined by

$$\ln P(p) := \frac{1}{D(p)} \left[\sum_{g=1}^{\ell} \alpha_g \ln p_g - \frac{1}{2} \sum_{g=1}^{\ell} \sum_{k=1}^{\ell} \beta_{gk} \ln p_g \ln p_k \right] \tag{6.38}$$

and $m(p;a)$ is the *equivalence scale* defined by

$$\ln m(p;a) := \frac{1}{D(p)} \sum_{g=1}^{\ell} \sum_{j=1}^{J} \gamma_{gj} (\ln p_g)\, a_j \tag{6.39}$$

In particular, note that $P(p)$ is homogeneous of degree one, like a price index, whereas $m(p;a)$ is homogeneous of degree zero in p.

From (6.37) it follows that the consumer's expenditure function is given by $e^*(p,u;a) = m(p;a)\, P(p)\, u^{1/D(p)}$. This implies that, no matter what the attribute vector a^R of the reference household may be, Lewbel's cost of characteristics index $c(p,a) = e^*(p,u;a)/e^*(p,u;a^R)$ is $m(p;a)/m(p;a^R)$. Because this is independent of u, his (IB) property is satisfied.

An Indirect Social Welfare Functional

The particular form of indirect SWFL proposed by Jorgenson and Slesnick (1983) is

$$W_\rho(p, y^N) \equiv \bar{V} - \kappa(p, a_N) \left[\frac{\sum_{i \in N} m(p;a^i) |V^*(p, y_i; a^i) - \bar{V}|^{-\rho}}{\sum_{i \in N} m(p;a^i)} \right]^{-\frac{1}{\rho}}$$

where $\bar{V}$ is the weighted average utility defined by

$$\bar{V} := \sum_{i \in N} m(p;a^i) V^*(p, y_i; a^i) / \sum_{i \in N} m(p;a^i)$$

Furthermore, ρ is a parameter satisfying $\rho \leq -1$, and for each price vector p and attribute profile a_N, the non-negative constant $\kappa(p, a_N)$ is the largest consistent with $W_\rho(p, y^N)$ being an increasing function of y_i, for each $i \in N$.[36] This requires that

$$\kappa(p, a_N) = \left[\lambda(p, a_N) \{ 1 + [\lambda(p, a_N)]^{-\rho-1} \} \right]^{1/\rho}$$

[36] This is a necessary amendment to Jorgenson's (1990) suggestion that $\kappa(p, a_N)$ should be as large as possible consistent with the Pareto principle. Indeed, note that even when $\kappa(p, a_N) = 0$ and $W_\rho(p, y^N) \equiv \bar{V}$, the SWFL is consistent with the Pareto principle only in the trival case when each $m(p;a^i)$ is independent of p, as would occur if preferences were homothetic. See Lewbel (1993) for a similar result.

where

$$\lambda(p, a_N) = 1 - [\min_{i \in N} m(p; a^i) / \sum_{i \in N} m(p; a^i)]$$

As Jorgenson and Slesnick (1983) point out, this form of SWFL is equity regarding. Indeed, given aggregate income Y, choosing each consumer i's income y_i to satisfy

$$y_i / m(p; a^i) = Y / \sum_{i \in N} m(p; a^i)$$

equates both utility levels $V^*(p, y_i; a^i)$ and marginal utilities of income, which are given by $\partial W_\rho / \partial y_i = \partial \bar{V} / \partial y_i = D(p)\, m(p; a^i) \bar{V} / y_i \sum_{h \in N} m(p; a^h)$.[37] However, because the weights $m(p; a^i)$ depend on prices, the SWFL is not generally Paretian. Indeed, it is non-Paretian even in the special case when $\rho = -\infty$ and so[38]

$$W = W_{-\infty}(p, y^N) \equiv \min_{i \in N} \{ m(p; a^i) V^*(p, y_i; a^i) \}$$

To summarize, in this translog model with aggregation conditions imposed, our four-step procedure is carried out as follows:

(SWFL) The greater the non-Paretian indirect welfare function $W_\rho(p, y^N)$, the better.

(CIG) Each consumer's cardinally fully comparable indirect utility function is given by $V_i(p, y_i) = V^*(p, y_i; a^i) = D(p) \ln[y_i / P(p)\, m(p; a^i)]$.

(OP) The observable proxies for each consumer's V_i are the attribute vectors a^i which, together with the common parameters α_g, β_{gk}, and γ_{gj}, determine the functions $m(p; a^i)$, $P(p)$ and $D(p)$.

(D) Data on aggregate demands, prices, and the aggregate statistics $\sum_{i \in N} y_i$, $\sum_{i \in N} y_i \ln y_i$, and $\sum_{i \in N} a^i_j\, y_i$ $(j = 1, 2, \ldots, J)$ can be used to estimate the common parameters α_g, β_{gk}, and γ_{gj} from (6.36). For each $\rho \geq 0$ the value of W_ρ is then determined from the joint distribution of households' incomes y_i and of attribute vectors a^i.

---

[37] Formally, in the extreme case when $\rho = -1$, the partial derivative $\partial W_\rho / \partial y_i$ does not exist when $V^*(p, y_i; a^i) = \bar{V}$. Nevertheless, even in this case, the optimal income distribution rule still equates both utility levels and marginal utilities of income.

[38] In fact, suppose that a (CFC) invariant SWFL is Paretian, anonymous, and satisfies independence of irrelevant utilities (IIU). Then it is fairly easy to extend the arguments of Hammond (1977b) in order to show that, except in the special case when consumers have identical preferences, the indirect social welfare function can only be equity-regarding in a special case when individual demand functions are more closely related than the translog system allows.

As remarked at the end of Section 6.2, the form of the equivalence scale function $m(p;a)$ depends on differences between households that determine the relative ethical values of their incomes, as well as on differences in their demand behaviour. Thus, $m(p;a)$ cannot be inferred from demand behaviour alone.

Restricted Equivalence Scales

Jorgenson and Slesnick (1983, 1987) show that the translog model is a special case of the equivalence scale model described in Section 6.2. Indeed, under the hypothesis that the symmetric coefficient matrix $B = (\beta_{gk})$ is invertible, then given any attribute vector a, one can define the corresponding equivalence scale vector $m = (m_1, m_2, \ldots, m_\ell)$ to satisfy[39]

$$\sum_{k=1}^{\ell} \beta_{gk} \ln m_k = -\sum_{j=1}^{J} \gamma_{gj}\, a_j \quad (g = 1, 2, \ldots, \ell) \tag{6.40}$$

Next, consider the reference household for which $m_g^R = 1$ (all g), and let a^R denote the corresponding attribute vector. It follows from (6.31) that, by adding $\sum_{j=1}^{J} \gamma_{gj}\, a_j^R$ to each α_g, we can normalize so that $a^R = 0$—this is equivalent to defining α_g as the expenditure share for good g of the reference household when faced with price vector $p = (1, 1, \ldots, 1)$ and with income level 1.

With this normalization, because (6.39) implies that $m(p;0) = 1$ for all p, equation (6.37) implies that the reference household's indirect utility function is $V^R(p,y) = V^*(p,y;0) = [y/P(p)]^{D(p)}$. When the equivalence scale vector m and the attribute vector a are related by (6.40), the second aggregation condition (6.34) obviously implies that

$$\sum_{g=1}^{\ell}\sum_{k=1}^{\ell} \beta_{gk}\, \ln m_k = 0 \tag{6.41}$$

From (6.35), it follows that

$$D(m_1\, p_1, m_2\, p_2, \ldots, m_\ell\, p_\ell) = D(p)$$

[39]The following equation has a different sign from the condition given by Jorgenson and Slesnick because of the way (6.29) has been specified here.

Now (6.38) implies that

$$\begin{aligned}
& D(p)\,\ln P(m_1\,p_1, m_2\,p_2, \ldots, m_\ell\,p_\ell) \\
= \quad & \sum_{g=1}^{\ell} \alpha_g\,\ln(m_g p_g) - \frac{1}{2}\sum_{g=1}^{\ell}\sum_{k=1}^{\ell} \beta_{gk}\,\ln(m_g p_g)\,\ln(m_k p_k) \\
= \quad & D(p)\,\ln P(p) - \sum_{g=1}^{\ell}\sum_{k=1}^{\ell} \beta_{gk}\,\ln p_g\,\ln m_k - \mu(m)
\end{aligned}$$

where

$$\mu(m) := \sum_{g=1}^{\ell} \alpha_g\,\ln m_g - \frac{1}{2}\sum_{g=1}^{\ell}\sum_{k=1}^{\ell} \beta_{gk}\,\ln m_g\,\ln m_k$$

Finally, this implies that

$$\begin{aligned}
& \ln V^R(m_1\,p_1, m_2\,p_2, \ldots, m_\ell\,p_\ell, y) \\
= \quad & D(p)\ln[y/P(p)] - \sum_{j=1}^{J} \gamma_{gj}\,(\ln p_g)\,a_j - \mu(m) \\
= \quad & V^*(p, y; a) - \mu(m) \qquad (6.42)
\end{aligned}$$

It follows from Roy's identity that each household's vector demand function is indeed exactly the same as in the (restricted) equivalence scale model.

Nevertheless, equation (6.41) is an important condition that must be imposed on the domain of allowable equivalence scale vectors m. The condition is restrictive because, when combined with the first aggregation condition (6.34), together with the obvious requirement that each $m_k > 0$, equation (6.41) confines the vector m to an $\ell-1$-dimensional cone within $\mathbb{R}^{\ell}_{++}$. For example, when $\ell = 2$ this cone reduces to the 45° half-line in $\mathbb{R}^2_{++}$ on which $m_1 = m_2 > 0$. By contrast, the usual equivalence scale model allows m to be any vector in the ℓ-dimensional cone $\mathbb{R}^{\ell}_{++}$.

More seriously, comparisons of different households' utility levels are clearly affected by the presence of the term $-\mu(m)$ in (6.42). In particular, an income distribution that equates the utility measure $V^*(p, y; a)$ for households with different attribute vectors will give rise to a lower value of the alternative utility measure $V^R(m_1\,p_1, m_2\,p_2, \ldots, m_\ell\,p_\ell, y)$ for households that happen to have a higher value of $\mu(m)$. Once again, this reflects how it is ethical values, not preferences revealed by demand behaviour, which determine the interpersonally comparable utility measure. In particular, ethical values must decide whether interpersonal comparisons should be based on the indirect utility functions $V^*(p, y; a)$, or on $V^R(m_1\,p_1, m_2\,p_2, \ldots, m_\ell\,p_\ell, y)$, or on some entirely different interpersonally comparable indirect utility function $\tilde{V}(p, y; a) = \psi(V^*(p, y; a); a)$—where $\psi(V; a)$ is allowed to be any function which is strictly increasing in V for each possible attribute vector a.

