Social welfare functionals and interpersonal comparability *

Claude d'Aspremont[†] and Louis Gevers[‡]

Abstract

This chapter reviews the SWFL approach to social choice. It does not attempt to be a complete and systematic survey of existing results, but to give a critical assessment of the main axioms and their role in filtering the ethically relevant information, in particular the measurability and comparability properties of individual evaluation functions. Social welfare functionals are defined formally together with closely related concepts. After adducing a good number of examples, we elaborate on the meaning of the SWFL domain of definition and we sketch some alternative approaches. Several types of axioms are considered; some of them are used to filter the relevant information while others express collective efficiency or equity requirements. Then, ti illustrate the various tradeoffs among these axioms, selected characterization results are presented; most of them are cast in what we call the formally welfarist framework. Finally, we have assembled some other characterizations which eschew either invariance properties or the formally welfarist framework. We discuss the treatment of two sets of social alternatives endowed with an enriched structure, viz. the set of classical exchange economies and the complete set of lotteries one can define on an abstract set of pure alternatives. As an introduction to the latter discuss, we elaborate on the difficulties raised by social evaluation when risks and uncertainty are taken explicitly into account.

Keywords: social welfare functionals, social welfare orderings, invariance axioms, utilitarianism, welfarism

JEL classification: D71, D81

---

*Chapter 10 in K. Arrow, A. sen and K. Suzumura (eds.), *Handbook of Social Choice and Welfare, Volume 1*, Amsterdam, Elsevier Science,459-541, 2002 – Reprint 1564.

We thank an anonymous referee for many useful remarks and W. Bossert, A. Dhillon, M. Fleurbaey, F. Gaspart, F. Maniquet, J.F. Mertens, K. Suzumura and J. Weymark for helpful discussions.

†CORE, Louvain-University, Louvain-la-Neuve, Belgium

‡University of Namur and CORE

Contents

1 Introduction **1**

2 Social welfare functionals and related concepts **5**

2.1 Definitions . . . 5

2.2 Examples . . . 9

2.2.1 Imposed SWFLs . . . 9

2.2.2 Dictatorship . . . 9

2.2.3 Maximin and leximin . . . 10

2.2.4 Pure utilitarianism, weighted utilitarianism and relative utilitarianism . . 11

2.2.5 Weighted rank utilitarianism and the generalised Gini family . . . 13

2.2.6 Nash's bargaining solution . . . 13

2.2.7 Borda's method of voting . . . 14

2.2.8 Majority voting . . . 15

2.3 Domain interpretation . . . 16

2.4 Some related concepts . . . 19

3 Axioms and their use **20**

3.1 Preliminary . . . 20

3.2 Invariance axioms . . . 21

3.2.1 Comparisons of evaluation levels . . . 23

3.2.2 Comparisons of evaluation differences . . . 26

3.2.3 Comparisons of evaluation indicator ratios . . . 32

3.2.4 Other invariance axioms . . . 34

3.3 Other information-filtering axioms and formal welfarism . . . 36

3.3.1 Pareto indifference, independence and neutrality . . . 36

3.3.2 Formal welfarism and invariance properties . . . 42

3.3.3 Separability, continuity and convexity properties . . . 44

3.3.4 Alternative approaches to formal welfarism 47

3.4 Pareto dominance principles and weak welfarism 49

3.5 Equity axioms 55

4 Independence and invariance-based characterizations 60

4.1 Restricting interpersonal level comparability 60

4.2 Full comparability: from cardinal to ordinal measurability 68

4.2.1 No separability 68

4.2.2 A modicum of Separability 71

4.2.3 Full Separability 73

4.3 Homothetic vs. translatable social welfare functionals 75

5 Discarding neutrality or invariance 77

5.1 Uncertainty and risk: from Harsanyi to Relative Utilitarianism 77

5.2 On some egalitarian social rankings 90

6 Conclusion 95

1 Introduction

Insofar as it probes the foundations of political constitutions, social choice theory deals with the relationship collective decisions or preferences ought to bear with individual preferences. Arrow (1951) launched the first systematic attack on this problem at a formal level. He aimed at generality, paid no attention to the specifics of usual economic models, and assumed that individual preferences could be of any shape whatsoever.

Although the set of problems raised by political decisions and those raised by social-evaluation judgements share the same basic formal structure, they ought to be distinguished sharply from each other. Political decisions are usually arrived at in groups, and individual preferences cannot be filtered; they must be accepted as they are while manipulation attempts cannot be excluded from communication channels with the center. Therefore, such things as election procedures and assembly rules are likely to be of central importance. Game-theoretic equilibrium concepts have been found relevant in this context, and equilibrium correspondences are a key link between individual preferences and collective decisions. Political equilibrium may often be interpreted as some form of compromise; in case new information becomes available before it is enacted, one does not expect the revised compromise to evolve from its predecessor in a very rational fashion.

In contrast, when an ordinary citizen attempts to take the standpoint of an ethical observer in order to formulate social-evaluation judgments, perhaps as an input towards establishing eventually an optimal voting strategy, he or she is in a position to tray new information more rationally; his or her attention is likely to be focused on the content of social outcomes and their consequences, at least as much as on the procedures followed to arrive at them. Moreover, the appropriate summary statistics describing what is ethically relevant from each individual's viewpoint do not necessarily have a utility interpretation, unless one is persuaded by the welfarist tradition. Even if this is the case, there is nothing schizophrenic when a voting procedure is approved by someone who is morally objecting to the consequences of some piece of legislation adopted in accordance with it. At least two kinds of reason may account for this. On the one hand, voting procedures have a more permanent character than ordinary legislative output;

on the other hand, the ethical observer may be relying on subjective prior information or on external information without paying a great deal of formal attention to manipulation.

Between the political body and the common citizen, public officials occupy an intermediate position: each of them should ideally be concerned with social evaluation, but, as Bergson (1954) writes, “the values to be taken as data are not those which might guide the official if he were a private citizen ... His one aim in life is to implement the values of other citizens as given by some rule of collective decision-making”. Bergson stresses that most pre-Arrovian welfare economists were discussing the ordinary citizen’s problem, whereas Arrow’s analysis of social choice is more relevant for public officials. Arrow (1963, p. 107), from whom we borrow the above quotation, marks his agreement with this interpretation. Indeed, his exclusive reliance on lists of individual preferences as informational basis of collective preference may be well suited for discussing the formal aspects of political decisions. But their economical use can hardly accommodate any concern for equity. For instance, when comparing income distributions, this concern might be expressed by the following doctrine: (1) individual utilities are all which matters for social evaluation, (2) they are concave functions of income, which is their single argument, and (3) their sum ought to be maximized by society. This simplified utilitarian approach is out of reach of the seminal Arrovian model. Indeed, it is based on individual preference relations which do not lend themselves to ethically meaningful interpersonal comparisons. Any one of two features may account for this shortcoming: too little structure is imposed on the set of social decisions or alternatives, and moreover, social preference between any two alternatives is required to depend only on individual preferences restricted to this pair. The latter principle was called by Arrow *independence of irrelevant alternatives.* It greatly contributes to informational parsimony. However, if it is weakened or deleted, the formal construction of interpersonal utility comparisons can be obtained as the by-product of a voting procedure, but the social ranking associated with it is unlikely to be adequate for social evaluation. A more meaningful construction can also be found in the literature; it relies on a set of alternatives endowed with a structure which is richer and less abstract: for instance, uncertainty is made explicit and social preferences are required to admit of an expected utility representation, or context-specific domain restrictions and equity arguments are brought in, as in the recent literature on axiomatic allocation theory.

Quite some years before these developments, Sen (1970) opened several paths branching out of Arrow's trail. Along the one we plan to follow, Sen reconsiders the problem of social evaluation from the viewpoint of an ethical observer who might be a private citizen. In contrast with the more recent approach we just sketched, he refrains from giving the set of alternatives any specific structure, and his innovation pertains to the informational basis of social-evaluation judgments: it is assumed to consist of all logically possible lists of individual *utility functions.* Except for the hardly significant case of individual preferences failing to be numerically representable, the above formal approach, which is crystallized in the concept of *social welfare functional* (SWFL), is more general than Arrow's, since individual utility levels or gains can be considered interpersonally comparable *a priori*, unless specific axioms restrict, for informational or moral reasons, the ethical observer's discriminating ability. However, generality involves a cost: there is a preparation stage, at which the ethical observer is to select an adequate list of *a priori* comparable utility representations of individual preferences, about which others do not necessarily agree. In the simple example already alluded to, the social evaluation of individual income vectors is very dependent on the degree of inequality aversion embodied in the concavity of each individual utility function, and thoroughly rational observers can be in total disagreement about this value judgment. Furthermore, SWFL theory can hardly help them solve their conflict, even though it does often provide assistance for finding out which value judgments are compatible and which are not. For example, it contains much clarification of the debate opposing utilitarianism with competing principles, but it cannot prescribe any complete ready-made recipe for every social-evaluation problem.

The major part of our survey is devoted to a review of developments of Sen's concept. The individual utility interpretation of its informational basis has been critized by Sen himself and by several political philosophers who stress that social-evaluation judgments should not be concerned exclusively with arbitrating the individuals' conflicts of interests as they are narrowly modeled by positive economic theories. But this criticism does not necessarily diminish the usefulness of the formal SWFL concept. Indeed, the latter can be reinterpreted as an application of multi-objective decision theory to the ethical observer's problem. It attempts to make explicit the link between the social evaluation of an alternative and its appraisal from the viewpoint of every individual in turn; the latter appraisals may themselves be linked with individual preferences, but they do not necessarily duplicate them. For instance, they could be represented by numerical indicators summarizing the individuals' doings and beings or their life expectancy or their set of opportunities. Stretching the use of an expression sparsely referred to in the literature, we shall can them *individual evaluation functions or indicators.* Aggregating them over the set of individuals is our main task in the present chapter.

Thus, the informal preparatory stage required before making use of the SWFL apparatus is not without advantage, as it makes for a versatile tool. Since the set of possible alternatives lacks any particular structure, the SWFL can accommodate the needs of an ethical observer who is more interested in appraising the instrumental value of social rules and political institutions than in focusing on single outcomes. It can also be found useful by someone persuaded by the piecemeal engineering approach of axiomatic allocation theorists; as they confine their analysis to a variety of specific economic environments, the SWFL can be thought of as a kind of residual tool for evaluating situations not yet adequately covered by the existing theoretical corpus.

However, it is also legitimate to adopt a positive interpretation of this body of literature: most people may be assumed to take up at least occasionally the position of an ethical observer. Empirical studies of professed evaluation judgments or decisions have started cropping up, and some of them are based on laboratory experiments.[1] A theory capable of structuring social-evaluation judgments is not only interesting in its own right but it can also usefully interact with the empirical part of an ambitious research program. There is a clear analogy with the theory of individual decisions under uncertainty.

To sum up, as they are defined by Sen (1970), social welfare functionals are maps determining the social ordering of the set of alternatives with help of a complete list of individual numerical indicators which are not necessarily interpreted as utility functions.[2]

[1] See for instance Bar-Hillel and Yaari (1984).

[2] A similar point of view is defended in Mongin and d'Aspremont (1988).

This chapter reviews the SWFL approach to social choice. It does not attempt to be a complete and systematic survey of existing results (this has been done elsewhere[3]), but to give a critical assessment of the main axioms and their role in filtering the ethically relevant information, in particular the measurability and comparability properties of individual evaluation functions. In Section 2, social welfare functionals are defined formally together with closely related concepts. After adducing a good number of examples, we elaborate on the meaning of the SWFL domain of definition and we sketch some alternative approaches. Section 3 is essentially devoted to the analysis of axioms considered in isolation; some of them are used to filter the relevant information while others express collective efficiency or equity requirements. Finally, to illustrate the various tradeoffs among these axioms, selected characterization results are presented in Section 4; most of them are cast in what we call the formally welfarist framework. In Section 5, we have assembled some other characterizations which eschew either invariance properties or the formally welfarist framework. We discuss at length the treatment of two sets of social alternatives endowed with an enriched structure, viz. the set of classical exchange economies and the complete set of lotteries one can define on an abstract set of pure alternatives. As an introduction to the latter discussion, we elaborate on the difficulties raised by social evaluation when risks and uncertainty are taken explicitly into account. Section 6 concludes.

2 Social welfare functionals and related concepts

2.1 Definitions

A society made up of a finite set $N = \{1, \cdots, n\}$ of individuals faces a set of possible social decisions or alternatives X consisting of at least three elements. Whether we interpret X as the set of all conceivably feasible alternatives or as the set of actually feasible social decisions in a narrowly defined situation, perhaps after excluding some alternatives because they violate legal

[3]The most recent survey, closest to this chapter, is Bossert and Weymark (2000). See also Sen (1977, 1979, 1986a), Blackorby et al. (1984), Lockwood (1984), d'Aspremont (1985), Moulin (1988), Roemer (1996), Mongin and d'Aspremont (1998).

or human rights, it is our task to evaluate the respective merits of its elements and to rank them from society's viewpoint. Any *ranking* (or *preference ordering*) of X is required to be *rational*, i.e. a complete and transitive binary relation over X:

Rationality of R:

$\forall\, x, y \in X, xRy$ *or* yRx, $\quad$ and $\forall\, x, y, z \in X$,

xRy and yRz $\quad$ implies xRz.

The set of all rankings one can define over X is denoted $\mathcal{R}$. For any $R \in \mathcal{R}$ and $x, y \in X$, xIy means indifference (xRy and yRx) and xPy strict preference (xRy and not yRx). The simplest ranking is the *trivial* one: $\forall\, x, y \in X, xIy$. In general, however, a ranking R may be a complicated object. Its handling is often facilitated if it admits of a faithful translation in the language of real numbers. In more formal terms, R is said to be *representable* by a numerical function u defined on X if and only if $\forall\, x, y \in X, xRy \Leftrightarrow u(x) \geq u(y)$. To avoid pre-committing the interpretation of u, we call it an *evaluation function.* To ensure its representation, we sometimes (when X is non denumerable and has a topological structure) assume in addition

Continuity of R:

$\forall\, x, y \in X,$ the sets $\{x \in X \mid xRy\}$ and $\{x \in X \mid yRx\}$ are closed in X.