6.5 An Extended Almost Ideal Demand System

The Almost Ideal Demand System

Yet another form of exact aggregation arises from the *almost ideal demand system* due to Deaton and Muellbauer (1980a, b). This is based on the family of indirect utility functions defined by

$$\ln V_i(p, y) = D(p)\ \ln[y/m_i P(p)] \tag{6.43}$$

Here, as at the end of Section 6.3, the positive constant m_i is household i's equivalence scale. Also, the income deflator $P(p)$ is defined by

$$\ln P(p) := \alpha_0 + \sum_{g=1}^{\ell} \alpha_g\ \ln p_g - \frac{1}{2} \sum_{g=1}^{\ell} \sum_{k=1}^{\ell} \beta_{gk}\ \ln p_g\ \ln p_k \tag{6.44}$$

and $D(p)$ is defined by

$$D(p) := \prod_{g=1}^{\ell} p_g^{-\delta_g} \tag{6.45}$$

The different parameters are assumed to obey the restrictions $\sum_{g=1}^{\ell} \delta_g = 0$, thus ensuring that $D(p)$ is homogeneous of degree zero, as well as $\sum_{g=1}^{\ell} \alpha_g = 1$ and $\sum_{g=1}^{\ell} \sum_{k=1}^{\ell} \beta_{gk} = 0$, thus ensuring that $P(p)$ is homogeneous of degree one. Hence, $V_i(p, y)$ is homogeneous of degree zero. Furthermore, as with the translog system, it loses no generality to assume that $\beta_{gk} = \beta_{kg}$ for all g, k. In addition, the expenditure function is given by $e_i(p, u) = m_i P(p) u^{1/D(p)}$. This implies that Lewbel's cost of characteristics index is just m_i/m^R, where m^R is the equivalence scale of the reference household. In particular, this system also satisfies Lewbel's (IB) property.

Using (6.30) again, the corresponding expenditure shares satisfy

$$\begin{aligned} w_g^i(p, y) &= -\frac{\partial \ln V^*/\partial \ln p_g}{\partial \ln V^*/\partial \ln y} \\ &= \alpha_g - \sum_{k=1}^{\ell} \beta_{gk}\ \ln p_k + \delta_g\ \ln[y/m_i\, P(p)]. \end{aligned} \tag{6.46}$$

Together, all the parameter restrictions described above evidently guarantee that the wealth shares do sum to one.

Multiplying each side of (6.46) by y_i, then adding over i, yields the aggregate expenditure equations

$$p_g\, H_g(p, y^N) = \left(\alpha_g - \sum_{k=1}^{\ell} \beta_{gk}\, \ln p_k - \delta_g\, \ln P(p)\right) Y + \delta_g \sum_{i\in N} y_i \ln(y_i/m_i)$$

Thus, aggregate demands depend only on prices, aggregate income Y, and the one additional aggregate statistic $\sum_{i\in N} y_i \ln(y_i/m_i)$.

An Affine Extension

Though Deaton and Muellbauer emphasize how their system already has a large number of parameters to estimate, nevertheless it seems to allow too little variation between different consumers, especially when compared to the translog demand system. Indeed, in the almost ideal demand system, each consumer is characterized by a single scalar m_i, whereas in the translog demand system each consumer is characterized by the vector of parameters a^i which determine the function $m(p; a^i)$. To compensate for this lack of heterogeneity, one possibility is to introduce an extra term into (6.43), as one does in going from homothetic to quasi-homothetic preferences. Specifically, instead of (6.43), assume that each consumer's indirect utility function satisfies

$$\ln V_i(p, y) = D(p)\, \ln\left(\frac{y - \gamma_i(p)}{m_i P(p)}\right) \tag{6.47}$$

where, as in the Gorman case of parallel linear Engel curves discussed in Section 6.3, it is assumed that each $\gamma_i(p)$ is an objective measure of consumer i's subsistence consumption expenditure. This should be a function that is homogeneous of degree one and also concave in p. Note that when $D(p) \equiv 1$ because each $\delta_g = 0$, then one has parallel linear Engel curves, but with a special translog price deflator.

Inverting (6.47) for each fixed p implies that consumer i's expenditure function satisfies

$$\ln[e_i(p, u) - \gamma_i(p)] = \ln[m_i P(p)] + \frac{1}{D(p)} \ln u \tag{6.48}$$

Differentiating (6.48) partially w.r.t. each p_g and then rearranging, the compensated demand functions $x^i_g(p, u)$ satisfy

$$\frac{p_g[x^i_g(p, u) - \gamma^i_g(p)]}{e_i(p, u) - \gamma_i(p)} = \frac{\partial \ln P}{\partial \ln p_g} - \frac{1}{D(p)}\, \frac{\partial \ln D}{\partial \ln p_g} \ln u$$

Substituting for $\ln u$ from (6.47), for $\partial \ln P/\partial \ln p_g$ from (6.44), and then for $\partial \ln D/\partial \ln p_g$ from (6.45), it follows that each consumer i's uncompensated demands $h_g^i(p, y_i)$ satisfy

$$\frac{p_g[h_g^i(p, y_i) - \gamma_g^i(p)]}{y_i - \gamma_i(p)} = \alpha_g - \sum_{k=1}^{\ell} \beta_{gk} \ln p_k + \delta_g \ln \left[\frac{y_i - \gamma_i(p)}{m_i P(p)} \right] \tag{6.49}$$

Multiplying each side of (6.49) by $y_i - \gamma_i(p)$ and then summing over i yields ℓ aggregate demand relations

$$\begin{aligned} p_g[H_g(p, y^N) - \Gamma'_g(p)] &= \left[\alpha_g - \sum_{k=1}^{\ell} \beta_{gk} \ln p_k - \delta_g \ln P(p) \right] [Y - \Gamma(p)] \\ &\quad + \delta_g \sum_{i \in N} [y_i - \gamma_i(p)] \ln \left[\frac{y_i - \gamma_i(p)}{m_i} \right] \end{aligned}$$

These equations allow all the parameters α_g, β_{gk}, and δ_g to be estimated, at least in principle, from sufficiently many independent observations of the aggregate demand vector H, of the price vector p, and of the two aggregate statistics $Y - \Gamma(p)$ and $\sum_{i \in N} [y_i - \gamma_i(p)] \ln([y_i - \gamma_i(p)]/m_i)$. Both the latter are derived from the empirical joint distribution of different consumers' incomes y_i, subsistence expenditures $\gamma_i(p)$, and equivalence scale parameters m_i.

An Equity-Regarding SWFL

Imposing intermediate dual comparability once again, as was done to derive (6.28), results in a social welfare function of the equity-regarding additively separable form

$$W(p, y^N) \equiv \sum_{i \in N} m_i \phi(V_i) \tag{6.50}$$

Or, more restrictively, putting $\phi(V) = \frac{1}{1-\rho} V^{1-\rho}$ implies that

$$W_\rho(p, y^N) \equiv \frac{1}{1-\rho} \sum_{i \in N} m_i \left[\frac{y_i - \gamma_i(p)}{m_i P(p)} \right]^{(1-\rho) D(p)}$$

where $\rho > 1$ in order to ensure that, for each fixed p, the function W is strictly concave in the income distribution y^N, no matter how large $D(p)$ may be; the usual condition $\rho > 0$ is insufficient.

Alternatively, inspired by Jorgenson and Slesnick's suggested use of a result due to Roberts (1980b), one could have

$$W(p, y^N) \equiv (1 - \kappa) V_* + \kappa \left[\frac{1}{M} \sum_{i \in N} m_i (V_i - V_*)^{1-\rho} \right]^{\frac{1}{1-\rho}}$$

where $V_* := \min_{i\in N} V_i$, $M := \sum_{i\in N} m_i$, and the constant parameter κ is chosen to satisfy $0 \leq \kappa \leq 1$.[40]

Thus, in this affine extension of the almost ideal demand system, our four-step procedure is carried out as follows:

(SWFL) The greater the (Paretian) indirect welfare function $W(p, y^N)$ given by $\sum_{i\in N} m_i \phi(V_i)$, as in (6.50), the better.

(CIG) Each consumer's cardinal indirect utility function is given by

$$V_i(p, y_i) = \left[\frac{y_i - \gamma_i(p)}{m_i P(p)}\right]^{D(p)}$$

where $P(p)$ and $D(p)$ are defined by (6.44) and (6.45) respectively.

(P) The observable proxies for each consumer's V_i are the equivalence scale parameter m_i and the subsistence expenditure function $\gamma_i(p)$.

(D) Data on aggregate demands, on prices, and on the joint distribution of consumers' incomes and their observable proxies can be used to estimate the common parameters α_g, β_{gk}, and δ_g. Given the particular function ϕ that was used in (6.50) to construct the SWFL, the value of W is then determined from this joint distribution of consumers' incomes and their observable proxies.

To conclude, we should repeat the *caveat* at the end of Section 6.3; the scalar parameters m_i are really welfare weights that are independent of consumer demand. Instead they need to be determined entirely by whatever ethical judgements are deemed relevant when contemplating how to make appropriate trade-offs between different households' incomes.

6.6 Assessment

All the examples in Sections 5 and 6 are intended to help reinforce the claims made in Section 5.1—namely, that steps (SWFL) and (CIG) are obviously purely normative, whereas step (OP) involves some normative judgements as well as information that can be inferred from individuals' (or households') observed behaviour. Only step (D) is purely factual. Thus, there are considerable subtleties involved in separating the normative values embodied in a social welfare objective from the relevant descriptive facts.

---

[40] See Jorgenson (1997b, pp. 66–69 and 199–200). Note that the arguments which Roberts uses to derive the form discussed by Jorgenson can easily be adapted to derive the alternative form $W \equiv (1-\kappa)V_* + \kappa\psi(\langle V_i - V_* \rangle_{i\in N})$, where $0 \leq \kappa \leq 1$, and ψ is an increasing and homogeneous of degree one function defined on $\mathbb{R}^N_+$.

7 A Generalized Utilitarianism

7.1 Introspective Approaches

The examples described so far do not include what we may call the "introspective" approaches. These rely on the idea that an ethical observer, possibly an ordinary individual, uses introspection to determine interpersonal comparisons of subjective utility, assuming that some form of extended sympathy enables this observer to imagine what it would be like to have other personal characteristics.[41] Discussions of such a possibility, and of how it is affected by scientific knowledge of the causal factors of well-being, appear in works such as Kolm (1972, 1994), Harsanyi (1976), Kaneko (1984), Hammond (1991a) and Broome (1993). Even though very interesting, this approach has not yet been incorporated in practical recipes permitting its full application in empirical studies.[42]

A related issue is that when different individuals make interpersonal comparisons, they may fail to agree.[43] As discussed by Suzumura (1983, 1996), by Roberts (1995, 1996, 1997), and by Nagahisa and Suga (1998), faced with diversity of ethical opinion, dictatorship of ethical values appears inevitable if one is going to have a (complete) social welfare ordering satisfying some form of Pareto criterion and of independence. But in this field as well, relaxing independence might be a promising way ahead, though as yet it has been little explored. After all, when society wants to compare (x, i) with (x, j), for instance, it may be relevant to take into account not only individual k's opinions of this pair, but also k's whole view about all interpersonal comparisons, because that might give some valuable information about how reasonable k's opinions are in general. These opinions may then be revealed as totally outlandish, which would justify discounting k's comparison of (x, i) with (x, j), even if this particular comparison is in tune with the opinion of a large majority. (This is the kind of discrimination that is performed, for instance, by the Condorcet criterion discussed in Section 5.8.) It would also make sense to look at the process by which k's opinions were formed, which would require a richer informational basis. Giving this important topic the attention it deserves, however, would take us too far away from the main topic of this chapter, and far beyond utility theory in general.

[41] Samuelson (1947, p. 91) applies the adjective "introspective" to the "concept of utility as a sensation". In other words, it is a person's *own* utility, resulting from that person's introspection. Instead, we are considering an external ethical observer's estimate of that person's utility, based on the observer's own process of introspection.

[42] At least, it has not unless one regards it as merely imposing more ethical structure on the capabilities approach described in Section 5.2.