In the sequel, any pair (x, i) will be called a *station* if it is an element of $X \times N$. Social evaluation has to rest somehow on information pertaining to this set, either directly or through individual evaluations. Any label $x \in X$ is assumed to convey directly a full description of all ethically relevant aspects of the social decision it designates, except for the other elements involved in the construction of the social ranking. On the other hand, the observer is assumed to be fully informed of individual evaluations by a real-valued function defined on $X \times N$ and called hereafter individual evaluation profile or, for short, *profile.* A typical profile is denoted U; if X is finite, it can also be thought of as an $|X| \times |N|$ matrix with generic element $U(x, i)$ lying at the intersection of row $U(x, \cdot) = U_x$ with column $U(\cdot, i) = U_i$. In any case, U_i will be called individual i's evaluation function or indicator, whereas U_x, the restriction of U to $\{x\}$, will be called an individual evaluation vector[4] of x, a point in *evaluation space* $\Re^N$, where $\Re^N$ stands for the real line.

Since we may want to accommodate every profile in a universal set, we define

$$\mathcal{U} = \{U \mid U : X \times N \to \Re\}.$$

We are now ready to define formally the main subject of this survey. Given any profile U in a subset[5] $\mathcal{D} \subset \mathcal{U}$, we are to recommend the social ranking of X that is best adapted to it. Following Sen (1970), who first fully articulated this concept, a *social welfare functional* (SWFL) is a map $F : \mathcal{D} \to \mathcal{R}$ with generic image $R_U = F(U)$. If x is ranked socially at least as high as y whenever the relevant profile is U, we shall write xR_Uy.

According to the context, various assumptions may apply to the domain $\mathcal{D}$ of F. Usually, they ensure that (1) $\mathcal{D}$ is not a singleton and (2) the set $\mathcal{H}(X, \mathcal{D}) = \{r \in \Re^N \mid \exists\, x \in X, \exists\, U \in \mathcal{D}$ such that $U_x = r\}$ fills the whole individual evaluation space, i.e. $\mathcal{H}(X, \mathcal{D}) = \Re^N$. The strongest one is:

Domain universality (UD): $\mathrm{D} = \mathrm{U}$.

But, in many cases this will be too demanding. A weaker assumption will then be used[6]:

Domain attainability (AD):
$\forall\, u, v, w \in \Re^N$,
$\exists\, x, y, z \in X, \exists\, U \in \mathcal{D}$ such that $U_x = u, U_y = v$ and $U_z = w$.

Interesting sufficient conditions for Domain attainability have been developed, adopting what we shall call the *single preference profile* approach to formal welfarism[7], and following the work

[4]In this chapter, vector inequalities are distinguished as follows: if $a = (a_1, \cdots, a_n)$ and $b = (b_1, \cdots, b_n)$, (1) '$a \leq b$' means 'for every i, $a_i \leq b_i$', (2) '$a \leq b$' means 'for every i, $a_i \leq b_i$ and there exists k such that $a_k < b_k$' and (3) '$a \ll b$' means 'for every i, $a_i < b_i$'.

[5]In this chapter, the notation '$\subset$' means 'is a subset of or equal to' and not 'is a strict subset of'.

[6]This condition is called "unrestricted individual utility profile" in d'Aspremont (1985) and "relative attainability" in d'Aspremont and Mongin (1997).

[7]Reference to the word "preference" is often omitted in the above expression. This practice is appropriate for the Arrovian social welfare function context, but it can be misleading in our SWFL context.

of Hammond (1976b), Kemp and Ng (1976), Parks (1976), Pollak (1979) and Roberts (1980b). To make the approach plausible, X is given a modicum of structure: for instance, Pollak (1979) assumes that X can be partitioned in separate subsets; within each subset, the alternatives differ only with respect to the allotment of one transferable private good among individuals, and the latter entertain self-oriented preferences with at least three distinct levels of satisfaction. Although individual preferences are assumed to remain fixed over the domain of F, every n-tuple of utility functions representing them is admitted. We call this a *Pollak domain*.

In some contexts, zero appears naturally as the greatest lower bound (resp. least upper bound) of every individual evaluation function so that $\mathcal{H}(X, \mathcal{D}) = \Re^N_+$ or $\Re^N_{++}$ (resp. $\Re^N_-$ or $\Re^N_{--}$). To illustrate, we rely on what can be called the pure income distribution model. Our n agents are assumed to be capable of consuming only one single good in non-negative amounts. As every individual's consumption set is $\Re_+$, society's possibility set is $\Re^N_+$. It is often assumed that each agent has an equal right to whatever quantity of good is available and that none can claim special treatment on account of his/her individual characteristics. In this setup, it seems natural to adopt a specific single preference profile, such that every agent is both self-oriented and insatiable. If we further assume that any n-tuple of utility representations consistent with the preference profile is relevant for our family of problems, we are facing an example of a Pollak domain, and $\mathcal{H}(X, \mathcal{D}) = \Re^N$. Yet, in the absence of risk and uncertainty, one can argue that measuring utility by consumption is the only appropriate convention, since it is the only one about which hard evidence can be adduced. Moreover, we attach importance to the units of measurement of consumption. If this argument is accepted, only one profile of identical evaluations is considered valid and $\mathcal{H}(X, \mathcal{D}) = \Re^N_+$, where $\mathcal{D}$ is a singleton, so that the SWFL is degenerate. The problem of ranking income distributions from a social viewpoint has elicited a whole body of literature, which is surveyed by Dutta (2002).

Among others, we shall be interested in SWFLs treating individuals according to the rank they occupy in the hierarchy of individual evaluation levels, a hierarchy which depends on both the profile at hand and the alternative being appraised. If $\mathcal{H}(X, \mathcal{D}) = \Re^N$, it will prove convenient to restrict some properties to the subset of well-ordered individual evaluation vectors.

For future reference, we register it here as $\mathcal{G}_N = \{w \in \Re^N \mid w_1 \leq w_2 \leq \cdots \leq w_{n-1} \leq w_n\}$.

2.2 Examples

Prior to elaborating further on the significance of the concepts used in the above definition, we proceed by describing some useful examples of SWFLs.

2.2.1 Imposed SWFLs

One can simply deny that the individual evaluation profile has any relevance for establishing a social ranking. The rest of the information concerning the alternatives, which is subsumed under the description of X, however minutely detailed it happens to be, is deemed sufficient to determine the social ranking. The latter is thus constant once X is given and the SWFL is said *imposed*. If our reader is persuaded by this extreme opinion, it is unlikely that he or she will be enthusiastic about the rest of this chapter, where profiles play a leading role.

2.2.2 Dictatorship

Another hardly more appealing family of SWFLs eschews any interpersonal comparison among the columns of any given profile: each one is based on an exogenously determined hierarchy among individuals, and the social ranking always reflects the relative individual evaluation of the agent enjoying the highest hierarchical rank, unless he or she values equally the alternatives under consideration. If this occurs, the relative social evaluation endorses the relative individual evaluation of the agent enjoying the next position, and so forth, until the subset is exhausted. Eventually, social indifference between two alternatives occurs only if they are equivalent from every agent's viewpoint. To each permutation of the set of agents, there corresponds an exogenous hierarchy among them and SWFL defined as above. Each of these $n!$ SWFLs is called *lexicographically dictatorial*. More generally, we shall say that an SWFL displays *weak dictatorship* if there exists an individual such that the social ranking always mimics his or her strict preference. This privilege cannot be based on talent or merit as it would become endogenous. This SWFL is characterized in Section 4.1.

2.2.3 Maximin and leximin

A way to introduce egalitarian consideration is to focus on the least favoured individual [see Kolm (1966) and Rawls (1971)] and, between two alternatives $x, y \in X$, to prefer strictly x to y whenever $\min_i U(x,i) > \min_i U(y,i)$. Based on this property, two SWFLs can be defined, corresponding to two ways to get a complete ordering for every profile U. One is the *maximin* principle requiring that, for any two alternatives $x, y \in X$, x be declared at least as good socially as y if and only if $\min_i U(x,i) \geq \min_i U(y,i)$. The other is the *leximin* rule (an expression coined by Sen). To define it, we associate with every given profile another one that is well ordered *for each alternative*: for each $x \in X$, we permute the elements of the row U_x and obtain $(U(x,i(1)), U(x,i(2)), \cdots, U(x,i(n)))$, where $i(k)$ is the individual who happens to be the kth-worst off at x (in formal terms, $i(\cdot)$ is a permutation such that $U(x,i(1)) \leq U(x,i(2)) \leq \cdots \leq U(x,i(n))$. Then xR_Uy according to the lexicon SWFL if and only if the lexicographically dictatorial SWFL, applied in the natural order $(1,2,\cdots,n)$ to the associated well-ordered profile, stipulates that x is at least as good as y socially. Characterizations of leximin are studied in Subsections 4.2.3 and 5.2.

More extreme egalitarian principles could be thought of: they would justify a sheer reduction in individual utility without benefit to anyone else on the ground that this pruning operation leads to smaller interpersonal utility differences. Although this cannot happen if the maximum or lexicon rules are adopted, both principles have been criticized because they recommend to give absolute priority to the slightest gain of the least advantaged agent over potentially very significant losses incurred by all other agents.

As a concluding observation, we note that the lexicon SWFL has meaning only if individual evaluation *levels* can be compared interpersonally. This is in complete contrast with our previous examples, and also with our next SWFL, which relies perhaps on the most popular social evaluation formula since they days when Bentham (1789) started propagating it.

2.2.4 Pure utilitarianism, weighted utilitarianism and relative utilitarianism

The SWFL is aid to be *purely utilitarian* if and only if $\forall x, y \in X, \forall U \in \mathcal{D}$,

$$xR_U y \Leftrightarrow \sum_{i=1}^{n} U(x,i) \geq \sum_{i=1}^{n} U(y,i).$$

Thus, $xR_U y$ if the sum of algebraic gains in individual evaluation is non-negative as society moves from y to x. We can also express the latter condition as requiring the total evaluation gain of the gainers to be at least equal to the total loss of the losers. Summing individual evaluation indicators or their first differences makes sense only if their units of measurement can be meaningfully compared across persons.

The family of rules called *weighted utilitarianism* also requires that the units of measurement of individual evaluation indicators be comparable with each other; as its qualifier suggests, it is parameterized by means of a vector of n individual weights, $\lambda \in \Re_+^N \setminus \{0\}$, and is such that, $\forall x, y \in X, \forall U \in \mathcal{D}$,

$$xR_U y \Leftrightarrow \sum_{i=1}^{n} \lambda_i U(x,i) \geq \sum_{i=1}^{n} \lambda_i U(y,i).$$

It is natural to assume all weights to be positive. If they are all equal, we are back to pure utilitarianism. Related characterizations can be found under Subsections 4.1, 4.2 and 4.3. Readers familiar with the theory of individual decision under risk and uncertainty will have noticed the close relation between expected utility and utilitarian principles. We elaborate on this topic in Subsection 5.1, where we also discuss relative utilitarianism, to which we turn next.

For this purpose, we restrict the SWFL domain $\mathcal{D}$ so that every individual evaluation function U_i always displays both a maximum and a minimum, denoted hereafter $\hat{u}_i$ and $\check{u}_i$, respectively. We also let every individual evaluation function in every profile in the SWFL domain undergo a positive affine transformation which is called in the sequel a *Kaplan normalization*[8]; in other words, to every $U \in \mathcal{D}$, we associate a normalized profile K_U defined on the set of stations as follows:

$$\forall U \in \mathcal{D}, \forall (x,i) \in X \times N,\ K_U(x,i) = \frac{U(x,i) - \check{u}_i}{\hat{u}_i - \check{u}_i}.$$

[8] Arrow (1963, p. 32) mentions that this normalization was suggested to him by Kaplan.

We are now in position to define *relative utilitarianism*: $\forall U \in \mathcal{D}, \forall x, y \in X$,

$$xR_U y \Leftrightarrow \sum_{i=1}^{n} K_U(x,i) \geq \sum_{i=1}^{n} K_U(y,i).$$

We have described in effect a two-stage procedure. As a result of the first-stage normalisation, the range of each $K_U(\cdot, i)$ extends from 0 to 1. At stage two, the utilitarian summation formula is applied to the first-stage output. Whether the units of the individual evaluation indicators belonging to any given profile are comparable or not is immaterial, because the normalisation process implies a new endogenous calibration.

Arrow criticizes relative utilitarianism for the fact that adding an alternative which everyone considers worst may play havoc at the top of the social ranking. This criticism seems to lose its cogency if the SWFL domain is based on a single preference profile and if there exists such a thing as the worst imaginable outcome from the viewpoint of every individual, which is an element of X even though it may involve an unfeasible exchange of individual characteristics.

More generally, as we hinted in Section 1, the utilitarian formula is consistent with a very wide array of ethical attitudes towards inequality that may be inherent to the description of alternatives in X, for instance when X consists of various allocations of private goods to a set of otherwise identical individuals. In this case, the degree of social inequality aversion is dictated to utilitarians by the degree of concavity of each individual evaluation indicator, i.e. by the decrease of successive first differences of each U_i with respect to income. The selection of a particular profile as ethically relevant in a given social conflict is thus a delicate matter, even if one is persuaded by relative utilitarianism. On the other hand, utilitarians have been aptly criticized for being insensitive to individual evaluation *levels* and their distribution, since only first differences with respect to this metric matter in their formulation. See e.g., Theorem 1.2 in Sen (1973, p. 20). We elaborate on this in subsection 5.1. The example we discuss next may be immune from the latter bias.

2.2.5 Weighted rank utilitarianism and the generalised Gini family

The family of SWFLs that goes by the name of weighted rank utilitarianism puts all individuals on the same footing, like the maximin principle and pure utilitarianism, both of which it encompasses. To parameters it, we rely on a vector λ of n non-negative weights; each weight is associated with a particular *rank* k and this is reflected in our generic notation λ_k. Hence, for weighted rank utilitarianism, there exists $\lambda \in \Re_+^N \setminus \{0\}$ such that $\forall\, x, y \in X, \forall\, U \in \mathcal{D}$,

$$xR_U y \Leftrightarrow \sum_{k=1}^{n} \lambda_k U(x, i(k)) \geq \sum_{k=1}^{n} \lambda_k U(y, i(k)).$$

In the above formula, $i(k)$ denotes again the name of the individual whose evaluation indicator is the kth-smallest for the social decision under study. Equity requires to treat any lower rank at least as well as any higher rank, by allocating to the former a weight at least equal to the weight allocated to the latter. By adding the clause $\lambda_1 > 0$, $\forall\, k \in N \setminus \{n\}$, $\lambda_k \geq \lambda_{k+1} \geq 0$, we satisfy this requirement and we define the *generalized Gini* family. The importance of this family of SWFLs is illustrated in our discussion of social decisions under risk and uncertainty (Subsection 5.1).

2.2.6 Nash's bargaining solution

Some aggregation rules exhibit a multiplicative form. An example is the *symmetric Nash bargaining solution* [following Nash (1950)], which is defined relative to some constant *status quo* point x_0 in X and is such that, $\forall\, x, y \in X, \forall\, U \in \mathcal{D}$ having $\forall\, i \in N$, $U(x, i) > U(x_0, i)$ and $U(y, i) > U(x_0, i)$,

$$xR_U y \Leftrightarrow \prod_{i=1}^{n} [U(x, i) - U(x_0, i)] \geq \prod_{i=1}^{n} [U(y, i) - U(x_0, i)].$$

We remark that our definition says nothing about the social ranking of alternatives weakly less preferred than x_0 by one or more individuals. The social ranking corresponding to the weighted version is not complete either; to wit: for some $\lambda \in \Re_+^N \setminus \{0\}$ and $x_0 \in X$ such that $\forall\, x, y \in X$,

$\forall U \in \mathcal{D}$ having $\forall i \in N$, $U(x,i) > U(x_0,i)$ and $U(y,i) > U(x_0,i)$,

$$xR_U y \Leftrightarrow \prod_{i=1}^{n}[U(x,i) - U(x_0,i)]^{\lambda_i} \geq \prod_{i=1}^{n}[U(y,i) - U(x_0,i)]^{\lambda_i}.$$

A characterization is offered in Subsection 4.3.

2.2.7 Borda's method of voting

Various methods of voting may be relied upon to define SWFLs. We give two examples. We begin with the *Borda method*, which cannot be used unless X is finite. To define it, we suppose first that $\mathcal{D}$ is such that every individual evaluation indicator represents a strict ordering on X. Thus, $\mathcal{D} \subset \{U \in \mathcal{U} \mid \forall i \in N, \forall x, y \in X, U(x,i) \neq U(y,i)\}$. In this case, we let $B(x,i)$ denote the number of alternatives of X which are less preferred than x by i for the given profile $U \in \mathcal{D}$. Formally,

$$\forall x \in X, \forall U \in \mathcal{D}, \forall i \in N, B(x,i) = \#\{y \in X \mid U(x,i) > U(y,i)\}.$$

In other words, $B(x,i)$ registers the number of victories of x when it is pitted successively against every other alternative. As we proceed to the social ranking, we shall maintain in the same spirit and put all individuals on the same footing. We shall say that x is ranked socially higher than y if and only if the total number of victories scored by x is greater than the corresponding number for y:

$$\forall x, y \in X, \forall U \in \mathcal{D},\ xR_U y \Leftrightarrow \sum_{i=1}^{n} B(x,i) \geq \sum_{i=1}^{n} B(y,i).$$

This voting method extends easily if the domain of the SWFL allows for individual indifference among alternatives. For any $x_0 \in X$, if there is no alternative indifferent with x_0, the definition of $B(x_0,i)$ is the same as above. If the indifference curve through x_0 consists of, say, k distinct alternatives $x_0, x_1, \cdots, x_{k-1}$ scoring by definition the same number g of victories, then we define for every h, $0 \leq h \leq k-1$, $B(xh,i) = g + \frac{k-1}{k}$.

We move next to the family of *generalised Borda methods*. Any member can be obtained by selecting an increasing transformation and by applying it to every $B(\cdot,i)$ for every individual and for every profile in $\mathcal{D}$. As before, a representation of the social ranking is obtained by summing the transformed numbers. All these voting rules share two features: (1) any two distinct evaluation functions representing the same ordering get mapped into the same individual representation, and (2) if any two alternatives are adjacent in any i's evaluation, the corresponding representation difference seems to be the result of a mechanical process which is foreign to equity considerations. Indeed, the social ranking is fully determined by the positions occupied by the alternatives in the individual rankings. The following example is tailored to criticize the Borda method, but it could have been fitted to criticize any other member of the generalised family. Suppose N consists of three selfish people who have an equal title to a unit cake. The set of alternatives is defined as some *finite* subset of the set of all non-negative triples such that the first individual gets less than one third, whereas the other two people share the balance evenly:

$$X \subset \{(x_1, x_2, x_3) \in \Re^3_+, \frac{1}{3} < x_2 = x_3 = \frac{1}{2}(1 - x_1)\}.$$

Following the Borda method, this is enough to conclude that xP_Uy if and only if $x_1 < y_1$, i.e. the poor bcome poorer and the rich richer, because they all have self-centered preferences. Although it may appear a decent aggregation procedure from the political viewpoint, the Borda method is a nonstarter in the social evaluation competition. Relative utilitarianism, which is not mechanically linked to intermediate positions in individual rankings, seems to fare better in this respect, because the Kaplan normalisation it implies leaves unchanged the concavity of the original individual evaluation functions.

2.2.8 Majority voting

According to the familiar *method of major voting*, an alternative is socially ranked above another one if the number of individual evaluations ranking the former strictly above the latter exceeds the number of strictly opposite evaluations. Formally, $\forall\, x, y \in X, \forall\, U \in \mathcal{D}, xR_Uy$ if and only if

$$\#\{i \in N \mid U(x,i) > U(y,i)\} \geq \#\{i \in N \mid U(y,i) > U(x,i)\}.$$

Ever since Condorcet's days it has been known that profiles must be selected with care if we want to be sure that the corresponding social rankings are transitive. As a rule, they are not the same as the social rankings obtained by the Borda method for the same profiles, unless $|X| = 2$.

However, the two social rankings coincide also in the situation we conjured up to criticize the implications of the Borda method. In conclusion, even though the method of majority voting is often considered superior ti is Borda rival as a voting procedure when it results in a transitive social ranking, it does not fare better as a tool for judging whether social outcomes are equitable.

2.3 Domain interpretation

The set of stations and the domain of definition of any SWFL are taken as a data set provided without deeper formal justification. Although this information is not without structure, it could prove very unwieldy. As we already indicated, the literature suggests several axioms mean to select, from an ethical point of view, the information which may matter and to delete unimportant details. They are partly motivated by the cost of gathering and processing information, and in particular, by the degree of precision deemed acceptable in the ethical observer's attempts to perform interpersonal evaluation comparisons. However, this selection process cannot be appraised without reference to the ethical interpretation of the data set: it depends on the ethical intuitions one may entertain about what can be morally relevant for the problem at hand, and this may be linked with views concerning the legitimate objectives of society and its domain of intervention, on the one hand, and each individual's private sphere of responsibility, on the other. General discussions of this topic may be found in Sen and Williams (1982) and in Sen (1997).

From one polar standpoint, individual evaluation functions are just representing individual preference relations over X, whatever they happen to be, self-oriented, altruistic or anti-social. This assumption is in agreement with the bulk of the modern theory of *positive* economics. Under ideal conditions, individual preferences can be estimated or even observed. One might wish to base an ethically agreeable social ranking on such a parsimonious information, that leaves simply no room for interpersonal comparisons of either welfare levels or welfare gains. As Arrow (1963, p. 112) points out, empirically distinguishable phenomena can be equated by our value judgments, whereas empirically undistinguishable states cannot be differentiated. Indeed, the pure individual preference interpretation of every profile U in $\mathcal{D}$ is implicit in Arrow's (1951) definition of the social welfare function. It shares with his other axioms the responsibility of his impossibility result.

Sen's (1970) definition of the SWFL is designed to circumvent this conclusion, by allowing greater flexibility in the selection of usable information, while maintaining a set of fully abstract alternatives. It does register interpersonal welfare comparisons, and the latter have at least potential ethical significance. Information of this nature may be indirectly based on other people's behavior, but Sen (1979) recognizes that one cannot rely on commonly accepted inference procedures. On the other hand, Arrow's views about empiricism are not so restrictive as the above judgment might lead some to suppose. Indeed, he adds that empirical experiments may be of an idealized type. In our context, he calls *extended sympathy* this source of evidence. In operational form, the most basic version of the corresponding judgment reads as follows: "it is better (in my judgment) to be myself in state x than to be you in state y" [Arrow (1963, p. 115)]. If this thought experiment is interpreted as an interpersonal exchange of characteristics which does not alter the identity of individuals, it requires a parallel extension of the set of possible alternatives. To be systematic and develop intuition about this kind of statement, it may be helpful to figure oneself behind some veil of ignorance such as those proposed by Vickrey (1945), Harsanyi (1955) and Rawls (1971), i.e. to pretend one does not know about one's personal traits and circumstances in actual life, while some more or less inscrutable impartial chance mechanism is about to allocate them to society's members.

Hammond (1991) requires the ethical observer to assume an even more elevated stance, as though he or she were capable of choosing not only who is to become a member of society, but also what endowment of individual characteristics must be allocated to any one member. This viewpoint involves a complete change of perspective. In Hammond's (1998) own words, "... comparisons of the utilities of different people to the chooser, rather than comparisons of different people's own utilities ...", are the ones which matter from his viewpoint. As common economic usage seems to have converged towards the latter interpretation of the word utility, we shall rather use the less specific word evaluation to fit either perspective.

When the description of the alternatives involves the allocation of idiosyncratic characteristics to every individual, the idealized experiment we have been discussing is of a highly subjective nature, and one must raise the question of the bound of this soul-searching method, assuming that the process of augmenting X reaches its own limit. Harsanyi (1955), Kolm (1972) and Hammond (1991) entertain the view of a unique limiting preference relation. After postulating a deterministic theory of individual preference formation, the first author went on to suggest that some kind of fundamental preference would indeed emerge as a common bedrock for all human beings, once they are stripped from their personal characteristics. Whether the postulated existence of such a Holy Grail can prevent a severe lack of unanimity among subjective experimentators, we leave our readers to decide after consulting the relevant literature.[9]

Independently of the conclusions reached about interpersonal comparability, we may be persuaded that each U_i is a representation of i's individual preferences. In this case, we say that the SWFL is *utility-based*. Nevertheless, profiles could be prevented from having a dominant influence on the social ranking. In this direction, the polar SWFLs are the imposed ones of example 2.2.1. In the opposite polar case, one can let profiles occupy center stage and eliminate completely the influence of whatever characteristics of the alternatives are not accounted for by the profile at hand. Hicks (1959) criticized[10] this viewpoint which he considered to be dominant among his fellow economists, and which he traced to the influence of A.C. Pigou's (1920) *Economics of Welfare*. Hicks created the world *welfarism* to designate this doctrine. Although he deals more explicitly with a multi profile approach, Sen (1979) is faithful to the same spirit, when he writes, "welfarism asserts that the goodness of a state of affairs depends ultimately only on the personal utilities in the respective states". He also stresses the limitations of welfarism, with reference to the SWFL context. In the *Tanner Lecture* delivered in 1979, Sen (1980) seems to have lost the hope to discover a minimally satisfactory welfarist SWFL, and he suggests a nonwelfarist approach that builds on some arguments put forth earlier by Rawls (1971). According to the welfare levels, as this would amount to meddling with basically private responsibilities. Instead, society should be concerned with the distribution of what he calls

[9] See Arrow (1977), Kaneko (1984), Broome (1993), Roberts (1995), Kolm (1996a) and Suzumura (1996).

[10] We are indebted to S.C. Kolm for pointing this out to us.

primary goods: basic liberties and "things that every rational man is presumed to want". Sen (1980) comes up with an alternative proposal. He agrees that people are responsible for their individual preferences; the latter are legitimately concerned with functioning's which can be achieved by consuming goods and by taking advantage of various social opportunities. Sen aptly remarks that people are not equally proficient at transforming goods and opportunities into functioning's: some can be very gifted and some others may be handicapped without being accountable for this state of affairs. Sen concludes that justice requires society to be primarily interested in the distribution of individual *capabilities*, *i.e.* sets of functioning opportunities.

We remark that anyone persuaded by these arguments, or for that matter, by Hammond's (1991) interpretation of individual evaluation, may still be interested in the formal SWFL approach. There does not seem to exist any good reason preventing us from reinterpreting every $U(x, i)$ as actually *ascribed* by the ethical observer to i if the outcome is x in consideration of what is desirable for society: for instance, it could be an index of the availability of primary goods, or an index measuring capabilities. This is why we call U_i an individual evaluation function instead of a utility function.

2.4 Some related concepts

SWFLs make up the focus of the present chapter. According to Sen's original definition, the social ranking R_U recommended by an SWFL is liable to change whenever there is some variation of the profile U given *a priori* as relevant for social evaluation. This is the only explicit independent variable. However, the SWFL setup can be extended in several directions that can be combined:

1. The set of individuals N may vary. This extension is mandatory if we want to evaluate social decisions influencing the size of society. We refer the reader to Blackorby et al. (2002).

2. The feasible set X may vary. In this case, the ambition of the ethical observer may stop short of uncovering a full social ranking, since he or she might simply look for a *solution*,

i.e. the subset of alternatives among which society ought to choose in consideration of any given pair (X, U) of some well-defined domain. The study of solutions is the main topic of axiomatic allocation theory[11]. In some cases, as in revealed consumer preference theory, the solution set coincides systematically with the set of alternatives which are best for some social ranking, and it may be possible to link properties of solutions to the properties of the social rankings associated with them. See for instance Young (1987) or Sprumong (1996).