[43] Actually, even when all do agree, still they may all be wrong. See Narens and Luce (1983).

This section will instead expound a theory of social choice, or ethical decision-making, that elaborates Hammond (1987b, 1991a, 1996). This theory encompasses the main features of the introspective approach, as well as some others—possibly including capabilities. It focuses on some formal requirements that one may want to impose on the SWFL when risk or uncertainty have to be taken into account. It also derives the various kinds of interpersonal comparison that may be implied by such requirements. In other words, instead of attempting to provide a full-blown theory of individual good and of the empirical basis of related interpersonal comparisons, as in the examples discussed so far, the idea is to derive the kinds of interpersonal comparison that inevitably have to be made, granted some basic ethical and decision-theoretic principles which determine the form of the SWFL. This project was inspired by and is quite close in spirit to Harsanyi (1955), although the framework and conclusions are more general, and closer to Broome's theory (1991) in particular. In addition, we will explain how the various kinds of interpersonal comparison relate to specific social decisions. This relationship is similar to, although possibly more sophisticated than, the equivalence between propositions [IC] and [SD] in the example of section 5.1 above.

7.2 Social and Personal Consequences

The objectively expected utility functions of Chapter 5, and the "consequentialist" normative arguments that were used to justify them, will now be applied to social decision problems. The result will be a form of utilitarianism that allows interpersonal comparisons to be interpreted as preferences for different personal characteristics, regardless of who may possess them.

First, given any $i \in N$, let X_i denote a copy of the set X whose members x_i are to be interpreted as i's *personalized social states*. As in the theory of public goods (Foley, 1970, p. 70; Milleron, 1972; etc.), it helps to imagine that we could somehow choose different social states $x_i \neq x_j$ for different individuals i and j, even though this may well be impossible in practice. Think how many social conflicts could be avoided if only everybody could be allowed to choose their own favourite social state! But the requirement that $x_i = x_j$ for all $i, j \in N$ can be imposed on the decision problem at a later stage.

In addition to social states in the conventional sense, it will be convenient to consider also for each $i \in N$ a space of *personal characteristics* $\theta_i \in \Theta_i$. Such characteristics determine *i's* preferences, interests, talents, and everything else (apart from the social state) which is ethically relevant in determining the welfare of the specific individual i. In Section 7.6, θ_i will even indicate whether or not individual i ever comes into existence.

For each individual $i \in N$, a *personal consequence* is a pair $z_i = (x_i, \theta_i)$ in the Cartesian product set $Z_i := X_i \times \Theta_i$ of personalized social states x_i and personal characteristics θ_i. Then, in a society whose membership N is fixed, a typical *social consequence* consists of a profile $z^N = (z_i)_{i \in N} \in Z^N := \prod_{i \in N} Z_i$ of such personal consequences—one for each individual member of society (both actual and potential). The consequence domain $Y = Z^N$ will consist of all such social consequences, with typical member $y = z^N$.

The theory of expected utility that was expounded and motivated in Chapter 5 of this *Handbook* can now be applied to the class of all decision problems with consequences in Z^N. The implication is the existence of a unique cardinal equivalence class of von Neumann–Morgenstern *social welfare functions* $w(y) \equiv w(z^N)$, defined on the space of social consequences, whose expected value should be maximized in every (finite) social decision problem. The only difference is that the consequence domain consists of social consequences. What is most important, however, is the idea that each personal consequence $z_i \in Z_i$ captures everything of ethical relevance to individual i. By definition, nothing else, not even some other individual's personal consequence, can possibly be relevant to i's welfare.

Diamond (1967) criticized Harsanyi for requiring that "social choice satisfies the axioms for expected utility maximization". In the famous example he proposed, there is a two-person society with $N = \{1, 2\}$, and with $\{0, 1\}$ as the common domain of personal consequences. It is assumed that both individuals benefit more from the personal consequence 1 than they do from the personal consequence 0. To take Broome's (1991, Section 5.7) version of this example, there is one kidney available for transplant, which both individuals need to survive. The von Neumann–Morgenstern social welfare function w is defined on the domain $\{0, 1\} \times \{0, 1\}$, and is assumed to satisfy the symmetry condition $w(1, 0) = w(0, 1)$. But then $w(1, 0) = \frac{1}{2} w(1, 0) + \frac{1}{2} w(0, 1)$, so there is social indifference between the sure outcome $(1, 0)$ and the even chance lottery $\frac{1}{2} \circ (1, 0) + \frac{1}{2} \circ (0, 1)$. This is true even though the even chance lottery seems clearly fairer than the first option of letting individual 1 enjoy the better personal consequence for sure, with individual 2 condemned to the personal consequence 0. As Broome carefully discusses, such criticism can be considerably blunted by assuming, as explained above, that personal consequences contain all relevant features of the situation, including fairness in the choice process that leads to the final outcome.

7.3 Individualistic Consequentialism

A general random social consequence is some joint probability distribution $\lambda \in \Delta(Z^N)$ over the product space Z^N of different individuals' personal consequences. Such personal consequences could be correlated between different individuals, or they could be independent. The extent of this correlation should be of no consequence to any individual, however. For, provided that everything relevant to individual $i \in N$ really has been incorporated in each personal consequence $z_i \in Z_i$, all that really should matter to i is the marginal distribution $\lambda_i \in \Delta(Z_i)$ of i's own consequences. This leads to the *individualistic consequentialism* hypothesis requiring any two lotteries $\lambda, \mu \in \Delta(Z^N)$ to be regarded as equivalent random consequences whenever, for every individual $i \in N$, the marginal distributions $\lambda_i = \mu_i \in \Delta(Z_i)$ of i's consequences are precisely the same. This means in particular that

$$\lambda_i = \mu_i \text{ (all } i \in N) \implies \mathbb{E}_\lambda\, w(z^N) = \mathbb{E}_\mu\, w(z^N)$$

—i.e., λ and μ must be indifferent according to the relevant expected utility criterion whenever the personal marginal distributions are all equal.

Succinctly stated, individualistic consequentialism amounts to requiring that only each individual's probability distribution of personal consequences be relevant when evaluating any social probability distribution. There is no reason to take account of any possible correlation between different individuals' personal consequences. From now on, therefore, individualistic consequentialism allows us to regard any lottery in $\Delta(Z^N)$ as adequately described by the profile $\lambda^N = \langle \lambda_i \rangle_{i \in N} \in \prod_{i \in N} \Delta(Z_i)$ of individuals' marginal distributions λ_i. That is, we identify $\Delta(Z^N)$ with $\prod_{i \in N} \Delta(Z_i)$.

For an ordinary description of personal consequences, this would certainly be a controversial claim, as discussed in detail by Broome (1991a, Section 9.3) and Broome (1991b, pp. 83–4). In the two-person society at the end of Section 7.2, for example, it requires society to be indifferent between the even-chance lottery $\frac{1}{2} \circ (1,1) + \frac{1}{2} \circ (0,0)$ and the alternative $\frac{1}{2} \circ (1,0) + \frac{1}{2} \circ (0,1)$, even though the first lottery guarantees that the outcome is egalitarian ex post, whereas the second guarantees extreme inequality. More generally, an egalitarian planner might prefer more egalitarian outcomes—that is, a positive correlation between levels of individual good so as to avoid situations with a large gap between winners and losers, when this makes no difference to individuals' *ex ante* prospects, or even at a small cost to individuals' *ex ante* well-being. As Broome notes, however, it is possible for social preferences that favour egalitarian outcomes to be incorporated into the measure of individual good.

Thus, when all ethically relevant social concerns are taken into account in the description of each individual prospect, individualistic consequentialism becomes innocuous. The counterpart of this is that the ethical content of the measure of individual good becomes disturbingly rich. But we have indeed assumed above that everything of ethical relevance has been included in z_i already.

7.4 Individual Welfarism

Consider any decision problem having the special property that there is only one individual $i \in N$ whose distribution of personal consequences is affected by any feasible decision. Hence, there must be a profile $\bar{\lambda}_{-i} \in \prod_{h \in N \setminus \{i\}} \Delta(Z_h)$ of fixed lotteries $\bar{\lambda}_h \in \Delta(Z_h)$ $(h \in N \setminus \{i\})$ for all other individuals, as well as a set $F_i \subset \Delta(Z_i)$ of feasible lotteries over i's personal consequences, such that the feasible set of lotteries is $F_i \times \{\, \bar{\lambda}_{-i} \,\} \subset \Delta(Z^N)$. A decision problem with this property will be called *individualistic*, or a *one-person situation*.

The second individualistic axiom which we shall use is *individual welfarism*. This requires that for each $i \in N$ there is a unique cardinal equivalence class of *individual welfare functions* $w_i(z_i)$ with the property that, in any individualistic decision problem having $F_i \times \{\, \bar{\lambda}_{-i} \,\} \subset \Delta(Z^N)$ as the feasible set of lotteries, the social decision should maximize the expected value $\mathbb{E}_{\lambda_i} w_i(z_i)$ of w_i w.r.t. λ_i over the set $F_i \subset \Delta(Z_i)$ of feasible probability distributions over i's personal consequences. In particular, the social decision should be independent of $\bar{\lambda}_{-i}$.

This last independence property is the key hypothesis here. The motivation is that, if only consequences to i are affected by any decision, the fixed consequences to all other individuals are ethically irrelevant—assuming, as required by individualistic consequentialism, that everything relevant to ethical decision making is already included in the consequences, and that only (distributions over) personal consequences matter.

Thus, whenever there is "no choice" in the personal consequences of all other individuals, the social objective becomes identical to the only affected individual's welfare objective. Note especially that individual welfarism poses no restrictions on what is allowed to count as part of a personal consequence and so to affect each individual's welfare. All it says is that, in "one-person situations", social welfare is effectively identified with that one person's individual welfare.