3. Both X and some singled out alternative $x^0 \in X$ may be included as independent variables together with U. Given N, we have here the basic elements of cooperative bargaining theory. In Nash's (1950) tradition, X is interpreted as the set of outcomes corresponding to as many feasible social agreements, whereas x^0 stands for the outcome resulting from the lack of unanimous agreement; moreover, an extended version of welfarism is adopted, so that bargaining problems are entirely described by the respective images of X and x^0 in individual utility space. This kind of structure may be further expanded to describe the fallback position of every subset of N, and it can be found relevant for social evaluation exercises. See for instance Yaari (1978). The limitations of welfarism in the bargaining context are stressed by Roemer (1986, 1990).

3 Axioms and their use

3.1 Preliminary

It is one of our main tasks to define various features of SWFLs which may be considered socially desirable. Formally, these value judgments are stated as axioms and they are used singly or in combination to reject uninteresting SWFLs. It seems natural to distinguish them following a two-way classification.

A first criterion refers to the axiom content. Some axioms are concerned with separating formally superfluous details from potentially paramount information, whereas others are reflec-

[11]This is treated in this Handbook by Moulin (2002), Thomson (2002) and Fleurbaey and Maniquet (2002).

tions of ethical intuitions concerning equity or justice, and it is useful to sub distinguish whether they may apply to conflicting situations or not.

The other criterion has to do with the number of profiles taken in consideration to state the axiom. Interprofile axioms are distinct from intraprofile statements. The latter actually define an incomplete binary relation over X on the basis of features that any U could display. They would also have meaning out of the SWFL context. For instance, when x dominates y in every individual evaluation, we may want the social evaluation of x versus y to reflect the unanimous dominance relation. This axiom may well register an ethical requirement, but it cannot handle conflicting situations.

When a single SWFL meets a number of axioms, we say that it is characterized by them. Instead of discussing exclusively the merits and demerits of SWFLs with the help of examples, we shall interpret them, in the next section, as compromises between axiomatic properties and compare various characterisations.

3.2 Invariance axioms

Two seemingly distinct individual evaluation profiles may turn out to be equivalent from the SWFL viewpoint, once what is deemed irrelevant has been left out. Although the two profiles do not look the same in mathematical terms, they may be considered equivalent because after all they contain the same usable information from the ethical observer's standpoint. If this kind of analysis is pursued systematically, we obtain a binary relation on $\mathcal{D}$ that is reflexive, symmetric and transitive: this is defined as an equivalence relation; as such, it generates a partition of $\mathcal{D}$ that is reflexive, symmetric and transitive: this is defined as an equivalence relation; as such, it generates a partition of $\mathcal{D}$ consisting of various subsets called equivalence classes, and any two profiles belonging to the same class have the same image according to F. In other words, the social ranking associated with some profile is required to remain invariant if the original profile is replaced with an equivalent one.

Several types of reasoned value judgments, including equity judgments, may account for dropping some profile features as negligible details. However, in the SWFL literature, the name invariance axiom is usually applied to a principle meant to restrict the *measurability and comparability*properties of individual evaluation profiles. These properties make up what Sen has called the SWFL *informational basis.* We shall not attempt to define the italicized expressions we have just used. If the exact magnitude of every number listed in a given profile were to matter for the definition of the corresponding social ranking, the ethical observer would face a decision problem very much like the ones described in economic consumption theory, where the agent is to choose between bundles of goods measured in exact quantities. It is usually recognized that individual evaluation counts do not involve such a degree of precision. We shall discuss some useful distinctions in this respect under the measurability heading. As it is often believed that a lesser degree of precision is achieved in interpersonal evaluation comparisons than interpersonally, further distinctions are called for under the comparability heading.

Once the superfluous information carried by a profile has been pruned, what remains is a set of *meaningful statements*, a notion formally developed in the measurement literature, [e.g., Krantz et al. (1971)], and brought to bear to our debate by Bossert (1991). From the invariance viewpoint, and two profiles are equivalent if they convex the same set of meaningful statements. Thus, together with the description of X, the latter set *potentially* encapsulates every bit of information about individual evaluations that is necessary and sufficient to generate an appropriate social ranking of X. We have emphasized the word potentially because some other axioms introduced at a later stage may concur to eliminating excessive information.

As our ethical intuitions concerning equity and justice are hardly helpful for appraising directly the merits and demerits of invariance axioms, we shall attempt to do it indirectly, by describing the social rankings implied jointly by invariance and other axioms taken together. In the same respect, axiom incompatibilities and tradeoffs may prove quite useful. For instance, we may be justified in rejecting the original Arrovian informational basis because it cannot accommodate the anonymity axiom.

We shall also find it convenient to interpret the equivalence relation between any two profiles as the outcome of a (possibly multiple) functional composition operation. In other words, moving from any profile to another that is equivalent may be considered as operating an *invariance* *transformation*. As will become clear, all the relevant ones imply an n-tuple of strictly *increasing* transformations, because the sets of meaningful statements associated with two profiles are never declared equivalent whenever there is an individual whose evaluation functions do not represent the same ordering. The systematic study of such transformations antedates Bossert's approach; it is best explained by Roberts (1980a). Other important contributions [Sen (1977), 1986b), Moulin (1983), Basu (1983)] pertain also these conceptual issues. We shall elaborate on some of them at the end of this section.

Measurability and comparability issues are also discussed in other contexts such as bargaining theory [see Shapley (1969), Shapley and Shubik (1975) and Shubik (1982, pp. 92–98)] and axiomatic allocation theory [see Sprumont (1996) and Fleurbaey and Maniquet (1996)].

To conclude these remarks, we introduce a new piece of notation. For any two stations $(x, i), (y, j) \in X \times N$, and any $U \in \mathcal{D}$, we shall denote the first difference $U(x, i) - U(y, j) = \Delta_U(x, i; y, j)$.

3.2.1 Comparisons of evaluation levels

We begin with the presentation of the most restrictive invariance axiom, which is implicit in the original Arrovian formulation. Each individual evaluation function is assumed to be only a representation of the individual's preference relation. In the behaviorist approach that is consistent with this interpretation, an individual preference relation in combination with the maximizing assumption is simply a handy analytical device capable of describing succinctly behavior under a variety of constraints. The map associating the latter behavior with each constraint set is as much a primitive as the analytical structure meant to facilitate its description. In the same spirit, U_i must be thought of as a kind of shorthand presentation of the underlying preference relation. For any two alternatives $x, y \in X$, the only kind of meaningful statement contained in U_i is the relation "at least as great as" between $U(x, i)$ and $U(y, i)$ and vice versa. If another individual utility function $V_i : X \to \Re$ implies the same set of relations, it has the same implications as the original U_i from the standpoint of positive theory. Alternatively, we can require the sign of any first difference $\Delta_U(x, i; y, i)$ to be the same as the sign of $\Delta_V(x, i; y, i)$.

Suppose this is the case; then, we can tabulate side by side the values taken by both U_i and V_i for each alternative in X and look at the figures in one column as functionally related to the corresponding numbers in the adjacent column. Prior to elaborating formally on this idea, we denote $U_i(X)$ (resp. $V_i(X)$) the image set of X by U_i (resp. V_i). In other words, we let $U_i(X) = \{r \in \Re \mid \exists\, x \in X$ such that $U(x,i) = r\}$ and similarly for $V_i(X)$. We are now ready to define the transformation $\varphi_i : U_i(X) \to V_i(X)$ such that $\forall\, x \in X,\ V(x,i) = \varphi_i(U(x,i))$. In effect, we have so defined a composition operation we can also write $V_i = \varphi_i \circ U_i$, and one quickly realizes that φ_i must be (strictly) increasing if U_i and V_i are required to be equivalent from the positive viewpoint.

In the reasoning we just went through, we started with a pair of equivalent utility functions and our conclusion pertained to a transformation. This order can obviously be reversed: if we are given any U_i and a set of increasing transformations defined on $U_i(X)$ we can generate a set of equivalent utility functions. Thus, according to modern positive theory, individual behavior is consistent with the maximization of any member of a set of utility functions which can be obtained from each other through increasing transformations: the latter invariance is known as an *ordinality and noncomparability* property (or, in terms of meaningful statements, *intrapersonal level comparability*). This is our first invariance axiom:

> *Invariance with respect to individual increasing transformations* (Inv($\varphi_i(U_i)$)*:*
> $\forall\, U, V \in \mathcal{D}$, for every n-tuple of increasing functions $(\varphi_i)_{i \in N}$,
> $R_U = R_V$ if $\forall\, i \in N, \varphi_i$ is defined on $U_i(X)$ and $\forall\, x \in X, V(x,i) = \varphi_i(U(x,i))$.

It says that the social ranking is invariant if individual evaluation functions undergo possibly distinct increasing transformations. The set of meaningful statements contained in any profile consists only of n-tuples of individual statements of the form described above in the particular case of individual i. To convince our reader that our last axiom implies, as it should, an equivalence relation on $\mathcal{D}$, we define a new set Φ as follows: Φ consists of all n-tuples of *increasing* functions $(\varphi_i)_{i \in N}$ that can be defined on $\times_{i \in N} U_i(X)$ with range $\times_{i \in N} V_i(X)$, and such that $\forall\, i \in N, \forall\, x \in X,\ V(x,i) = \varphi_i(U(x,i))$. Thus, Φ is generated by testing successively all pairs of profiles U, V in $\mathcal{D}$. We are now in a position to adduce three arguments: (i) If the conditions defined in the axiom are met, there exists for each i an increasing transformation φ_i^{-1} that is defined on $V_i(X)$ and such that $U_i = \varphi_i^{-1}(V_i)$, so that the relation between U and V is symmetric; (ii) For every triple $U, V, W \in \mathcal{D}$, if $\forall\, i \in N$, φ_i (resp. ξ_i) is an increasing transformation defined on $U_i(X)$ with range $V_i(X)$ (resp. on $V_i(X)$ with range $W_i(X)$) and satisfying the last condition stated in the axiom, then the joint composition operation $(\varphi_i \circ \xi_i)_{i \in N}$ results in a third increasing transformation which is by definition an element of Φ, so that the relation on $\mathcal{D}$ is also transitive; (iii) Finally, reflexivity is implied by the other two properties. Thus we have an equivalence relation on $\mathcal{D}$.

Let us move next to less restrictive principles, by introducing comparability among individual evaluations. For this purpose, we shall interpret any profile as an expression of extended sympathy. In the simplest case, we content ourselves with *level comparability*, as the ethical observer supplies the missing link between the individual evaluation orderings in order to obtain no less and no more than a single ordering over $X \times N$. The corresponding set of meaningful statements may thus be limited to the following: for any two stations $(x, i), (y, j) \in X \times N$, the individual evaluation of the former is at least as great (as small) as the individual evaluation of the latter. Alternatively, we can register for every ordered pair of stations the sign of the first difference between their individual evaluation scores. We observe that we have so defined a superset of the set of meaningful statements that would be implied by our previous axiom $\text{Inv}(\varphi_i(U_i))$.

Suppose that any two profiles U and V happen to deliver the same set of meaningful statements according to our new principle, so that they are equivalent. We can then tabulate side by side the values taken by either profile for each station in $X \times N$. In effect, we are facing the graph of a transformation φ with domain $U(X, N) = \{r \in \Re \mid \exists\, (x, i) \in X \times N \textit{ such that } U(x, i) = r\}$ and similarly defined range $V(X, N)$; we observe that it must be increasing. Since the reciprocal is also true, we can conclude that we are measuring individual evaluations on a *common ordinal scale* (leading to what has been called a property of *co-ordinality*, or, in terms of meaningful statements, of *interpersonal level comparability*). We will use it as an axiom of

Invariance with respect to common increasing transformations (Inv($\varphi(U_i)$)*:*

$\forall\, U, V \in \mathcal{D},$ for every increasing function φ defined on $U(X, N)$,

$R_U = R_V$ if $\forall\,(x, i) \in X \times N, V(x, i) = \varphi(U(x, i))$.

As we compare this axiom with the previous one, we remark at once that any two profiles declared equivalent by Inv($\varphi(U_i)$) must also be declared equivalent by Inv($\varphi(U_i)$). Hence, the partition of $\mathcal{D}$ in equivalence classes generated by the former must be finer than the one generated by the latter. Given any profile $U \in \mathcal{D}$, the subset of profiles considered equivalent in the former case must be contained in the subset corresponding to INV($\varphi_i(U_i)$), notwithstanding the seemingly opposite assertion we just made about sets of meaningful statements. To get an idea of the magnitudes involved in the refinement process, let us consider an example: suppose $|X| = s$, a finite number. Suppose also that $\mathcal{D}$ consists of all the profiles in the universal domain made out of ns distinct numbers, so that indifferences never occur. Then, the finer partition consists of $(ns)!$ cells, whereas the coarser one consists of $(s!)^n$ subsets; these numbers are equal if $n = 1$, but the ratio of the former to the latter increases by a factor greater than $(n+1)^s$ each time we add an individual to an n-member society. An analogy between a profile and a set of n geographic maps may also help comparing Inv($\varphi(U_i)$) with Inv($\varphi_i(U_i)$). If the latter axiom is adhered to, any profile looks like n disconnected contour maps with ascending directions indicated, whereas the n countour maps would be fully connected in the former case.

3.2.2 Comparisons of evaluation differences

For usual geographic purposes, information about ascending directions would have to be supplemented at least by information about the gradient steepness. Similarly, having profiles measured on an ordinal scale might be deemed insufficient, and information about the sign of first differences in individual evaluation provided by our first two invariance axioms would have to be completed in a variety of manners to which we shall turn next. All of them involve first differences in evaluation counts. In the four following axioms, we find it simpler to assume the same degree of comparability between individual evaluation functions as within them, a feature shared with $\text{Inv}(\varphi(U_i))$.