7.5 Utilitarianism

Individual welfarism has a much more powerful implication, however, when it is combined with individualistic consequentialism as defined in Section 7.3. To see this, define the expected utility functions $U : \Delta(Z^N) \to \mathbb{R}$ and $U_i : \Delta(Z^i) \to \mathbb{R}$ by $U(\lambda^N) := \mathbb{E}_{\lambda^N} w(z^N)$ and $U_i(\lambda_i) := \mathbb{E}_{\lambda_i} w_i(z_i)$ $(i \in N)$ respectively. Now fix any profile $\bar{\lambda}^N \in \Delta(Z^N)$. As before, let n denote the number of individuals in the set N. Following an argument due to Fishburn (1970, p. 176), note that for all $\lambda^N \in \Delta(Z^N)$ one has the equality

$$\sum_{i \in N} \frac{1}{n} \, (\lambda_i, \bar{\lambda}_{-i}) = \frac{n-1}{n} \, \bar{\lambda}^N + \frac{1}{n} \, \lambda^N$$

between the two probability mixtures on each side of the equation, and so between the expected utilities of these two mixtures. Because the expected utility function U must preserve such probability mixtures, it follows that

$$\frac{1}{n}\sum_{i\in N} U(\lambda_i, \bar{\lambda}_{-i}) = \frac{n-1}{n}\, U(\bar{\lambda}^N) + \frac{1}{n}\, U(\lambda^N)$$

Therefore

$$U(\lambda^N) = \sum_{i\in N} U(\lambda_i, \bar{\lambda}_{-i}) - (n-1)\, U(\bar{\lambda}^N)$$

But individual welfarism implies that $U(\lambda_i, \bar{\lambda}_{-i})$ and $U_i(\lambda_i)$ must be cardinally equivalent functions of λ_i. So, for each $i \in N$, there exist real constants α_i and β_i, with $\beta_i > 0$, such that

$$U(\lambda_i, \bar{\lambda}_{-i}) = \alpha_i + \beta_i\, U_i(\lambda_i)$$

for all $\lambda_i \in \Delta(Z^i)$. Therefore

$$U(\lambda^N) = \sum_{i\in N} [\alpha_i + \beta_i\, U_i(\lambda_i)] - (n-1)\, U(\bar{\lambda}^N) = \bar{\alpha} + \sum_{i\in N} \beta_i\, U_i(\lambda_i)$$

where $\bar{\alpha} := \sum_{i\in N} \alpha_i - (n-1)\, U(\bar{\lambda}^N)$. Hence, there must exist an additive constant $\bar{\alpha}$ and a set of positive multiplicative constants β_i $(i \in N)$ such that

$$w(z^N) \equiv \bar{\alpha} + \sum_{i\in N} \beta_i\, w_i(z_i)$$

Then, however, since the individual and social welfare functions are only unique up to a cardinal equivalence class, for each $i \in N$ we can replace the individual welfare function $w_i(z_i)$ by the cardinally equivalent function $\tilde{w}(z_i) := \beta_i\, w_i(z_i)$, and the social welfare function $w(z^N)$ by the cardinally equivalent function $\tilde{w}(z^N) := w(z^N) - \bar{\alpha}$. The result is that

$$\tilde{w}(z^N) = w(z^N) - \bar{\alpha} = \sum_{i\in N} \beta_i\, w_i(z_i) = \sum_{i\in N} \tilde{w}_i(z_i)$$

This takes us back to the simple addition of individual "utilities", as in classical utilitarianism, once these utilities have all been suitably normalized. Because of this possible normalization, we shall assume in the future that

$$w(z^N) \equiv \sum_{i\in N} w_i(z_i).$$

Note, however, that these utility functions are by no means the same as those in other more traditional versions of utilitarianism discussed in previous sections. They are merely representations of appropriate ethical social decisions in decision problems affecting just one individual. There need not be any relationship to classical or other concepts of utility such as happiness, pleasure, absence of pain, preference satisfaction, etc. Indeed, the functions should probably be thought of more as indicators of individual ethical value (to the social planner) rather than as any measure of individual utility or even welfare. It follows that this approach is compatible with many different conceptions of individual good—perhaps even with all "monistic" conceptions. It therefore encompasses many ethical theories. This is a major difference from Harsanyi's (1955) utilitarian theory.[44] On the other hand, the additive structure of that theory is preserved, as its use of the expected utility criterion to choose among lotteries.

7.6 Personal Non-Existence

So far the set of individuals N has been treated as fixed. Yet many ethical issues surround decisions affecting the size of future generations, as well as the precise characteristics of those individuals who will come into existence. That is, both the number and the composition of the set N are of great ethical significance. Thus, it would seem that N itself should be treated as variable consequence along with z^N, as indeed it was in Hammond (1988). For some of the most recent work on the ethics of variable population, see Blackorby, Bossert and Donaldson (1995, 1996, 1997a, b, 1998) and several other articles by the same authors.

A simpler alternative to the arguments in these papers, however, is to treat "non-existence" for any individual $i \in N$ as a particular personal characteristic $\theta_i^0 \in \Theta_i$ which i could have, and then to define N as the set of all potential rather than actual individuals. In this way, N is partitioned into the two sets $N^* := \{ i \in N \mid \theta_i \neq \theta_i^0 \}$ of actual individuals who do come into existence, and $N^0 := \{ i \in N \mid \theta_i = \theta_i^0 \}$ of individuals whose potential existence remains unrealized.

Actually, not much generality is lost by doing this, for the following reason. Assuming that only a finite number of individuals can ever be born before the world comes to an end (as seems quite reasonable, despite many economists' fondness for models of steady state growth, etc.), one can regard each identifier $i \in N$ as just an integer used to number all the individuals who come into existence, more or less in the temporal order of their birth. Everything that is really relevant about an individual i, including date of birth, can be included in i's personal characteristic θ_i. Accordingly, every individual who is ever born

[44] Harsanyi's approach has been hotly debated. See Weymark (1991) and Mongin and d'Aspremont (1998) for syntheses.

certainly gets numbered. Also, unless all the maximum possible number of individuals does actually come into existence, there will be "unused" numbers which refer to potential rather than actual individuals.

For those individuals $i \in N^0$ who never come into existence, the concept of individual welfare hardly makes any sense. In decision-theoretic terms, this means that non-existent individuals are not affected by social decisions—all social decisions are the same to them (except for decisions giving rise to a positive probability of their coming into existence, of course). Consider now, for any $i \in N$, an individualistic decision problem whose feasible set F_i has the property that $\lambda_i \in F_i$ only if $\lambda_i(X_i \times \{\theta_i^0\}) = 1$—i.e., the probability of i not existing is always 1, no matter what decision is taken. Since all consequences in F_i are the same to this almost surely non-existent individual, this suggests that all social decisions with consequences in F_i are equally ethically appropriate from the point of view of individual i alone. This suggestion motivates the assumption that, for some constant w_i^0, individual i's welfare function $w_i(z_i)$ should satisfy $w_i(x_i, \theta_i^0) = w_i^0$ for all $x_i \in X_i$. Thus, w_i^0 can be regarded as the constant "welfare of non-existence", which is entirely independent of the social state or any aspect of any social consequence in which i never exists.[45]

After making this assumption, one additional useful normalization of individuals' welfare functions is possible. Replace each $w_i(z_i)$ by the function

$$\tilde{w}_i(z_i) := w_i(z_i) - w_i^0$$

This function is cardinally equivalent because a constant has merely been subtracted. Then, of course, $\tilde{w}_i(x_i, \theta_i^0) = 0$ for all $x_i \in X_i$, and so $\tilde{w}_i(z_i) = 0$ whenever $i \in N^0$. Similarly, replace $w(z^N) \equiv \sum_{i \in N} w_i(z_i)$ by the cardinally equivalent function

$$\tilde{w}(z^N) := w(z^N) - \sum_{i \in N} w_i^0$$

Then, however,

$$\tilde{w}(z^N) \equiv \sum_{i \in N} [w_i(z_i) - w_i^0] \equiv \sum_{i \in N} \tilde{w}_i(z_i) \equiv \sum_{i \in N^*} \tilde{w}_i(z_i)$$

where N^* is the set of individuals who ever come into existence. So only individuals in the set N^* need be considered when adding all individuals' welfare levels.

[45] Blackorby *et al.*, in the works cited previously, prefer to call w_i^0 the *critical level* of i's utility; for them, a life has zero utility, by definition, when the individual is no better or worse off than by never having been born.

Once again, it will be assumed from now on that this normalization has been carried out. Because $N^* = \{ i \in N \mid \theta_i \neq \theta_i^0 \}$, it follows that

$$w(x^N, \theta^N) = w(z^N) \equiv \sum_{i \in N^*} w_i(z_i) = \sum_{i \in N^*} w_i(x_i, \theta_i)$$

Maximizing this social objective is formally identical to classical utilitarianism. But as already pointed out above, the resemblance is only formal because the individual welfare functions $w_i(z_i)$ mean something quite different. In particular, the zero level of this function is, by its very construction, just the minimum level of individual welfare at which it is ethically appropriate to cause the individual to come into existence.[46] This does much to dilute the strength of Parfit's (1984) "repugnant conclusion", which is that classical utilitarianism recommends creating very many extra individuals who are barely able to live above a subsistence level set so low that anyone who was forced to live below it would prefer not to have been born at all. Here we can escape the repugnant conclusion because there is nothing to prevent the ethical values embodied in the normalized individual welfare function $w_i(z_i)$ from making w_i positive only if individual i would actually be quite well off if allowed to come into existence. The fact that the personal consequence z_i makes individual i glad to be alive is not by itself sufficient to make $w_i(z_i)$ positive, though many might argue that is a necessary condition.

Note too that having $w_i(z_i)$ positive would only be a sufficient condition on its own for wanting i to exist if i's existence could somehow be brought about without interfering with anybody else. Yet children cannot exist without having (or having had) parents. So the personal benefits (or costs) to i of coming into existence have to be weighed against any costs and benefits to other individuals, especially i's parents, etc. Some further discussion of such issues occurs in Hammond (1988).

7.7 Revealed Interpersonal Comparisons

Equipped with these social preferences, we are now in a position to see how interpersonal comparisons of utility relate to concrete social decisions consis-

[46]A similar construction is used by Dasgupta (1993, ch. 13), who also provides a much more thorough philosophical discussion. The zero level in his approach, as well as in that outlined above, corresponds to the "critical level" considered by Blackorby *et al.* In this connection, Blackorby, Bossert and Donaldson (1998, p. 17) are somewhat misleading when they claim that the approach presented here uses "individual 'preferences' that cover states in which the person does not exist"—although this may not be entirely clear from the paper Hammond (1988) which they cite. In fact, the approach presented here uses ethical *social* preferences throughout, even for decisions affecting only one individual. So this treatment of non-existence is only one of many important ways in which social preferences differ from individual preferences.

tent with such preferences. As pointed out in Hammond (1991a), there are interpersonal comparisons embodied in the social welfare function $w(z^N) = \sum_{i \in N^*} w_i(z_i)$, and simply by looking at some specific social decisions, one may be able to deduce what interpersonal comparisons of w_i are implied. Indeed, the level comparison $w_h(z_h) > w_i(z_i)$ means that society is better off creating individual h with personal consequence z_h rather than individual i with personal consequence z_i. Furthermore, the difference comparison $w_h(z_h) - w_h(z'_h) < w_i(z'_i) - w_i(z_i)$, which is of course equivalent to $w_h(z_h) + w_i(z_i) < w_h(z'_h) + w_i(z'_i)$, really does mean that moving h from z_h to z'_h and i from z_i to z'_i produces a benefit to society (if nobody else is affected). If there is a loss to h, this must be outweighed by the gain to i. Alternatively, if there is a loss to i, this must be outweighed by the gain to h.

Actually, even welfare ratios acquire meaning. For $w_h(z_h)/w_i(z_i)$ can be regarded as the marginal rate of substitution between individuals like h facing personal consequence z_h and individuals like i facing personal consequence z_i. If this ratio is greater than 1, for instance, then society could gain by creating more individuals like h and fewer like i. And if $w_h(z_h)/w_i(z_i) = 10$, this means that society should be indifferent between creating 10 extra individuals like i and one extra individual like h. Thus, the claim that a Brahmin has 10 times the utility (or welfare) of an Untouchable does have meaning, even if most of us would regard the kinds of decision implied by such a claim as highly unethical and obnoxious.[47]

So we have a "cardinal ratio scale" measure of individual welfare, with "cardinal full comparability" of both welfare levels and differences, as well as a clearly defined zero level of welfare. Yet, according to the theory expounded above, of all the SWFLs considered by Roberts (1980b) which have this property, only the simple sum is ethically appropriate. The social welfare functional is no longer left indeterminate, therefore, as usually happens in the SWFL approach to social choice theory.

Of course, this extra determinacy of the functional form comes at a high price, since now all the indeterminacy has been displaced into the individual welfare function, which has been left unspecified here. In Sen's version of social choice theory with interpersonal comparisons, as well as in Harsanyi's version of utilitarianism, the measure of individual's well-being (capability, or utility)

---

[47]Robbins (1938, p. 636) attributes to Sir Henry Maine a story of a Brahmin who, upon meeting a Benthamite, was moved to say: "I am ten times as capable of happiness as that untouchable over there." See also Sen (1973, pp. 81–2). The Brahmin's statement appears extremely obnoxious, but actually is not immediately relevant to any social decision, except insofar as it was addressed to a Benthamite. After all, the statement is about the capacity for happiness rather than about any ethical measure of individual welfare; though Benthamites, by definition, confuse the two, there is no reason for anybody else to do so.

has a rather precise content, and, although laden with ethical values (the selection and weighing of functionings in Sen's theory, the laundering of antisocial features of individual preferences in Harsanyi's approach), it is assumed to be given when the construction of the social criterion is envisioned. Here instead, on the basis of minimal ethical principles, we have focused on the mathematical structure of the social criterion, and derived from it the need for an individual measure of well-being. This leaves for later the discussion of all other relevant ethical values needed to construct this measure.

From this last subsection one can see that observing all the social planner's decisions would be enough to deduce the underlying welfare measures, if not the underlying value judgements (because different value judgements might yield the same measure). But implementing such a "revealed preference" approach does require a perfect social planner to be available. If such were the case social choice theory would become—perhaps fortunately—an entirely pointless exercise. In the absence of such a guide, additional hard ethical issues must be faced squarely.

In other words, this section has considered only the case when there is a single interpersonal ordering, or when ethical decisions are made by some kind of benevolent ethical dictator. It does not consider what is implied by the divergence of ethical opinions that seems inevitable in any real human society, notwithstanding the arguments of Harsanyi (1955) and others. Indeed, suppose that all individuals subscribe to the above theory, but have divergent ethical values. Then they will have different cardinal equivalence classes of the "von Neumann–Morgenstern ethical value functions" whose expected value they think it is right to maximize. Now, this is exactly the setting for Sen's (1970a, Theorem 8*2, pp. 129–30) cardinal extension of Arrow's impossibility theorem. Moreover, as shown by Bordes, Hammond and Le Breton (forthcoming), there is little hope that one can escape the need to dictate ethical values even if one restricts the domain of admissible profiles of different individuals' opinions concerning what the ethical value function should be. Unless, that is, one relaxes the independence conditions, or else admits interpersonal comparisons of ethical values in a way that allows some weighted average of different individuals' versions of the ethical value function to be constructed. But in both cases some hard ethical choices will ultimately have to be made anyway, either by consensus, or dictatorially.

8 Concluding Remarks

Sen (1970a) pioneered the social welfare functional approach to social choice theory, which many others have followed during the ensuing three decades. This approach allows the social welfare ordering to depend on broader information than the profiles of individual orderings that form the domain of an Arrow social welfare function. In particular, a social welfare functional could accommodate interpersonal comparisons of utility. This approach was very useful in pointing out how the iron grip of Arrow's "dictatorship theorem" could be relaxed provided that one admitted interpersonal comparisons, thereby allowing Arrow's independence of irrelevant alternatives condition to be replaced by some form of "independence of irrelevant interpersonal comparisons", as in Hammond (1991b).

In retrospect, however, the social welfare functional approach can now be seen as having several quite serious defects. One is the failure to explain how interpersonal comparisons of utility are to be interpreted, since the informational basis was assumed to be exogenously given. A second arises once we know what interpersonal comparisons mean, and any ethically appropriate interpersonal ordering (or corresponding invariance class of utility transformations) has been specified. For example, suppose interpersonal comparisons have the same interpretation as in Section 7, and so give rise to a unique corresponding invariance class of common ratio scale (CRS) measurable utilities. As Roberts (1980b) in particular makes clear, this still leaves scope for an enormous variety of different SWFLs. Thus, even with interpersonal comparisons that are this complete, the SWFL approach is still far away from determining an unambiguous procedure for embodying such comparisons in the social welfare functional that generates the social welfare ordering.

In fact, a profile of individual utility functions is typically determined only up to an invariance class that contains many functions representing the same preferences. For this reason, interpersonal comparisons of utility, thought of as comparisons of different individuals' uniquely specified utilities, make little sense until put into a more appropriate framework. For instance, it seems better to rely directly on the interpersonal ordering $\tilde{R}$ on $X \times N$ considered in Section 4, instead of deriving it from utility functions as was done there. The ordering $\tilde{R}$, when represented by a single interpersonal utility function $\tilde{U}$ on $X \times N$, gives meaning to comparisons between utility levels $\tilde{U}(x, i)$ and $\tilde{U}(y, j)$ for any pair (x, i) and (y, j) in $X \times N$. A second alternative considers the interpersonal ordering $\tilde{R}$ on the set $\Delta(X \times N)$ of simple lotteries over $X \times N$ which, if it can be represented by the expected value of each von Neumann–Morgenstern utility function in a cardinal equivalence class, gives meaning to comparisons between utility differences, and even to comparisons between ratios of utility differences. In this way, the fundamental concept becomes the interpersonal ordering that is represented by an interpersonal utility function which happens to give meaning to interpersonal comparisons of utility, rather than starting out with different individuals' utility functions which one then tries to compare.

Though this direct use of an interpersonal ordering seems a definite improvement, it is formally equivalent, and still leaves us with the question of what this ordering is meant to represent, and how it and the interpersonal comparisons it implies should be reflected in social preferences. Rather than face these questions directly, Section 7 attempted to lay out the details of a comprehensive ethical decision theory, based on consequentialist principles of the kind that were discussed in Chapters 5 and 6 of Volume 1 of this *Handbook*. This leads to a form of utilitarianism requiring the maximization of the expected sum of individual utility functions that all lie within one common cardinal equivalence class. Indeed, by considering what variations might be desirable in the set of individuals who come into existence, these utilities can be given a meaningful zero level and determined up to a common ratio scale. The utility functions constructed in this way, however, reflect each person's relative ethical value, rather than what most social choice and decision theorists have generally understood as their utility. Individual preferences and individual values are important considerations affecting the measure of an individual's ethical value, but they do not determine it uniquely; other considerations that are relevant to ethical decision-making also have to be included, and sometimes even allowed to predominate. Even the way in which an individual's ethical value does depend on that individual's preferences, or on any utility function which represents those preferences, may be quite indirect or convoluted.

Suppose, for instance, the relevant information is restricted to the sphere of individual preferences, in the way that is familiar to most economists. Even then, we have seen that it is possible in principle to define individual ethical value without having to decipher private mental states, simply by allowing the preference profile to be incorporated into the social preference ordering in a broader way than any allowed by Arrow's independence condition. In terms of the theory proposed in Section 7, this simply amounts to letting each personal characteristic θ_i include any ethically relevant description of the whole individual preference relation (amongst other things). Obviously, this entails violating Arrow's independence condition with respect to ordinary preferences. Nevertheless, as Samuelson (1987) and others have advocated, there is every reason to follow this path. While much of the literature after Sen (1970a) has explored the possibilities opened up by shedding the Ordinal Non-Comparability straitjacket, it turns out that no less interesting possibilities are permitted by relaxing (IIU). The various examples provided by the Nash SWFL, by relative utilitarianism, by fairness criteria, and by voting rules have amply shown that there are some important social choice contexts in which individual preferences alone may provide a sufficient informational basis to bypass Arrow's impossibility result.

Whether this informational basis is sufficient to allow for all *relevant* aspects of individual situations, however, remains to be discussed. After all, Arrow's theorem involves only weak ethical requirements, and satisfying most of them does not automatically guarantee that the social preferences under consideration are attractive. It is not practical convenience, but ethical reasoning, which should decide whether it is important to make social decisions depend on, say, the distribution of "levels" of subjective satisfaction in the population.

Another important lesson emerges from these various explorations that go beyond (IIU), especially when one compares the Condorcet criterion of the voting model and the Walrasian SWFL of the fair division model. This is that the very description of each social state, when it includes more structure and detail, can suggest a suitable basis for expressing reasonably sophisticated value judgements. By contrast, the more abstractly each state is described, the less rich are the ethical values that can be expressed and discussed in relation to the social choice problem. For example, it is impossible even to talk about equal shares of resources within an abstract framework where one egalitarian allocation is simply called x, while a second extremely inegalitarian one hides behind the name y. Thus, as one confronts the difficult problem of constructing an ethically appropriate measure of individual good, a more concrete description of social life may suggest how to start filling in the many blanks that Section 7 leaves wide open. Such an extended description would at least lay bare the hard issues about what really matters when evaluating and comparing different individuals' fates, and how these fates will be affected by policy decisions. By contrast, the abstract model is inadequate even for formulating the questions related to these hard issues, let alone for finding the answers. Of course, specifying each social state in more detail creates extra possibilities for describing the ethically relevant aspects of different individuals' personal situations. This in turn provides a richer information basis for making whatever interpersonal comparisons are ethically relevant.

Acknowledgments

Sections 2 and 7 of this chapter in particular are based on two lectures delivered by Peter Hammond at the Institut International de Paris La Défense during June, 1995. His thanks to Alain Trannoy and Serge Kolm for inviting him to give those lectures, and for the interesting discussion with them and other members of the audience that ensued. Section 7 is adapted from parts of Hammond (1996). An earlier incomplete draft of this chapter was prepared during Hammond's visit in Autumn 1998 to the Institute of Economic Research at Hitotsubashi University, sponsored by the Japanese Ministry of Education. He expresses his gratitude to the members of that institute, and especially to his host, Kotaro Suzumura.

For fruitful discussions with at least one co-author, sometimes many years ago, both of us would also like to thank Chuck Blackorby, Walter Bossert, David Donaldson, Louis Gevers, Dale Jorgenson, Mamoru Kaneko, Michel Le Breton, Arthur Lewbel, François Maniquet, Martin Rechenauer, John Roemer, Yves Sprumont, Kotaro Suzumura, Koichi Tadenuma, William Thomson, and John Weymark, as well as Serge Kolm for his comments on an earlier draft. Finally, our thanks to Jacqueline Seignan for invaluable bibliographic research assistance.

References

Arneson, R. (1989). Equality and Equal Opportunity for Welfare. *Philosophical Studies*, 56:77–93.

Arneson, R. (1990). Liberalism, Distributive Subjectivism, and Equal Opportunity for Welfare. *Philosophy and Public Affairs*, 19:158–194.

Arrow, K. J. (1950). A Difficulty in the Concept of Social Welfare. *Journal of Political Economy*, 58:328–346. Reprinted in Arrow (1983), chapter 1, pp. 1–29.

Arrow, K. J. (1951, 1963). *Social Choice and Individual Values*. Yale University Press, New Haven, 2 edition.

Arrow, K. J. (1971). A Utilitarian Approach to the Concept of Equality in Public Expenditure. *Quarterly Journal of Economics*, 85:409–415.

Arrow, K. J. (1977). Extended Sympathy and the Possibility of Social Choice. *American Economic Review (Papers and Proceedings)*, 67:219–225. Reprinted in Arrow (1983), chapter 11, pp. 147–161.

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

Arrow, K. J., Chenery, H. B., Minhas, B. S., and Solow, R. M. (1961). Capital-Labor Substitution and Economic Efficiency. *Review of Economics and Statistics*, 43:225–250.

Atkinson, A. B. (1970). On the Measurement of Inequality. *Journal of Economic Theory*, 2:244–263.

Atkinson, A. B. (1995). *Public Economics in Action. The Basic Income–Flat Tax Proposal*. Oxford University Press, Oxford.