For instance, the set of meaningful statements could be an exhaustive list of sentences of the following form: the net individual evaluation gain obtained by moving from (y, j) to (x, i) is at least as great (as small) as the net gain obtained by moving from (z, ℓ) to (w, k). In this case, each comparison involves two ordered pairs of stations, possibly all distinct, and the full list of $\Delta_U(x, i; y, j)$'s mrs up the range of an ordering representation over $(X \times N)^2$, which in turn includes the definition of a unique ordering on $(X \times N)$. This partition of $\mathcal{D}$ was introduced by Bossert (1991) under the name of *strong interpersonal difference comparability.*

If any two profiles U and V in $\mathcal{D}$ are equivalent in this sense, we may again observe that there exists a transformation $\psi : U(X, N) \to V(X, N)$ such that $V(x, i) = \psi(U(x, i))$ for every station $(x, i) \in X \times N$, and where in Basu's (1983) terminology, ψ is *first-difference preserving*: $\forall\, s_1, s_2, s_3, s_4 \in U(X, N), s_1 - s_2 \geq s_3 - s_4$ if and only if $\psi(s_1) - \psi(s_2) \geq \psi(s_3) - \psi(s_4)$. Moreover, we may note that the reciprocal statement is also true.

Alternatively, we may rephrase the equivalence condition between U and V in $\mathcal{D}$ as follows: there exists a domain $\Delta\, U = \{r \in \Re \mid \exists\, (x, i), (y, j) \in X \times N$ such that $\Delta_u(x, i; y, j) = r\}$, and an increasing function $\varphi : \Delta\, U \to \Re$ such that for any two stations $(x, i), (y, j) \in X \times N, \Delta_V(x, i; y, j) = \varphi(\Delta_u(x, i; y, j))$.

This possibility is almost trivial to prove: By assumption, we are given U and V in $\mathcal{D}$ and they are equivalent. Suppose first that there exists $\psi : U(X, N) \to V(X, N)$ which is first-difference preserving; then for any two stations $(x, i), (y, j) \in X \times N$, observe that φ is a well-defined increasing function. Suppose next that there exists $\varphi : \Delta\, U \to \Re$ as in the reciprocal statement to be proved; then we define $\psi : U(X, N) \to V(X, N)$ such that $V(x, i) = \psi(U(x, i))$ for every station $(x, i) \in X \times N$, and we observe immediately that for any two stations $(x, i), (y, j) \in X \times N$, $\Delta_V(x, i; y, j) = \varphi(\Delta_U(x, i; y, j)) = \psi(U(x, i)) - \psi(U(y, j))$. Since φ is increasing, ψ must be first-difference preserving.

We can now proceed by introducing formally the axiom, meant to capture strong interpersonal difference comparability.

Invariance with respect to common first-difference preserving transformations (Inv($\varphi(\Delta U)$))*:*
$\forall U, V \in \mathcal{D}$, for every increasing function φ defined on ΔU,
$R_U = R_V$ if $\forall (x,i),(y,j) \in X \times N, V(x,i) - V(y,j) = \varphi(U(x,i) - U(y,j))$.

Much more discriminating would be a set of meaningful statements delivered for any two stations as follows: $\Delta_U(x,i;y,j) = c$, for $c \in \Re$ and $(x,i),(y,j) \in X \times N$. In other words, the full list of ordered pairs of stations and the exact numerical magnitude of the first differences in evaluation counts would be of potential interest for the ethical observer. Thus, if any $U, V \in \mathcal{D}$ are equivalent according to this principle, we must have $\forall (x,i),(y,j) \in X \times N, \Delta_U(x,i;y,j) = \Delta_V(x,i;y,j)$. Obviously, the latter equality is not only necessary, but it is also sufficient to make sure that $U, V \in \mathcal{D}$ are equivalent.

As Bossert (1991) remarks, the same equality condition is in turn satisfied if and only if there exists some real number a such that $\forall (x,i) \in X \times N, V(x,i) = a + U(x,i)$. To prove necessity, we fix some $(y,j) \in X \times N$, and let $V(y,j) - U(y,j) = a$. We observe that, by assumption, $\forall (x,i) \in X \times N, V(x,i) - V(y,j) = U(x,i) - U(y,j)$, so that $V(x,i) - U(x,i) = V(y,j) - U(y,j) = a$, as was to be proved. In conclusion, V is equivalent to U if and only if they are the same up to a change of origin, a property also known as *translation-scale measurability* with *full comparability*. This is the same as having individual evaluation counts measured in common natural units, while their common origin is arbitrary. The corresponding axiom can be stated formally as follows:

Invariance with respect to common changes of origin (Inv($a + U_i$))*:*
$\forall U, V \in \mathcal{D}, \forall a \in \Re, R_U = R_V$ if $\forall (x,i) \in X \times N, V(x,i) = a + U(x,i)$.

We turn next to an informational basis that lies in between the two last ones. According to this intermediate principle, the set of meaningful statements consists of sentences formulated as follows: given any four stations $(x,i),(y,j),(w,k),(z,\ell) \in X \times N$, such that $\Delta_U(w,k;z,\ell) \neq 0$, the ratio of $\Delta_U(x,i;y,j)$ to $\Delta_u(w,k;z,\ell)$ amounts exactly to c, for some $c \in \Re$. Any two profiles $U, V \in \mathcal{D}$ are thus equivalent for this principle if and only if $\forall (x,i),(y,j),(w,k),(z,\ell) \in X \times N$, the ratio of their corresponding first differences is equal, assuming that $\mathrm{sgn}\{\Delta_U(w,k;z,\ell)\} =$

$\text{sgn}\{\Delta_V(w, k; z, \ell)\} \neq 0$.

When the last assumption is satisfied for two pairs of stations, say (w, k) and (z, ℓ), this condition is in turn satisfied, if and only if V is a *positive affine transformation* of U, or, in other words, there exist two real numbers a, b with $b > 0$, such that $\forall\,(x, i) \in X \times N, V(x, i) = a + bU(x, i)$. To prove necessity, we avail ourselves of $\Delta_U(w, k; z, \ell)$ and $\Delta_V(w, k; z\ell)$ and we let b stand for the ratio of the latter to the former. By assumption, $b > 0$. Moreover, we fix some $(y, j) \in X \times N$, and let $V(y, j) - bU(y, j) = a$. We observe that, by assumption, $\forall\,(x, i) \in X \times N$, $(V(x, i) - V(y, j))/(\Delta_V(w; k; z, \ell)) = (U(x, i) - U(y, j))/(\Delta_U(w, k; z, \ell))$, so that $V(x, i) - V(y, j) = b(U(x, i) - U(y, j))$ and $V(x, i) = bU(x, i) + (V(y, j) - bU(y, j)) = a + bU(x, i)$, as was to be proved.

Under these assumptions, profiles U and V are said to be measured on a *cardinal scale* with *full comparability.* The origin and the scale of individual evaluation counts are both common and arbitrary. The axiom (corresponding to a property called *co-cardinality*) can be stated formally as follows:

Invariance with respect to common positive affine transformations (Inv$(a + bU_i)$)*:*
$\forall\, U, V \in \mathcal{D}, \forall\, a, b \in \Re$, with $b > 0$,
$R_U = R_V$ if $\forall\,(x, i) \in X \times N, V(x, i) = a + bU(x, i)$.

It is easy to see that any positive affine transformation is first-difference preserving (so that Inv$(\varphi(\Delta_U))$ implies Inv$(a + bU_i)$) whereas the reciprocal statement is not always true. Think for instance of $U_i(X) = \{1, 2, 4\}$ and $V_i(X) = \{1, 2, 5\}$. Basu (1983) provides an interesting domain condition that is sufficient for equivalence, viz. that $U_i(X)$ be dense on a nontrivial connected subset of $\Re$. This condition is met, for example, if the ethical observer is concerned with a set of alternatives and a domain of profiles consistent with the definition of a family of Arrow-Debreu economies, but it cannot be satisfied if the SWFL is defined on an unrestricted domain [see also Fishburn, Marcus-Roberts and Roberts (1988) and Fishburn and Roberts (1989)]. The following table recapitulates the five invariance properties described so far. Arrows indicate inclusion relations between sets of meaningful statements. To list the latter in summary form, we have used the language of first differences.

Transformation	Meaningful statement
$a + U_i$	$\Delta_u(x, i; y, j)$
$\downarrow$	
$a + bU_i$	$\Delta_u(x, i; y, j)/\Delta_U(w, k; z, \ell)$
$\downarrow$	
$\varphi(\Delta\, U)$	$\mathrm{sgn}\{\Delta_U(x, i; y, j) - \Delta_U(w, k; z, \ell)\}$
$\downarrow$	
$\varphi(U_i)$	$\mathrm{sgn}\{\Delta_U(x, i; y, j)\}$
$\downarrow$	
$\varphi_i(U_i)$	$(\mathrm{sgn}\{\Delta_U(x, i; y, i)\})_{i \in N}$.

Except for our very first invariance axiom, we have relied so far on interpersonal evaluation comparability. This feature may be hardly desirable for anyone wishing to remain as close as possible to hard data and commonly accepted inference procedures. Therefore, we shall sample next some weakenings of interpersonal comparability. For this purpose, let us reconsider how $\mathrm{Inv}(\varphi_i(U_i))$ relates to $\mathrm{Inv}(\varphi(U_i))$: instead of treating a profile as a single function defined on $X \times N$, we look at it as a n-tuple of individual evaluation functions defined on X; instead of comparing the respective evaluations of any two stations, meaningful statements are restricted to comparing them from the viewpoint of one single individual at a time. Similarly, the invariance property is no longer based on a single transformation defined on a typical image set $U(X, N)$, but it calls for a n-tuple of transformations, each one being applied to an individual image set $U_i(X)$.

It is straightforward to remove formally interpersonal comparability from $\mathrm{Inv}(a + U_i)$ by relying on the same pattern of reasoning. In this case, profiles are said to be *translation-scale measurable*, without additional qualifier, and we obtain

Invariance with respect to individual changes of origin ($\mathrm{Inv}(a_i + U_i)$)*:*
$\forall\, U, V \in \mathcal{D}, \forall\, (a_i) \in \Re^N, R_U = R_V$ if $\forall\, (x, i) \in X \times N, V(x, i) = a_i + U(x, i)$.

In this case, the unit of measurement cannot be changed, and a natural interpretation would hold it interpersonally comparable.

In contrast, as we turn to cardinality with full comparability, we notice that the latter feature may be erased thoroughly, but the process need not be so radical. We have pointed out in due time that the set of meaningful statements corresponding to Inv$(a + bU_i)$ pertains to ratios of first differences in individual evaluations involving up to four distinct individuals. If we decide to throw out of the list any ratio involving more than one individual at a time, the end result is a new axiom prescribing

> *Invariance with respect to individual positive affine transformations* (Inv$(a_i + b_iU_i)$)*:*
> $\forall\, U, V \in \mathcal{D}, \forall\, (a_i) \in \Re^N, \forall\, (b_i) \in \Re^N_{++}$,
> $R_U = R_V$ if $\forall\, (x, i) \in X \times N, V(x, i) = a_i + b_iU(x, i)$.

Indeed, we observe that the origin and the scale of each individual evaluation function are arbitrary. In other words, each evaluation is measured on a *cardinal scale without interpersonal comparability.* If the SWFL is utility-based, this can perhaps be defended when uncertainty is made explicit and individuals are assumed to display von Neumann-Morgenstern rationality. Moreover, as we reconsider the role of individual preference profiles in the positive theory of rational social interactions, we remember that one of its commonly accepted tools is the Nash noncooperative equilibrium concept based on mixed strategies, whose conclusions are invariant under independent positive affine transformations of each individual utility function in the profile on which equilibrium is based. The same invariance property was invoked by Nash (1950) to justify his bargaining solution, and it could be used as an interesting SWFL informational restriction.

Alternatively, we may content ourselves with a less severe shearing operation and consider significant any ratio involving up to two individuals, provided each first difference refers to one single person at a time. We get

Invariance with respect to common rescaling and individual change of origin (Inv($a_i + bU_i$))*:*
$\forall\, U, V \in \mathcal{D}, \forall\,(a_i) \in \Re^N, \forall\, b \in \Re_{++}$,
$R_U = R_V$ if $\forall\,(x, i) \in X \times N, V(x, i) = a_i + bU(x, i)$.

In this case, each individual evaluation is measured on a *cardinal scale*; there is *unit comparability* among individuals, but level comparability is excluded. This axiom may look less directly intuitive than either Inv($a + bU_i$) or Inv($a_i + b_iU_i$); it can appeal to someone who wishes to concede as little as possible to interpersonal comparability.

3.2.3 Comparisons of evaluation indicator ratios

Among the invariance axioms we have been discussing, none does restrict *a priori* the sign of individual evaluation counts. However, there may be interesting contexts where they are all positive for each profile in the SWFL domain, while they refer to a worst (possibly unattainable) alternative which is used as a natural origin. Or they might all be of negative sign by reference to a best (possibly unattainable) alternative used in similar fashion; in this case, referring to cost figures seems to be closer to usual conventions. For instance, a firm manager may judge the relative effort level of two salesmen by looking at their sales ratio, the same manager may also rely on the cost ratio to evaluate the relative performance of two workers who deliver the same quantity of output. Suppose now we want to compare two profit centers, one of which is in the red whereas the other one earns a positive amount. In this case, it would seem odd to compute the ratio of two figures of opposite sign. We present next a set of meaningful statements that is hardly intuitive on domains which mix up positive and negative signs. A typical domain that meets the requirement of being homogeneous with respect to sign will be denoted $\mathcal{D}^*$.