Bandyopadhyay, T. (1988). Revealed Preference Theory, Ordering and the Axiom of Sequential Path Independence. *Review of Economic Studies*, 55:343–351.

Barone, E. (1908). Il ministero della produzione nello stato colletivista. *Giornale degli Economisti*, 37:267–293 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).

Barten, A. P. (1964). Family Composition, Price, and Expenditure Patterns. In Hart, P., Mills, G., and Whitaker, J. K., editors, *Econometric Analysis for National Economic Planning (16th Symposium of the Colston Society)*, pages 277–292. Butterworth, London.

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

Berg, M. and Lepelley, D. (1994). On Probability Models in Voting Theory. *Statistica Neerlandica*, 48:133–146.

Bergson, A. (1954). On the Concept of Social Welfare. *The Quarterly Journal of Economics*, 68:233–252.

Black, D. (1948). On the Rationale of Group Decision Making. *Journal of Political Economy*, 56:23–34.

Blackorby, C., Bossert, W., and Donaldson, D. (1995). Intertemporal Population Ethics: Critical-Level Utilitarian Principles. *Econometrica*, 63:1303–1320.

Blackorby, C., Bossert, W., and Donaldson, D. (1996). Intertemporally Consistent Population Ethics: Classical Utilitarian Principles. In Arrow, K. J., Sen, A. K., and Suzumura, K., editors, *Social Choice Re-Examined*, volume 2, chapter 13, pages 137–162. Macmillan, London. IEA Conference Volume, no. 117.

Blackorby, C., Bossert, W., and Donaldson, D. (1997a). Birth-Date Dependent Population Ethics: Critical-Level Principles. *Journal of Economic Theory*, 77:260–284.

Blackorby, C., Bossert, W., and Donaldson, D. (1997b). Critical-Level Utilitarianism and the Population-Ethics Dilemma. *Economics and Philosophy*, 13:197–230.

Blackorby, C., Bossert, W., and Donaldson, D. (1998). Uncertainty and Critical-Level Population Principles. *Journal of Population Economics*, 11:1–20.

Blackorby, C. and Donaldson, D. (1988). Cash versus Kind, Self-Selection, and Efficient Transfers. *American Economic Review*, 78:691–700.

Blackorby, C. and Donaldson, D. (1990). A Review Article: The Case Against the Use of the Sum of Compensating Variations in Cost-Benefit Analysis. *Canadian Journal of Economics*, 23:471–494.

Blackorby, C. and Donaldson, D. (1991). Adult-Equivalence Scales, Interpersonal Comparisons of Well-Being, and Applied Welfare Economics. In Elster, J. and Roemer, J. E., editors, *Interpersonal Comparisons of Well-Being*, chapter 6, pages 164–199. Cambridge University Press, Cambridge.

Blackorby, C. and Donaldson, D. (1993a). Adult-Equivalence Scales and the Economic Implementation of Interpersonal Comparisons of Well-Being. *Social Choice and Welfare*, 10:335–361.

Blackorby, C. and Donaldson, D. (1993b). Household Equivalence Scales and Welfare Comparisons: A Comment. *Journal of Public Economics*, 50:143–146.

Blackorby, C., Donaldson, D., and Weymark, J. (1984). Social Choice with Interpersonal Utility Comparisons: A Diagrammatic Introduction. *International Economic Review*, 25:327–356.

Blackorby, C., Donaldson, D., and Weymark, J. (1990). A Welfaristic Proof of Arrow's Theorem. *Recherches Economiques de Louvain*, 56:259–286.

Blaug, M. (1992). *The Methodology of Economics*. Cambridge University Press, Cambridge.

Border, K. C. (1985). More on Harsanyi's Utilitarian Cardinal Welfare Theorem. *Social Choice and Welfare*, 1:279–281.

Bordes, G., Hammond, P. J., and Le Breton, M. (1997). Social Welfare Functionals on Restricted Domains and in Economic Environments. Stanford University, Department of Economics Working Paper No. 97-023. Available at `http://www-econ.stanford.edu/faculty/workp/index.html`.

Bossert, W. (1991). On Intra- and Interpersonal Utility Comparisons. *Social Choice and Welfare*, 8:207–219.

Broome, J. (1990). Bolker-Jeffrey Expected Utility Theory and Axiomatic Utilitarianism. *Review of Economic Studies*, 57:477–502.

Broome, J. (1991a). *Weighing Goods: Equality, Uncertainty and Time*. Basil Blackwell, Oxford.

Broome, J. (1991b). Utilitarian Metaphysics? In Elster, J. and Roemer, J. E., editors, *Interpersonal Comparisons of Well-Being*, chapter 3, pages 70–97. Cambridge University Press, Cambridge.

Broome, J. (1993). A Cause of Preference is not an Object of Preference. *Social Choice and Welfare*, 10:57–68.

Broome, J. (1998). Extended Preferences. In Fehige, C. and Wessels, U., editors, *Preferences*, pages 271–287. W. de Gruyter, Berlin. Reprinted as chapter 3 of J. Broome (1999), *Ethics Out of Economics*, Cambridge University Press, Cambridge.

Chipman, J. S. (1987). Compensation Tests. In Eatwell, J., Milgate, M., and Newman, P., editors, *The New Palgrave Dictionary of Economics*. Macmillan, London.

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

Christensen, L. R., Jorgenson, D. W., and Lau, L. J. (1975). Transcendental Logarithmic Utility Functions. *American Economic Review*, 65:367–383. Reprinted as chapter 1 of Jorgenson (1997a).

Clark, A. and Oswald, A. J. (1994). Unhappiness and Unemployment. *Economic Journal*, 104:648–659.

Clark, A. and Oswald, A. J. (2002). A Simple Statistical Method for Measuring How Life Events Affect Happiness. *International Journal of Epidemiology*, 31:1139–1144.

Cohen, G. A. (1989). On the Currency of Equal Justice. *Ethics*, 99:906–944.

Cohen, G. A. (1993). Equality of What? On Welfare, Goods, and Capabilities. In Nussbaum, M. C. and Sen, A. K., editors, *The Quality of Life*, pages 9–29. Clarendon Press, Oxford.

Cooter, R. and Rappoport, P. (1984). Were the Ordinalists Wrong about Welfare Economics? *Journal of Economic Literature*, 22:507–530.

Cooter, R. and Rappoport, P. (1985). Reply to I.M.D. Little's Comment. *Journal of Economic Literature*, 23:1189–1191.

Coughlin, P. J. (1986). Rights and the Private Pareto Principle. *Economica*, 53:303–320.

Coulhon, T. and Mongin, P. (1989). Social Choice Theory in the Case of Von Neumann-Morgenstern Utilities. *Social Choice and Welfare*, 6:175–187.

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

d'Aspremont, C. (1985). Axioms for Social Welfare Orderings. In Hurwicz, L., Schmeidler, D., and Sonnenschein, H., editors, *Social Goals and Social Organization*, chapter 2, pages 19–76. Cambridge University Press, Cambridge.

d'Aspremont, C. and Gevers, L. (1977). Equity and the Informational Basis of Collective Choice. *Review of Economic Studies*, 44:199–209.

Davis, J. B. (1990). Cooter and Rappoport on the Normative. *Economics and Philosophy*, 6:139–146.

Davis, J. B. (1991). Interpretation of Interpersonal Utility Comparisons—Positive, Normative or Descriptive? *Journal of Income Distribution*, 1:73–90.

De Meyer, B. and Mongin, P. (1995). A Note on Affine Aggregation. *Economics Letters*, 47:177–183.

Deaton, A. (1975). *Models and Projections of Demand in Post-War Britain*. Chapman and Hall, London.

Deaton, A. and Muellbauer, J. (1980a). *Economics and Consumer Behavior*. Cambridge University Press, Cambridge.

Deaton, A. and Muellbauer, J. (1980b). An Almost Ideal Demand System. *American Economic Review*, 70:312–326.

DeMeyer, F. and Plott, C. R. (1970). The Probability of a Cyclical Majority. *Econometrica*, 38:345–354.

Dhillon, A. and Mertens, J. F. (1999). Relative Utilitarianism. *Econometrica*, 67:471–498.

Di Tella, R., MacCulloch, R. J., and Oswald, A. J. (2001). Preferences over Inflation and Unemployment: Evidence from Surveys of Happiness. *American Economic Review*, 91:335–341.

Diamond, P. A. (1967). Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility: Comment. *Journal of Political Economy*, 75:765–766.

Diamond, P. A. and Mirrlees, J. A. (1971). Optimal Taxation and Public Production I: Production Efficiency and II: Tax Rules. *American Economic Review*, 61:8–27 and 261–278.

Dixit, A. and Norman, V. (1980). *Theory of International Trade*. James Nisbet, Welwyn, Herts.

Dixit, A. and Norman, V. (1986). Gains from Trade without Lump-Sum Compensation. *Journal of International Economics*, 21:99–110.

Dworkin, R. (1981). What is Equality? Part 2: Equality of resources. *Philosophy & Public Affairs*, 10:283–345.

Dworkin, R. (2000). *Sovereign Virtue. The Theory and Practice of Equality*. Harvard University Press, Cambridge, Mass.

Easterlin, R. (1974). Does Economic Growth Improve the Human Lot? Some Empirical Evidence. In David, P. A. and Reder, M. W., editors, *Nations and Households in Economic Growth: Essays in Honour of Moses Abramowitz*. Academic Press, New York.

Easterlin, R. (1995). Will Raising the Income of All Increase the Happiness of All? *Journal of Economic Behavior and Organization*, 27:35–48.

Ebert, U. (1997). Social Welfare When Needs Differ. *Economica*, 64:233–244.

Fishburn, P. C. (1970). *Utility Theory for Decision Making*. John Wiley, New York.

Fleurbaey, M. (1995). Equal Opportunity or Equal Social Outcome? *Economics and Philosophy*, 11:25–55.

Fleurbaey, M. and Maniquet, F. (forthcoming). Utilitarianism versus Fairness in Welfare Economics. In Salles, M. and Weymark, J. A., editors, *Justice, Political Liberalism and Utilitarianism: Themes from Harsanyi and Rawls*. Cambridge University Press, Cambridge.

Foley, D. K. (1970). Lindahl's Solution and the Core of an Economy with Public Goods. *Econometrica*, 38:66–72.

Frank, R. H. (1997). The Frame of Reference as a Public Good. *Economic Journal*, 107:1832–1847.

Geary, R. C. (1949–50). A Note on 'A Constant Utility Index of the Cost of Living'. *Review of Economic Studies*, 18:65–66.

Gehrlein, W. V. (1997). Condorcet's Paradox and the Condorcet Efficiency of Voting Rules. *Mathematica Japonica*, 45:173–199.

Gehrlein, W. V. and Fishburn, P. C. (1976). Condorcet's Paradox and Anonymous Preference Profiles. *Public Choice*, 26:1–18.

Gevers, L. (1986). Walrasian Social Choice: Some Simple Axiomatic Approaches. In Heller, W. P., Starr, R. M., and Starrett, D. A., editors, *Social Choice and Public Decision Making: Essays in Honor of Kenneth J. Arrow. Vol. 1.* Cambridge University Press, Cambridge.

Glewwe, U. (1997). Household Equivalence Scales and the Measurement of Inequality: Transfers from the Poor to the Rich Could Decrease Inequality. *Journal of Public Economics*, 44:211–216.

Goodin, R. (1986). Laundering Preferences. In Elster, J. and Hylland, A., editors, *Foundations of social choice theory*, chapter 3, pages 75–102. Cambridge University Press, Cambridge.

Gorman, W. M. (1953). Community Preference Fields. *Econometrica*, 21:63–80.

Gorman, W. M. (1961). On a Class of Preference Fields. *Metroeconomica*, 13:53–56. Reprinted in C. Blackorby and A.F. Shorrocks (eds.) (1995), *Separability and Aggregation: Collected Works of W.M. Gorman, Volume 1*, Clarendon Press, Oxford.

Gorman, W. M. (1976). Tricks with Utility Functions. In Artis, M. J. and Nobay, A. R., editors, *Essays in Economic Analysis*. Cambridge University Press, Cambridge.

Guesnerie, R. (1995). *A Contribution to the Pure Theory of Taxation*. Cambridge University Press, Cambridge.

Hammond, P. J. (1975). A Note on Extreme Inequality Aversion. *Journal of Economic Theory*, 11:465–467.

Hammond, P. J. (1976). Equity, Arrow's Conditions, and Rawls' Difference Principle. *Econometrica*, 44:793–804.

Hammond, P. J. (1977a). Dynamic Restrictions on Metastatic Choice. *Economica*, 44:337–350.

Hammond, P. J. (1977b). Dual Interpersonal Comparisons of Utility and the Welfare Economics of Income Distribution. *Journal of Public Economics*, 7:51–71.

Hammond, P. J. (1979). Equity in Two Person Situations: Some Consequences. *Econometrica*, 47:1127–1135.

Hammond, P. J. (1980). Dual Interpersonal Comparisons of Utility and the Welfare Economics of Income Distribution: A Corrigendum. *Journal of Public Economics*, 14:105–106.

Hammond, P. J. (1982). Liberalism, Independent Rights and the Pareto Principle. In Cohen, L. J., Łoś, J., Pfeiffer, H., and Podewski, K.-P., editors, *Logic, Methodology and the Philosophy of Science, VI*, pages 607–620. North-Holland, Amsterdam.

Hammond, P. J. (1987a). Social Choice: the Science of the Impossible? In Feiwel, G. R., editor, *Arrow and the Foundations of the Theory of Economic Policy*, chapter 1B, pages 116–134. New York University Press, New York.

Hammond, P. J. (1987b). On Reconciling Arrow's Theory of Social Choice With Harsanyi's Fundamental Utilitarianism. In Feiwel, G. R., editor, *Arrow and the Foundations of the Theory of Economic Policy*, chapter 4, pages 179–222. New York University Press, New York.

Hammond, P. J. (1988). Consequentialist Demographic Norms and Parenting Rights. *Social Choice and Welfare*, 5:127–145.

Hammond, P. J. (1990). Theoretical Progress in Public Economics: A Provocative Assessment. *Oxford Economic Papers*, 42:6–33. Also in P.J.N. Sinclair and M.D.E. Slater (eds.) (1991), *Taxation, Private Information and Capital*, Clarendon Press, Oxford.

Hammond, P. J. (1991a). Interpersonal Comparisons of Utility: Why and How They Are and Should Be Made. In Elster, J. and Roemer, J. E., editors, *Interpersonal Comparisons of Well-Being*, chapter 7, pages 200–254. Cambridge University Press, Cambridge.

Hammond, P. J. (1991b). Independence of Irrelevant Interpersonal Comparisons. *Social Choice and Welfare*, 8:1–19.

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

Hammond, P. J. (1994). Money Metric Measures of Individual and Social Welfare Allowing for Environmental Externalities. In Eichhorn, W., editor, *Models and Measurement of Welfare and Inequality*, pages 694–724. Springer-Verlag, Berlin.

Hammond, P. J. (1995). Social Choice of Individual and Group Rights. In Barnett, W. A., Moulin, H., Salles, M., and Schofield, N., editors, *Social Choice, Welfare, and Ethics*, chapter 3, pages 55–77. Cambridge University Press, Cambridge.

Hammond, P. J. (1996). Consequentialist Decision Theory and Utilitarian Ethics. In Farina, F., Hahn, F., and Vannucci, S., editors, *Ethics, Rationality, and Economic Behaviour*, pages 92–118. Clarendon Press, Oxford.

Hammond, P. J. (1999). Multilaterally Strategy-Proof Mechanisms in Random Aumann-Hildenbrand Macroeconomies. In Wooders, M., editor, *Topics in Game Theory and Mathematical Economics: Essays in Honor of Robert J. Aumann*, pages 171–187. American Mathematical Society, Providence, RI.

Hammond, P. J. (2000). Reassessing the Diamond/Mirrlees Efficiency Theorem. In Hammond, P. J. and Myles, G. D., editors, *Incentives, Organization, and Public Economics: Papers in Honour of Sir James Mirrlees*, chapter 12, pages 193–216. Oxford University Press.

Hammond, P. J. and Sempere, J. (1995). Limits to the Potential Gains from Economic Integration and Other Supply Side Policies. *Economic Journal*, 105:1180–1204.

Harberger, A. C. (1971). Three Basic Postulates for Applied Welfare Analysis. *Journal of Economic Literature*, 9:785–797.

Harberger, A. C. (1978). On the Use of Distributional Weights in Social Cost Benefit Analysis. *Journal of Political Economy*, 86:S87–S120.

Harsanyi, J. C. (1955). Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility. *Journal of Political Economy*, 63:309–321. Reprinted in Phelps (1973) and Harsanyi (1976).

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

Harsanyi, J. C. (1982). Morality and the Theory of Rational Behaviour. In Sen, A. K. and Williams, B., editors, *Utilitarianism and Beyond*, chapter 2, pages 39–62. Cambridge University Press and Editions de la MSH, Cambridge and Paris.

Harsanyi, J. C. (1987). Interpersonal Utility Comparison. In Eatwell, J., Milgate, M., and Newman, P., editors, *The New Palgrave Dictionary of Economics*. Macmillan, London.

Hennipman, P. (1988). A New Look at the Ordinalist Revolution: Comments on Cooter and Rappoport. *Journal of Economic Literature*, 26:89–91.

Hennipman, P. (1992). Hicks, Robbins, and the Demise of Pigovian Welfare Economics: Rectification and Amplification. *Southern Economic Journal*, 59:88–97.

Hicks, J. R. (1959). *Essays in World Economics*. Clarendon Press, Oxford.

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

Isbell, J. R. (1959). Absolute Games. In Tucker, A. W. and Luce, R. D., editors, *Contributions to the Theory of Games*, pages 357–396. Princeton University Press, Princeton.

Jorgenson, D. W. (1990). Aggregate Consumer Behavior and the Measurement of Social Welfare. *Econometrica*, 58:1007–1040. Reprinted as chapter 1 of Jorgenson (1997b).

Jorgenson, D. W. (1997a). *Welfare Volume 1: Aggregate Consumer Behavior*. MIT Press, Cambridge, Mass.

Jorgenson, D. W. (1997b). *Welfare Volume 2: Measuring Social Welfare*. MIT Press, Cambridge, Mass.

Jorgenson, D. W., Lau, L. J., and Stoker, T. M. (1982). The Transcendental Logarithmic Model of Aggregate Consumer Behavior. In Basmann, R. L. and Rhodes, G., editors, *Advances in Econometrics, Vol. 1*, pages 97–238. JAI Press, Greenwich, Conn. Reprinted as chapter 8 of Jorgenson (1997a).

Jorgenson, D. W. and Slesnick, D. T. (1983). Individual and Social Cost-of-Living Indices. In Diewert, W. E. and Montmarquette, C., editors, *Price Level Measurement*, pages 241–323. Statistics Canada, Ottawa. Reprinted as chapter 2 of Jorgenson (1997b).

Jorgenson, D. W. and Slesnick, D. T. (1987). Aggregate Consumer Behavior and Household Equivalence Scales. *Journal of Business and Economic Statistics*, 5:219–232. Reprinted as chapter 5 of Jorgenson (1997b).

Kahneman, D., Wakker, P. P., and Sarin, R. (1997). Back to Bentham? Explorations of Experienced Utility. *Quarterly Journal of Economics*, 112:375–406.

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

Kaneko, M. (1984). On Interpersonal Utility Comparisons. *Social Choice and Welfare*, 1:165–175.

Kaneko, M. and Nakamura, K. (1979). The Nash Social Welfare Function. *Econometrica*, 47:423–435.

Kemp, M. and Ng, Y. K. (1987). Arrow's Independence Condition and the Bergson-Samuelson Tradition. In Feiwel, G. R., editor, *Arrow and the Foundations of the Theory of Economic Policy*, chapter 5, pages 223–242. New York University Press, New York.

Klein, L. R. and Rubin, H. (1947–48). A Constant Utility Index of the Cost of Living. *Review of Economic Studies*, 15:84–87.

Kolm, S.-C. (1968). The Optimal Production of Social Justice. In Margolis, J. and Guitton, H., editors, *Public Economics: An Analysis of Public Production and Consumption and their Relations to the Private Sector*, pages 145–200. Macmillan, London.

Kolm, S.-C. (1972). *Justice et équité*. Editions du CNRS, Paris. Translated as *Justice and Equity* (1998), MIT Press, Cambridge, Mass.

Kolm, S.-C. (1994). The Meaning of Fundamental Preferences. *Social Choice and Welfare*, 11:193–198.

Kolm, S.-C. (1995). The Economics of Social Sentiments: The Case of Envy. *Japanese Economic Review*, 46:63–87.

Kolm, S.-C. (1996). *Modern Theories of Justice*. Mass: MIT Press, Cambridge.

Lau, L. J. (1982). A Note on the Fundamental Theorem of Exact Aggregation. *Economics Letters*, 9:119–126.

Lewbel, A. (1989). Household Equivalence Scales and Welfare Comparisons. *Journal of Public Economics*, 39:377–391.

Lewbel, A. (1991). Cost of Characteristics Indices and Household Equivalence Scales. *European Economic Review*, 35:1277–1293.

Lewbel, A. (1993). Household Equivalence Scales and Welfare Comparisons: Reply. *Journal of Public Economics*, 50:147–148.

Little, I. M. D. (1957). *A Critique of Welfare Economics*. Oxford University Press, Oxford.

May, K. O. (1952). A Set of Independent Necessary and Sufficient Conditions for Simple Majority Decisions. *Econometrica*, 20:680–684.

May, K. O. (1985). Robert Cooter and Peter Rappoport, 'Were the Ordinalists Wrong about Welfare Economics?': A Comment. *Journal of Economic Literature*, 23:1186–1188.

Mayston, D. J. (1974). *The Idea of Social Choice*. Macmillan, London.

Mayston, D. J. (1980). Where Did Prescriptive Welfare Economics Go Wrong? In Currie, D. A. and Peters, W., editors, *Contemporary Economic Analysis, Vol. 2*, chapter 7, pages 175–216. Croom-Helm, London.

McGarvey, D. (1953). A Theorem on the Construction of Voting Paradoxes. *Econometrica*, 21:608–610.

McKelvey, R. D. (1976). Intransitivities in Multidimensional Voting Models and Some Implications for Agenda Control. *Journal of Economic Theory*, 12:472–482.

Milleron, J. C. (1972). Theory of Value with Public Goods: A Survey Article. *Journal of Economic Theory*, 5:419–477.

Mirrlees, J. A. (1971). An Exploration in the Theory of Optimum Income Taxation. *Review of Economic Studies*, 38:175–208.