The new set of meaningful statements consists of propositions revealing the exact magnitude of every well-defined ratio of individual evaluation counts. Thus, U and V are equivalent for this principle if and only if we have $(V(x, i))/V(y, j)) = (U(x, i))/(U(y, j))$ for any two stations (x, i) and (y, j), with $\mathrm{sgn}(U(y, j)) = \mathrm{sgn}(V(y,)) \neq 0$, and the latter condition holds true if and only if there exists $b \in \Re_{++}$ such that for every $(x, i) \in X \times N, V(x, i) = bU(x, i)$. We omit the easy proof, which is analogous to the previous ones. In this case, individual evaluations are said to be measured on a *ratio-scale with full comparability.* Tsui and Weymark (1997) note that percentage changes in individual evaluations might be substituted for ratios in the above developments without altering the conclusion. The latter is captured formally in our axiom of

Invariance with respect to positive similarity transformation (Inv(bU_i))*:*
$\forall\, U, V \in \mathcal{D}^*, \forall\, b \in \Re_{++}, R_U = R_V$ if $\forall\,(x,i) \in X \times N, V(x,i) = bU(x,i)$.

It should be noted that anyone wishing to rely exclusively on the transformation approach to invariance may be interested in similarity transformations over an extended domain and reject the sign restriction inherent in $\mathcal{D}^*$.

We may consider also *ratio-scale measurability without comparability.* The relevant set of meaningful statements discloses the exact magnitude of every well-defined ratio $(U(x,i))/(U(y,i))$ for each individual separately. Therefore, any two profiles U and V in $\mathcal{D}^*$ are equivalent for this principle if and only if $\forall\, i \in N$, we have $(V(x,i)) + (V(y,i)) = (U(x,i))/(U(y,i))$ for any two alternatives x and y, with $\mathrm{sgn}(U(y,i)) = \mathrm{sgn}(V(y,i)) \neq 0$, and the latter condition holds true if and only if there exists $b_i \in \Re_{++}$ such that $V(x,i) = b_iU(x,i)$ for every $x \in X$. The corresponding axiom is

Invariance with respect to individual positive similarity transformations (Inv(b_iU_i))*:*
$\forall\, U, V \in \mathcal{D}^*, \forall\,(b_i) \in \Re^N_{++}, R_U = R_V$ if $\forall\,(x,i) \in X \times N, V(x,i)b_iU(x,i)$.

It should be stressed that our new axiom retains implicitly a piece of information that is interpersonally comparable. Indeed, the origin cannot be changed arbitrarily, and a zero count can be attributed to the evaluation of the worst (or perhaps the best) alternative by every individual, an alternative which is not necessarily the same for everybody. As we noted for Inv($b\,U_i$), the sign restriction of $\mathcal{D}^*$ might be rejected by someone who does not care about intuitive meaningful statements.

3.2.4 Other invariance axioms

To illustrate a mixed-invariance principle, we combine the restrictions imposed by both $\mathrm{Inv}(a_i + b\,U_i)$ and $\mathrm{Inv}(\varphi(U_i))$, to wit

Mixed invariance ($\mathrm{Inv}((a_i + b\,U_i)\&\varphi(U_i))$)*:*
$\forall\, U, V \in \mathcal{D}, \forall\,(a_i) \in \Re^N, \forall\, b \in \Re_{++}$,
for every increasing function φ defined on $U(X, N)$,
$R_U = R_V$ if $\forall (x, i) \in X \times N, V(x, i) = \varphi(U(x, i)) = a_i + bU(x, i)$.

Interpersonal comparisons of levels and gains are allowed, but at most two individuals are involved in any comparison of utility *gains*, in contrast with $\mathrm{Inv}(a + bU_i)$. The corresponding set of meaningful statements is obtained by operating the union of the two sets implied by the parent axioms.[12]

More exotic invariance axioms could be added to our list. The following example has little interest by itself: it is included only to illustrate some boundaries of the invariance transformation approach. At first glance, the opening statement looks symmetrical to $\mathrm{Inv}(a_i + bU_i)$.

Invariance with respect to individual rescaling and common change of origin
($\mathrm{Inv}(a + b_i U_i)$) :
$\forall\, U, V \in \mathcal{D}, \forall\, a \in \Re, \forall\,(b_i) \in \Re^N_{++}$,
$R_U = R_V$ if $\forall\,(x, i) \in X \times N, V(x, i) = a + b_i U(x, i)$.

Dixit (1980) points out rightly that the inverse transformation is not necessarily increasing, so that the relation between U and V implied by the transformation fails to be symmetric. It follows that profiles can hardly be considered equivalent, unless the equivalence criterion is redefined in a way that is more roundabout than in the previous cases. Thus, U and V may be declared *directly* equivalent if either the condition stated in the axiom holds or if, instead, the last clause is replaced with $U(x, i) = a + b_i V(x, i)$. Furthermore, they may be declared

[12]This is to be distinguished from the requirement that the SWFL should satisfy more than one invariance axiom.

equivalent if they are either directly equivalent, or there exists in $\mathcal{D}$ a sequence $U, U', U'', \cdots, V$ such that any two successive profiles in the sequence are directly equivalent. If this stricture is accepted, we obtain a well-defined partition of $\mathcal{D}$, which is consistent with $\text{Inv}(a + b_i U_i)$, but we failed to derive its implications in terms of meaningful statements.

On occasion, we shall rely on weakened versions of some invariance axioms; we introduce them when we need. At this stage, we feel that we have been already taxing too much our reader's imagination. Let us recapitulate their implication relations. As we compared $\text{Inv}(\varphi(U_i))$ with $\text{Inv}(\varphi_i(U_i))$, we already pointed out that the latter axiom is stronger than the former, whereas the set of meaningful statements delivered by any profile is always richer in the former case than in the latter; hence, the ethical observer is more susceptible to being misled by potentially slippery information. This inclusion relation between sets of meaningful statements associated with pairs of invariance axioms is displayed in Figure 1 by means of arrows pointing towards the smaller set: it is not a complete relation, but nevertheless it proves quite useful.

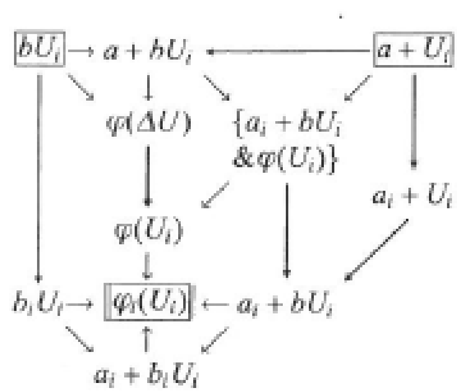

Figure 1: Invariance transformations

We make at once the following remarks: the Arrovian axiom $\text{Inv}(\varphi_i(U_i))$ is dominated by every single one of our remaining ten axioms, and there are two undominated and thus non-comparable axioms, viz. $\text{Inv}(a+U_i)$ and $\text{Inv}(bU_i)$. One can conjure up an axiom of independence with respect to identity transformations which would dominate both of them, and which is implicitly built in the SWFL definition, but this would hardly help. Anyone interested in characterizing a specific SWFL may be tempted to choose an invariance axiom that lies as close as possible to $\text{Inv}(\varphi_i(U_i))$, whereas someone interested in uncovering the implications of a richer information structure will attempt to maintain unchanged a set of basic axioms while combining them successively with weaker and weaker invariance axioms. By moving closer and closer either to $\text{Inv}(a + U_i)$ or to $\text{Inv}(bU_i)$, one is bound to admit a larger and larger set of SWFLs consistent with the axioms already listed. Moreover, it proves illuminating to define ordered subsets or types of axioms and to study the tradeoffs between weaker axioms of one type and stronger axioms of another type.

3.3 Other information-filtering axioms and formal welfarism

3.3.1 Pareto indifference, independence and neutrality

If information gathering and processing is costly, one is tempted to focus attention exclusively on what seems mot important. Independently of invariance principles, our next sequence of axioms prescribes the deletion of what may appear as unnecessary details in the formulation of social evaluation judgments. This is what the axioms called Pareto indifference, Binary independence, Neutrality, Continuity and Separability are meant to capture. We introduce also an important property of SWFLs we call formal welfarism; furthermore, we are in position to characterize the set of SWFLs satisfying this property.

We begin by restricting the analyst's attention exclusively to the content of the relevant individual evaluation profile. This can be expressed as a single-profile property, known as the Pareto indifference principle; it says that any two alternatives are socially indifferent if they are represented by the same image in the evaluation space:

Pareto indifference (PI):

$\forall\, U \in \mathcal{D}, \forall\, x, y \in X,\ xI_U y$ if $U_x = U_y$.

Thus, as long as $U_x = U_y$, the particular consequences involved in the description of the alternatives cannot affect the social ranking.

We introduce next another method for eliminating superfluous details when dealing with a subproblem. Suppose indeed that the social ranking we are interested in is restricted to a particular pair of alternatives; does this judgment require any information regarding other alternatives which could be feasible? A negative answer is best expressed as an inter profile property. Two profiles U, V and two alternatives x, y are involved; we further assume that their respective images in the evaluation space are the same; in other words, $U_x = V_x, U_y = V_y$. Then, we require the social rankings R_U and R_V to be the same with respect to the pair (x, y), independently of the discrepancies existing between U and V over the other alternatives.

Binary independence (BIN):
$\forall V \in \mathcal{D}, \forall x, y \in X,\ xR_Vy$ if $\exists U \in \mathcal{D}$ such that $V_x = U_x, V_y = U_y$ and xR_Uy.

In other words, if the restrictions of two profiles to a given pair of alternatives cannot be distinguished, the relative social ranking of the two alternatives under focus must be unique. Without some similar principle, the exact definition of X or the exact application of the definition would always be problematic. Directly adapted from the Arrovian axiom of Independence of irrelevant alternatives, it has become a workhorse of the bulk of the SWFL literature. Even though it does not imply any invariance axiom, unlike its Arrovian model, it has been criticized as being too constraining. For instance, it makes the SWFL less suited for the discussion of rights; games in various forms provide more appropriate tools of analysis. See, for instance, Gaertner et al. (1992) and Suzumura (2002). In contexts closer to the SWFL, several authors recently managed to derive social rankings despite the fact that they weaken Binary independence: we shall report briefly about their characterization results at the end of this chapter.

We turn next to the discussion of welfarism and its relation with the two properties we just described. Drawing our inspiration from Blackorby et al. (1990), we consider first the image of any profile $U \in \mathcal{D}$ in $\Re^N$ interpreted as the evaluation space, and we denote it $U(x)$. Formally:

$$U(X) = \{u \in \Re^N \mid \exists x \in X, u = U_x\}.$$

Letting $\mathcal{R}^*_U$ denote the set of orderings on $U(X)$, with typical element R^*_U, we shall say that F

displays *profile-dependent welfarism* if and only if $\forall\, U \in \mathcal{D}, \exists\, R_U^* \in \mathcal{R}_U^*$ such that

$$\forall\, u, v \in U(X), \forall\, x, y \in X, \langle u = U_x \text{ and } v = U_y\rangle \Rightarrow \langle u R_U^* v \Leftrightarrow R_U y\rangle.$$

The interpretation seems clear: the association between R_U and R_U^* is one to one; to study the restriction of F to U, one can equivalently rely on R_U or on R_U^*, since each version may be translated into the other without ambiguity. In particular, R_U is then said to be *fully recoverable* from R_U^*.

We proceed with the analysis of Pareto indifference, which has some remarkable implications by itself. We stressed already one of them: viz., for any given profile, neither the intrinsic characteristics of an alternative nor its name can have any bearing on its social ranking: all that matters is its individual evaluation vector; this idea may also be re-expressed formally by a property called

Intraprofile neutrality (IAN):
$\forall\, U \in \mathcal{D}, \forall\, x, y, x', y' \in X$ such that $U_x = U_{x'}$ and $U_y = U_{y'}$, $\;xR_U y$ iff $x'R_U y'$.

Intraprofile neutrality implies Pareto indifference, because we may choose $x' = y$ and $y' = x$ in the last statement. On the other hand, it is also implied by Pareto indifference, due to the transitivity of any social ranking R_U. Thus, we are led to the following

Theorem 3.1 *For all $\mathcal{D} \subset \mathcal{U}$, Pareto indifference is equivalent to Intraprofile neutrality.*

To clarify the relation between Pareto indifference and profile-dependent welfarism, it is convenient and always possible to define on $U(X)$ a binary relation $\hat{R}_U^*$ as follows:

$$\forall\, u, v \in U(X), \quad u\hat{R}_U^* v \text{ iff } \exists\, x, y \in X \text{ such that } u = U_x, v = U_y \text{ and } xR_U y.$$

We observe at once that $\hat{R}_U^*$ lacks one of the characteristics of the relation R_U^* used to define profile-dependent welfarism. Indeed, $xP_U y$ can hold true despite the fact that $u = U_x, v = U_y$ and $u\hat{I}_U^* v$; in other words, recoverability is not guaranteed by $\hat{R}_U^*$. Let us elaborate on this. The relation $\hat{R}_U^*$ inherits reflexivity from R_U. Suppose we reject Pareto indifference; then, there must exist some x and y such that $xP_U y$ and $U_x = U_y$; in this case, by reflexivity, we register $U_y \hat{I}_U^* U_x$, whereas xP_Uy, so that R_U cannot be fully recoverable from $\hat{R}^*_U$. Partial recoverability is however warranted in the following sense: $u\hat{P}^*_Uv$, $U_x = u$ and $U_y = v$ together imply xP_Uy. Indeed, any conclusion to the contrary would immediately contradict $u\hat{P}^*_Uv$. Failing Pareto indifference, the bluntness of $\hat{R}^*_U$ proves unfortunately infectious:

Lemma 3.2 *For any $U \in \mathcal{D}, \forall\, x, y, z \in X,$ such that xP_UyP_Uz, $U_x = U_z$ implies $U_x\hat{I}^*_UU_y\hat{I}^*_UU_z$.*

Proof: By assumption and by definition, we obtain $U_x\hat{R}^*_UU_y\hat{R}^*_UU_z$, whereas $U_x\hat{I}^*_UU_z$, by reflexivity, so that $U_x\hat{I}^*_UU_y$. By a similar argument, $U_y\hat{I}^*_UU_z$. ■

Having shown that Pareto indifference is necessary for profile dependent welfarism, we can prove it to be also sufficient:

Theorem 3.3 *For all $\mathcal{D} \subset \mathcal{U}$, Pareto indifference is equivalent to profile-dependent welfarism.*

Proof: Suppose woe have PI and $u\hat{R}^*_Uv$ for some $u, v \in U(X)$; we want to show that xR_Uy if $U_x = u$ and $U_y = v$. By definition, $u\hat{R}^*_Uv$ only if there exist $x', y' \in X$ with $x'R_Uy'$, $U_{x'} = u = U_x$ and $U_{y'} = v = U_y$. There remains to apply the previous theorem. ■

Of course, the ordering R^*_U depends in general on U, the profile selected in $\mathcal{D}$. For instance, R^*_U can agree with utilitarianism for some profiles in $\mathcal{D}$, whereas it reflects the lexicon principle for all other profiles in $\mathcal{D}$. If the SWFL is utility-based, profile-dependent welfarism seems to agree with what Hicks (1959) probably meant when he coined the word welfarism. By dropping profile dependence, Sen developed a more useful but demanding concept. As he was dealing exclusively with utility-based SWFLs, he relied on the same word as Hicks. We shall take the liberty to twist Sen's definition of welfarism because we believe the scope of the SWFL tool is wider: we shall say that an SWFL is welfarist in a formal sense if the relative goodness of two states of affairs can be *entirely* judged by comparing the goodness of their respective individual *evaluation* vectors, independently from the other aspects of the profile at hand. In this case, each social ranking R_U on X can be construed as emanating from a unique ordering defined on the relevant subset of the individual evaluation space. The various rankings on X remind us of spokes, whereas the unique ordering to which they are linked is analogous to a hub. This property will prove invaluable in the sequel because it is easier to characterise an ordering or to study its representation than to deal directly with a multi profile structure. What axioms are necessary and sufficient to obtain formal welfarism? This is the question which we address next.

As a preamble, we define a *Social welfare ordering* (SWO); for lack of a less ambiguous word, this is simply how we designate an ordering on the evaluation space ($\Re^N, \Re^N_+$ or $\Re^N_{++}$ as the case may be). We shall further say that F displays *formal welfarism* if and only if we can associate with F an SWO R^* with the following property:

$$\forall\, u, v \in \mathcal{H}(X, \mathcal{D}),\ \forall\, x, y \in X, \forall\, U \in \mathcal{D},$$
$$\langle u = U_x \text{ and } v = U_y \rangle \Rightarrow \langle u R^* v \Leftrightarrow x R_U y \rangle.$$

In this case, we say that R^* is the *formally welfarist associate* of F. We remark that, whatever the profile U considered, R_U is fully recoverable from R^*; under these circumstances, one also has $x I_U y$ if $u I^* v$ and $x P_U y$ if $u P^* v$: there is thus a one-to-one link between the SWFL and its SWO associate under formal welfarism. For example, the purely utilitarian SWO (i.e. the formally welfarist associate of the purely utilitarian SWFL) is defined by

$$\forall\, u, v \in \Re^N,\ u R^* v \Leftrightarrow \sum_{i=1}^{n} u_i \geq \sum_{i=1}^{n} v_i.$$

Similarly, the formally welfarist associate of the leximin principle (resp. of its inequitable mirror-image called *leximax*) is defined by:

$$\forall\, u, v \in \Re^N, u P^* v \text{ iff } \exists\, k \in N, \quad \text{such that } u_{i(k)} > v_{i(k)}, \text{ and}$$
$$\forall\, h \in N, h < k (\text{resp. } h > k), u_{i(h)} = v_{i(h)},$$

where $u_{i(\cdot)}$ (resp. $v_{i(\cdot)}$) denotes the non-decreasing re-ordering of the vector u (resp. v) according to the permutation $i(\cdot)$ such that $i(k)$ is the kth-worst off individual: $u_{i(1)} \leq u_{i(2)} \leq \cdots \leq u_{i(n)}$ (resp. $v_{i(1)} \leq v_{i(2)} \leq \cdots \leq v_{i(n)}$).

Suppose we are given some SWFL F and we want to find out whether it displays formal welfarism. As a first step, we associate with F a reflexive binary relation $\hat{R}^*$ defined on $\mathcal{H}(X, \mathcal{D})$ as follows[13]: $\forall\, u, v \in \mathcal{H}(X, \mathcal{D})$,

$$u \hat{R}^* v \Leftrightarrow \langle \exists\, U \in \mathcal{D}, \exists\, x, y \in X \text{ such that } u = U_x, v = U_y \text{ and } x R_U y \rangle.$$

[13]Alternatively, $u \hat{R}^* v \Leftrightarrow \exists\, U \in \mathcal{D}$ such that $u, v \in U(X)$ and $u \hat{R}^*_U v$.

We have to make sure that every R_U can be retrieved from $\hat{R}^*$. For this purpose, it is convenient to define another SWFL property, which is often used as an axiom in its own right. In the literature, it is referred to as

Strong Neutrality (SN) :
$\forall\, U \in \mathcal{D}, \forall\, x, y \in X,$
xR_Uy if there are $x', y' \in X$ and $V \in \mathcal{D}$ such that $V_{x'} = U_x, V_{y'} = U_y, x'R_Vy'$.

We can state at once an obvious remark, which does not require any domain condition pertaining to F:

Theorem 3.4 *Formal welfarism implies Strong neutrality.*

However, for the characterisation coming next, a rich domain condition (e.g., the domain is Pollak) is maintained (see Section 2.1):

Theorem 3.5 (Formal welfarism) *Suppose F satisfies Domain universality or attainability. Then it satisfies Strong neutrality if and only if it displays formal welfarism, whereas R^* is an ordering over $\mathcal{H}(X, \mathcal{D} = \Re^N$.*

Proof: One direction follows from the preceding theorem. For the other direction, we first prove full recoverability. Take any $u, v \in \mathcal{H}(X, \mathcal{D})$, any $U \in \mathcal{D}$, any $x, y \in X$, such that $u = U_x, v = U_y$ and $u\hat{R}^*v$. We want to show that SN implies xR_Uy. Now, by definition of $\hat{R}^*$, there exists $V \in \mathcal{D}$, $x', y' \in X$ such that $u = V_{x'}$, $u = V_{y'}$ and $x'R_Vy'$. Applying SN, we obtain xR_Uy. We have indicated earlier that $\hat{R}^*$ is reflexive over $\mathcal{H}(X, \mathcal{D})$. To establish the completeness and transitivity of $\hat{R}^*$, we select any $u, v, w \in \Re^N$. As F satisfies UD or AD, there exists a profile $U \in \mathcal{U}$, and $x, y, z \in X$ such that $u = U_x, v = U_y, w = U_z$. Because R_U is an ordering, xR_Uy or yR_Uz so that $u\hat{R}^*_v$ or $v\hat{R}^*u$. We may w.l.o.g. assume xR_Uy and yR_Uz, and obtain xR_Uz, so that $u\hat{R}^*v$, $v\hat{R}^*w$ and $u\hat{R}^*w$ by applying the definition of $\hat{R}^*$. ■

Whatever the domain of definition of the SWFL, Strong neutrality implies Intraprofile neutrality and thus, Pareto indifference; to see this, simply add to the statement of Strong neutrality the following clause: $U = V$. Another useful implication of Strong neutrality may be obtained by adding the orthogonal requirement $x = x', y = y'$ to its statement; without surprise, we obtain the axiom of Binary independence:

Theorem 3.6 *For all $\mathcal{D} \subset \mathcal{U}$, Strong neutrality implies both Pareto indifference and Binary independence.*

If the domain of F is universal, a reciprocal statement can also be proved. It has been called welfarism theorem by Sen (1977), who derived it by strengthening a result of d'Aspremont and Gevers (1977).

Theorem 3.7 *If $\mathcal{D} = \mathcal{U}$, Pareto indifference and Binary independence in conjunction imply Strong neutrality, and hence, formal welfarism.*

Proof: Suppose we are given two profiles $U, V \in \mathcal{U}$, two pairs (x, y) and (x', y') of alternatives in X, and two vectors $u, v \in \Re^N$ such that $U_x = V_{x'} = u$, $U_y = V_{y'} = v$, and xR_Uy. We have to show that $x'R_Vy'$. Since we always assume that $|X| \geq 3$, we may choose $r \in X \setminus \{y, y'\}$, with r a third alternative if the two pairs coincide, and construct profiles U^1, U^2 and U^3 such that $U^1_x = U^1_r = u$, $U^1_y = v$, $U^2_r = r$, $U^2_y = U^2_{y'} = v$, $U^3_{x'} = U^3_r = u$ and $U^3_{y'} = v$. Then letting "$\rightrightarrows$" mean "implies by BIN", whereas "$\Rightarrow$" means "implies by PI and transitivity", we get: $xR_Uy \rightrightarrows xR_{U^1}y \Rightarrow rR_{U^1}y \rightrightarrows rR_{U^2}y \Rightarrow rR_{U^2}y' \rightrightarrows rR_{U^3}y' \Rightarrow x'R_{U^3}y' \rightrightarrows x'R_Vy'$. We have thus SN. Formal welfarism follows from Theorem (3.5). ■

3.3.2 Formal welfarism and invariance properties

All axioms we have been studying in this and the previous subsection have in common the goal of filtering information. Whenever some of them taken together imply formal welfarism, the associated SWO properties are inherited from axioms pertaining to the parent SWFL. In most cases, there is a one-to-one link between the parent axiom and the corresponding SWO property. To designate the latter, we shall use an asterisk flanking their parent's name, be it in full or in shortened version. In particular, each invariance axiom stated so far for SWFLs translates (in Section 3.2) as a relation between *pairs of points* in $\Re^N$. As the proof of their equivalence is almost trivial, we shall content ourselves to quote the ten main axioms, relying only on their short names. Only one of them requires some care. This is invariance with respect to common first-difference preserving transformation, the domain of definition of which matters a lot, as Basu (1983) points out. Towards defining it properly, we first pick any point $u \in \Re^N$ and let $T(u) = \{t \in \Re \mid \exists\, i \in N$ such that $u_i = t\}$. For any pair $u, v \in \Re^N$, the appropriate domain is defined as $Y(u,v) = T(u) \cup T(v)$.

- $\mathrm{Inv}^*(\varphi(\Delta u)) : \forall\, u, v \in \Re^N$, and any first-difference preserving function $\psi : Y(u,v) \to \Re$, i.e., $\forall\, s_1, s_2, s_3, s_4 \in Y(u,v), \psi(s_1) - \psi(s_2) \leq \psi(s_3) - \psi(s_4) \Leftrightarrow s_1 - s_2 \leq s_3 - s_4$,

$$uR^*v \Leftrightarrow (\psi(u_1), \cdots, \psi(u_n))R^*(\psi(v_1), \cdots, \psi(v_n)).$$

  We remark that we could in particular choose $s_3 = s_4 \in Y(u,v)$ in the above statement, and observe that ψ must be increasing. The next imd on our list is the axiom of invariance with respect to common increasing transformation:

- $\mathrm{Inv}^*(\varphi(u_i)) : \forall\, u, v \in \Re^N$, and for any increasing function $\varphi : Y(u,v) \to \Re$,

$$uR^*v \Leftrightarrow (\varphi(u_1), \cdots, \varphi(u_n))R^*(\varphi(v_1), \cdots, \varphi(v_n)).$$

  We observe that our explicit reference to $Y(u,v)$ in the last statement is in fact superfluous; for simplicity, $\Re$ could have been used as domain of definition of φ. The next invariance axiom lend itself to a similar remark; to economize on notation, we state the following version:

- $\mathrm{Inv}^*(\varphi_i(u_i))$: for any n-tuple of increasing functions $(\varphi_i)_{i \in N}$ defined on $\Re$, and $\forall\, u, v \in \Re^N$,

$$uR^*v \Leftrightarrow (\varphi_1(u_1), \cdots, \varphi_n(u_n))R^*(\varphi_1(v_1), \cdots, \varphi_n)).$$

  The remaining invariance axioms can be simply stated as follows:

- $\mathrm{Inv}^*(a + u_i) : \forall\, a \in \Re$, and $\forall\, u, v \in \Re^N$,

$$uR^*v \Leftrightarrow (a + u_1, \cdots, a + u_n)R^*(a + v_1, \cdots, a + v_n).$$

- $\text{Inv}^*(a + bu_i) : \forall\, a, b \in \Re$, with $b > 0$, and $\forall\, u, v \in \Re^N$,

$$uR^*v \Leftrightarrow (a + bu_1, \cdots, a + bu_n)R^*(a + bv_1, \cdots, a + bv_n).$$

- $\text{Inv}^*(a_i + u_i) : \forall\, (a_i) \in \Re^N$, and $\forall\, u, v \in \Re^N$,

$$uR^*v \Leftrightarrow (a_1 + u_1, \cdots, a_n + u_n)R^*(a_1 + v_1, \cdots, a_n + v_n).$$

- $\text{Inv}^*(a_i + b_iu_i) : \forall\, (a_i) \in \Re^N$, $\forall\, (b_i) \in \Re^N_{++}$, and $\forall\, u, v \in \Re^N$,

$$uR^*v \Leftrightarrow (a_1 + b_1u_1, \cdots, a_n + b_nu_n)R^*(a_1 + b_1v_1, \cdots, a_n + b_nv_n).$$

- $\text{Inv}^*((a_i + bu_i)\&\varphi(u_i)) : \forall\, u, v, u', v' \in \Re^N$, $\forall\, (a_i) \in \Re^N, \forall\, b \in \Re_{++}$, and for every increasing function φ defined on $\Re$ such that, $\forall\, i \in N, u'_i = a_i + bu_i = \varphi(u_i)$ and $v'_i = a_i + vb_i = \varphi(v_i)$,

$$u'R^*v' \Leftrightarrow uR^*v.$$

- $\text{Inv}^*(bu_i)$: $\forall\, b \in \Re_{++}$ and $\forall\, u, v \in \Re^N_{++}$,

$$uR^*v \Leftrightarrow buR^*bv.$$

- $\text{Inv}^*(b_iu_i) : \forall\, (b_i) \in \Re^N_{++}$ and $\forall\, u, v \in \Re^N_{++}$,

$$uR^*v \Leftrightarrow (b_1u_1, \cdots, b_nu_n)R^*(b_1v_1, \cdots, b_nv_n).$$

Notice that the last two axioms have been, for simplicity, defined on a subset of $\Re^N$ with no change of sign, i.e. $\Re^N_{++}$.