Mirrlees, J. A. (1982). The Economic Uses of Utilitarianism. In Sen, A. K. and Williams, B., editors, *Utilitarianism and Beyond*, chapter 3, pages 63–84. Cambridge University Press and Editions de la Maison des Sciences de l'Homme, Cambridge and Paris.

Mongin, P. and d'Aspremont, C. (1998). Utility Theory and Ethics. In Barberà, S., Hammond, P. J., and Seidl, C., editors, *Handbook of Utility Theory, Vol. I*, chapter 10, pages 371–482. Kluwer, Dordrecht.

Moulin, H. and Thomson, W. (1996). Axiomatic Analysis of Resource Allocation Problems. In Arrow, K. J., Sen, A. K., and Suzumura, K., editors, *Social Choice Re-Examined, Vol. 1*, chapter 5, pages 101–120. Macmillan, London. IEA Conference Volume, no. 117.

Myers, D. G. (1993). *The Pursuit of Happiness: Who is Happy and Why?* Avon, New York.

Myerson, R. (1985). Bayesian Equilibrium and Incentive Compatibility: An Introduction. In Hurwicz, L. and Schmeidler, D., editors, *Social Goals and Social Organization*, chapter 8, pages 229–259. Cambridge University Press, Cambridge.

Nagahisa, R. and Suga, K. (1998). Impossibility Theorems with Interpersonally Comparable Welfare Levels: The 'Extended Sympathy' Approach Reconsidered. *Japanese Economic Review*, 49:171–190.

Narens, L. and Luce, R. D. (1983). How We May Have Been Misled into Believing in the Interpersonal Comparability of Utility. *Theory and Decision*, 15:247–260.

Nash, J. F. (1950). The Bargaining Problem. *Econometrica*, 18:155–162.

Ng, Y. K. (1996). Happiness Surveys: Some Comparability Issues and an Exploratory Survey Based on Just Perceivable Increments. *Social Indicators Research*, 38:1–27.

Nussbaum, M. C. (1993). Non-Relative Virtues: An Aristotelian Approach. In Nussbaum, M. C. and Sen, A. K., editors, *The Quality of Life*, pages 242–269. Clarendon Press, Oxford.

Oswald, A. J. (1997). Happiness and Economic Performance. *Economic Journal*, 107:1815–1831.

Parfit, D. (1987). *Reasons and Persons*. Oxford University Press, Oxford.

Pazner, E. (1979). Equity, Nonfeasible Alternatives and Social Choice: A Reconsideration of the Concept of Social Welfare. In Laffont, J. J., editor, *Aggregation and Revelation of Preferences*. North-Holland, Amsterdam.

Pazner, E. and Schmeidler, D. (1978). Egalitarian Equivalent Allocations: A New Concept of Economic Equity. *Quarterly Journal of Economics*, 92:671–687.

Peters, H. J. M. (1992). *Axiomatic Bargaining Game Theory*. Kluwer, Dordrecht.

Phelps, E. S. (1973). *Economic Justice*. Penguin Books, Harmondsworth.

Plott, C. R. (1973). Path Independence, Rationality, and Social Choice. *Econometrica*, 41:1075–1091.

Pratt, J. W. (1964). Risk Aversion in the Small and in the Large. *Econometrica*, 32:122–136.

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

Robbins, L. (1932). *An Essay on the Nature and Significance of Economic Science*. Macmillan, London.

Robbins, L. (1938). Interpersonal Comparisons of Utility: A Comment. *Economic Journal*, 48:635–641.

Robbins, L. (1981). Economics and Political Economy. *American Economic Review, Papers and Proceedings*, 71:1–10.

Roberts, K. W. S. (1980a). Possibility Theorems with Interpersonally Comparable Welfare Levels. *Review of Economic Studies*, 47:409–420.

Roberts, K. W. S. (1980b). Interpersonal Comparability and Social Choice Theory. *Review of Economic Studies*, 47:421–439.

Roberts, K. W. S. (1995). Valued Opinions or Opinionated Values: The Double Aggregation Problem. In Basu, K., Pattanaik, P. K., and Suzumura, K., editors, *Choice, Welfare and Development: A Festschrift in Honour of Amartya K. Sen*, chapter 8, pages 141–165. Oxford University Press, Oxford.

Roberts, K. W. S. (1996). Interpersonal Comparisons of the Extended Sympathy Type and the Possibility of Social Choice: Discussion. In Arrow, K. J., Sen, A. K., and Suzumura, K., editors, *Social Choice Re-examined, Vol. 2*, pages 230–236. Macmillan, London. IEA Conference Volume, no. 117.

Roberts, K. W. S. (1997). Objective Interpersonal Comparisons of Utility. *Social Choice and Welfare*, 14:79–96.

Roemer, J. E. (1993). A Pragmatic Theory of Responsibility for the Egalitarian Planner. *Philosophy and Public Affairs*, 22:146–166.

Roemer, J. E. (1998). *Equality of Opportunity*. Harvard University Press, Cambridge, Mass.

Rosenbaum, E. F. (1995). Interpersonal Comparisons of Utility: Positive, Normative or Value-Laden? *Journal of Economic Methodology*, 2:239–257.

Roth, A. E. (1979). *Axiomatic Models of Bargaining*. Springer Verlag, Berlin.

Samuelson, P. A. (1947). *Foundations of Economic Analysis*. Harvard University Press, Cambridge, Mass.

Samuelson, P. A. (1947–48). Some Implications of Linearity. *Review of Economic Studies*, 15:88–90.

Samuelson, P. A. (1956). Social Indifference Curves. *Quarterly Journal of Economics*, 70:1–22.

Samuelson, P. A. (1977). Reaffirming the Existence of 'Reasonable' Bergson–Samuelson Social Welfare Functions. *Economica*, 44:81–88.

Samuelson, P. A. (1987). Sparks from Arrow's Anvil. In Feiwel, G. R., editor, *Arrow and the Foundations of the Theory of Economic Policy*, chapter 3, pages 154–178. New York University Press, New York.

Schick, F. (1971). Beyond Utilitarianism. *Journal of Philosophy*, 68:657–666.

Schüssler, R. (1998). Wish You Were Me: A Reply to Broome and a Comment on Harsanyi's Extended Preference Theory. In Fehige, C. and Wessel, U., editors, *Preferences*, pages 288–297. W. de Gruyter, Berlin.

Scitovsky, T. (1941). A Note on Welfare Propositions in Economics. *Review of Economic Studies*, 9:77–88.

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

Sen, A. K. (1970b). The Impossibility of a Paretian Liberal. *Journal of Political Economy*, 78:152–157.

Sen, A. K. (1973). *On Economic Inequality*. Clarendon Press, Oxford.

Sen, A. K. (1977). On Weights and Measures: Informational Constraints in Social Welfare Analysis. *Econometrica*, 45:1539–1572.

Sen, A. K. (1979). Personal Utilities and Public Judgements: or What's Wrong with Welfare Economics? *Economic Journal*, 89:537–558.

Sen, A. K. (1980). Equality of What? In McMurrin, S. M., editor, *The Tanner Lectures on Human Values, Vol. I*. The University of Utah Press and Cambridge University Press, Salt Lake City and Cambridge.

Sen, A. K. (1985). *Commodities and Capabilities*. North-Holland, Amsterdam.

Sen, A. K. (1986). Social Choice Theory. In Arrow, K. J. and Intriligator, M. D., editors, *Handbook of Mathematical Economics, Vol. III*, chapter 22, pages 1073–1181. North-Holland, Amsterdam.

Sen, A. K. (1987). *On Ethics and Economics*. Clarendon Press, Oxford.

Sen, A. K. (1992). *Inequality Reexamined*. Clarendon Press, Oxford.

Shorrocks, A. F. (1995). Inequality and Welfare Evaluation of Heterogeneous Income Distributions. Discussion Paper 447, University of Essex.

Simmons, P. J. (1980). Discussion. In Currie, D. A. and Peters, W., editors, *Contemporary Economic Analysis, Vol. 2*, pages 217–220. Croom-Helm, London.

Simon, J. A. (1974). Interpersonal Welfare Comparisons Can be Made — and Used for Redistribution Decisions. *Kyklos*, 27:63–98.

Slesnick, D. T. (2001). *Consumption and Social Welfare: Living Standards and Their Distribution in the United States*. Cambridge University Press, Cambridge.

Stone, J. R. N. (1954). Linear Expenditure Systems and Demand Analysis: An Application to the Pattern of British Demand. *Economic Journal*, 64:511–527.

Strasnick, S. (1977). Ordinality and the Spirit of the Justified Dictator. *Social Research*, 44:668–690.

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

Suppes, P. (1966). Some Formal Models of Grading Principles. *Synthese*, 6:284–306.

Suzumura, K. (1983). Resolving Conflicting Views of Justice in Social Choice. In Pattanaik, P. K. and Salles, M., editors, *Social Choice and Welfare*, chapter 8, pages 125–149. North-Holland, Amsterdam.

Suzumura, K. (1996). Interpersonal Comparisons of the Extended Sympathy Type and the Possibility of Social Choice. In Arrow, K. J., Sen, A. K., and Suzumura, K., editors, *Social Choice Re-Examined, Vol. 2*, chapter 15, pages 202–229. Macmillan, London. IEA Conference Volume, no. 117.

Thomson, W. (1988). A Study of Choice Correspondences in Economies with a Variable Number of Agents. *Journal of Economic Theory*, 46:237–254.

Thomson, W. (1994). *Bargaining Theory: The Axiomatic Approach*. Academic Press, San Diego.

van Parijs, P. (1995). *Real Freedom for All*. Oxford University Press, Oxford.

Veenhoven, R. (1993). *Happiness in Nations: Subjective Appreciation of Life in 56 Nations 1946–1992*. Erasmus University Risbo, Rotterdam.

Vickrey, W. S. (1945). Measuring Marginal Utility by the Reactions to Risk. *Econometrica*, 13:319–333.

Weymark, J. A. (1984). Arrow's Theorem with Social Quasi-Orderings. *Public Choice*, 42:235–246.

Weymark, J. A. (1991). A Reconsideration of the Harsanyi–Sen Debate on Utilitarianism. In Elster, J. and Roemer, J. E., editors, *Interpersonal Comparisons of Well-Being*, chapter 8, pages 255–320. Cambridge University Press, Cambridge.

Weymark, J. A. (1993). Harsanyi's Social Aggregation Theorem and the Weak Pareto Principle. *Social Choice and Welfare*, 10:209–221.

Weymark, J. A. (1995). Further Remarks on Harsanyi's Social Aggregation Theorem and the Weak Pareto Principle. *Social Choice and Welfare*, 12:87–92.

Wilson, R. B. (1972). Social Choice without the Pareto Principle. *Journal of Economic Theory*, 5:478–486.

Winkelmann, L. and Winkelmann, R. (1998). Why Are the Unemployed So Unhappy? Evidence from Panel Data. *Economica*, 65:1–15.

Young, H. P. (1974). An Axiomatization of the Borda Rule. *Journal of Economic Theory*, 9:43–52.

Young, H. P. (1988). Condorcet's Theory of Voting. *American Political Science Review*, 82:1231–1244.

Young, H. P. (1994). *Equity. In Theory and Practice*. Princeton University Press, Princeton.

Young, H. P. (1995). Optimal Voting Rules. *Journal of Economic Perspectives*, 9:51–64.

Young, H. P. and Levenglick, A. (1978). A Consistent Extension of Condorcet's Election Principle. *SIAM Journal of Applied Mathematics*, 35:285–300.