3.3.3 Separability, continuity and convexity properties

We turn next to another instance of parsimonious attitude towards information: what do we do when some individuals are completely unconcerned by the issues at stake, so that their evaluation function remains constant over X? Can the social ranking be affected by the level of their constant evaluation count? If we answer in the negative, we express it as an inter profile statement known as Separability axiom:

Separability (SE): $\forall U, V \in \mathcal{D}$,
$R_U = R_V$ if $\exists M \subset N$ such that $\forall i \in M, U_i = V_i$
whereas, $\forall j \in N \setminus M,\ \forall x, y \in X,\ U(x, j) = U(y, j)$ and $V(x, j) = V(y, j)$.

If we accept this idea, we need not worry too much about the precise definition of N, as long as all interested individuals are included. We notice here some analogy with the axiom of Binary independence which allows the decision-maker not to worry about the precise definition of X. We remark also that $\mathrm{Inv}(a_i + u_i)$ or any stronger invariance axiom implies Separability.

The SWO translation, directly implied from Separability and formal welfarism, is as follows:

Separability (SE*):* $\forall u, v, u', v' \in \Re^N$,
$uR^*v \Leftrightarrow u'R^*v'$, if $\exists M \subset N$ such that $\forall i \in M, u_i = v_i$ and $u'_i = v'_i$
whereas, $\forall j \in N \setminus M, u_j = u'_j$ and $v_j = v'_j$.

We introduce next the SWO version of a weakening of Separability which proves interesting in the sequel. It restricts this property to the set of well-ordered vectors $\mathcal{G}_N$ defined at the end of Section 2.1:

$\mathcal{G}_N$-*Separability** $(\mathcal{G}_n - \mathrm{SE}^*)$: $\forall u, v, u', v' \in \mathcal{G}_N$,
$uR^*v \Leftrightarrow u'R^*v'$, if $\exists M \subset N$ such that $\forall i \in M, u_i = v_i$ and $u'_i = v'_i$
whereas, $\forall j \in N \setminus M, u_j = u'_j$ and $v_j = v'_j$.

At this stage, we turn to an axiom relieving potential anxiety about slight measurement errors in evaluation profiles: we shall require such errors to have a limited bearing on the corresponding social rankings. This is explicitly an interprofile concern, which requires defining a distance between numerical functions over a given domain. We state the following as an example:

Continuity (of F) (C):

$\forall\, x, y \in X, \forall\, U^0 \in \mathcal{D}$ and any sequence $(U^\ell)_{\ell=1}^\infty \subset \mathcal{D}$ converging pointwise to U^0,

if $\forall\, \ell \geq 1, xR_{U^1}y,$ then $xR_{U^0}y$.

In the literature, continuity seems to have always been used in a formally welfarist framework. Hence, it is defined directly as a condition on the social welfare ordering R^*:

Continuity (of R*) (C*):*

$\forall\, v \in \Re^N,$ the sets $\{u \in \Re^N \mid uR^*v\}$ and $\{u \in \Re^N \mid vR^*u\}$ are closed in $\Re^N$.

We have the following:

Lemma 3.8 (Continuity) *Suppose an SWFL F satisfies Domain universality, Binary independence, Pareto indifference and Continuity: then F has a continuous formally welfarist associate R^*.*

Proof: Suppose R^* does not satisfy C^*. Then, for some $v \in \Re^N$, say, the set $\{u \in \Re^N : uR^*v\}$ is not closed, and it is possible to find a sequence $(u^\ell)_{\ell=1}^\infty$ in $\Re^N$, converging to some u^0, such that $u^\ell R^* v$ for $\ell \geq 1$, but vP^*u^0. Now, we may choose $x, y \in X$ and, by UD, construct a sequence $(U^\ell)_{\ell=1}^\infty$ such that $U_x^\ell = u^\ell$ and $U_y^\ell = v$, for $\ell \geq 1$, and converging pointwise to some profile U^* such that $U_x^0 = u^0$ and $U_y^0 = v$. Then, by the last theorem, $\forall\, \ell \geq 1, xR_{U^1}y$, but $yP_{U^0}x$, that is, F does not satisfy C. ∎

In this result, we ensure formal welfarism through Domain universality. If we wanted to use attainability instead, then the continuity axiom on F would have to be modified, and some topological structure put on X, in order to derive the continuity of R^*.

We conclude this section with another regularity axiom. According to Convexity, society is never hurt if it moves toward a better alternative, wherever it may have to stop along the way on feasibility grounds. IN the SWO context, the formal version is a straightforward application of the general mathematical definition:

*Convexity**:

$\forall \alpha \in \Re_{++}, \alpha > 1, \forall u, v, w \in \Re^N$,

$\langle w = \alpha u + (1-\alpha)v$ and $uR^* v \langle \Rightarrow wR^* v$.

The SWFL version is hardly less transparent:

Convexity:

$\forall U \in \mathcal{D}, \forall x, y, z \in X, \forall \alpha \in \Re_{++}, \alpha < 1$,

$\langle xR_U y$ and $U_z = \alpha U_x + (1 - \alpha U_y\rangle \Rightarrow \langle zR_U y\rangle$.

3.3.4 Alternative approaches to formal welfarism

The importance of formal welfarism makes it worthwhile to study the relationships of Strong neutrality with seemingly weaker axioms.

What if, instead of permuting names among individuals, we permute labels among alternatives? One can surmise that this arbitrary operation should not affect the social ranking of the alternatives once they have been properly relabelled. The corresponding axiom is obtained by weakening Strong neutrality or by strengthening Intraprofile neutrality.

Interprofile neutrality (IRN):

Consider any permutation $\sigma : X \to X$ and any $U, V \in \mathcal{D}$,

such that $\forall x, y \in X, V_y = U_x$ whenever $y = \sigma(x)$;

then, $\forall x, y \in X, xR_U y$ if $\sigma(x)R_V\sigma(y)$.

Our readers are invited to check again the examples provided in Section 2.2: the imposed SWFLs are the only ones violating Interprofile neutrality, as they violate both Pareto indifference and Intraprofile neutrality. Although Interprofile neutrality is necessary to obtain formal welfarism, it fails to be sufficient.

The Borda method (Subsection 2.2.7) illustrates this point: it satisfies Interprofile neutrality, but it violates Binary independence and, hence, Strong neutrality. Suppose $X = \{x, y, z\}$, $N = \{i, j\}$ and $U(x, i) = U(y, j) = 1$, whereas $U(y, i) = U(x, j) = 2$. Then the Borda social ranking of x and y and the relation it induces between $U_x = (1, 2) = u$ and $U_y = (2, 1) = v$

are highly influenced by U_z, the evaluation vector of the third alternative. Indeed, xI_Uy if both x and y either Pareto-dominate z or are Pareto-dominated by z, or if U is such that there is no Pareto domination at all. If z is Pareto-dominated by x (resp. y) alone we get xP_Uy (resp. yP_Ux). This pattern gets reversed if Pareto domination goes the other way.

We can easily build on the above example and define a Borda SWFL with domain $\mathcal{D} = \{U, V\}$, for which uP^*_Uv coexists with vP^*_Vu, consistently with profile-dependent welfarism, but in opposition to the Binary independence requirement. Let us inquire about the consequences of this state of affairs for $\hat{R}^*$ as defined in Subsection 3.3.1: we immediately realise that $u\hat{I}^*v$ by definition of $\hat{R}^*$. In this case, neither R_U nor R_V could be recovered from $\hat{R}^*$, be it directly or not, because $\hat{R}^*$ lacks discriminating power.

The following hypothetical example suggests that Interprofile neutrality might be strengthened without going all the way to Strong neutrality. Suppose X is a large set and we are asked to check whether for some pair of profiles (U, V), for some permutation σ of X and for every pair of alternatives (x, y), with $y = \sigma(x)$, $V_y = U_x$. In other words, we wish to find out whether the first sentence in the statement of Interprofile neutrality holds true. Imagine now that $V_y = U_x$ for all pairs but one. We surmise that the SWFL would seem abnormal if it were to prescribe both a utilitarian R_U and a lexicon R_V. As a possible strengthening of Interprofile neutrality, the following axiom could be relied on for some integer m, where $2 < m < |X|$:

m-ary Neutrality :
Consider any two subsets $A, B \subset X$, such that $|A| = |B| = m$,
any bijection $\sigma : A \to B$,
and any $U, V \in \mathcal{D}$, such that $\forall\, x \in A, \forall\, y \in B, V_y = U_x$ whenever $y = \sigma(x)$;
then, $\forall\, x, y \in A, xR_Uy$ if $\sigma(x)R_V\sigma(y)$.

Which number m should we select? If $m_1 < m_2$, m_1-ary Neutrality obviously implies m_2-ary Neutrality, whatever the domain of the SWFL may be. On the other hand, Strong neutrality and Interprofile neutrality correspond respectively to $m = 2$ and to $m = |X|$, two values we chose to exclude *a priori* to prevent ambiguity.

Theorem 3.9 *For all $\mathcal{D} \subset \mathcal{U}$, for all integers m_1, m_2 such that $2 < m_1 < m_2 < |X|$, Strong neutrality implies m_1-ary Neutrality, which implies m_2-ary Neutrality, which implies Interprofile neutrality, which implies Intraprofile neutrality.*

If the domain of the SWFL is universal, a partial reciprocal statement is also true:

Theorem 3.10 *If $\mathcal{D} = \mathcal{U}$, if $2 < m_1 < m_2 < |X|$, m_2-ary Neutrality implies m_1-ary Neutrality, which implies Strong neutrality.*

As Strong neutrality implies Binary independence m-ary Neutrality implies m-ary Independence, a property we introduce next:

> *m-ary Independence:*
> Consider any subset $A \subset X$, such that $|A| = m$,
> and any $U, V \in \mathcal{D}$, such that $\forall\, x \in A, V_x = U_x$;
> then, $\forall\, x, y \in A$, $xR_U y$ if $xR_V y$.

It is interesting to note that the last two theorems have an analogue dealing with m-ary Independence, the method of proof being analogous. See d'Aspremont and Gevers (1977), who draw their inspiration from Blau (1971).

Theorem 3.11 *For all $\mathcal{D} \subset \mathcal{U}$, for all integers m_1, m_2 such that $2 < m_1 < m_2 < |X|$, Binary independence implies m_1-ary Independence, which implies m_2-ary Independence.*

As before, a partial reciprocal statement holds true if the SWFL domain is universal:

Theorem 3.12 *If $\mathcal{D} = \mathcal{U}$, if $2 < m_1 < m_2 < |X|$, m_2-ary Independence implies m_1-ary Independence, which implies Binary Independence.*

3.4 Pareto dominance principles and weak welfarism

At this stage, it is useful to bring in the familiar Pareto dominance principles. Their popularity among economists and the criticisms recently levied against them warrant some discussion. We start with a formal introduction, before woe movie on to interpretive comments.

The least controversial axiom is known as Weak Pareto; it grants a higher social rank to an alternative beating another one on every evaluation count:

Weak Pareto (WP) :

$\forall\, U \in \mathcal{D}, \forall\, x, y \in X,\; xP_U y$ if $U_x \gg U_y$.

Most economists seem to agree with a stronger dominance principle, called hereafter Strong Pareto:

Strong Pareto (SP) :

$\forall\, U \in \mathcal{D}, \forall\, x, y \in X,\; xR_U y$ if $U_x \geq U_y$ and $xP_Y y$ if $U_x > U_y$.

The latter property implies Pareto indifference. We retained this formulation for the sake of simplicity, and not on grounds of logical necessity. On the other hand, Pareto indifference is also implied by Weak Pareto, Continuity and Universal domain taken together:

Theorem 3.13 *If an SWFL satisfies Weak Pareto, Continuity and Domain universality, it satisfies also Pareto indifference.*

Proof: Consider any $U^0 \in \mathcal{U}$, any $x, y \in X$, such that $U_x^0 = U_y^0$; by Universal domain, we can form a sequence of profiles $(U^\ell)_{\ell=1}^{\infty} \subset \mathcal{U}$ converging point wise to U^0, and such that $\forall\, \ell, U_x^\ell \gg U_x^0$, whereas $U_y^\ell = U_y^0$. By Weak Pareto we obtain $\forall\, \ell \geq 1,\; xR_{U^1} y$. By Continuity, $xR_{U^0} y$ follows. Since y can be interchanged with x in the above argument, we observe that $xI_{U^0} y$. ■

If we adopt the formally welfarist structure of the previous section, our weak version translates as follows:

Weak Pareto (WP*):*

$\forall\, u, v \in \Re^N$, if $u \gg v$ then $uP^* v$.

On the other hand, as Pareto indifference is subsumed by welfarism, the strong axiom becomes

Strict Pareto (SP*):*

$\forall\, u, v \in \Re^N$, if $u > v$ then $uP^* v$.

To discuss this set of axioms, we shall have to refer to the SWFL purpose and to the meaning imparted to individual evaluation vectors. In what follows, we shall maintain the *value judgment* interpretation of the Pareto principles, which tell us how influential individual evaluations must be when they are in agreement about the relative merits of any two alternatives; indeed, *under these circumstances*, the description of the alternatives cannot affect the social ranking, and the ethical observer's own influence is banished. As it transpires, we move quite far from the spirit of imposed SWFLs.

In the sequel, we have to distinguish whether the SWFL we analyze is utility-based or not. If the SWFL is utility-based, we need distinguish again the constitutional design problem from the social evaluation process. We realize that we can accept every unanimous individual preference profile at face value in the former case, whereas we cannot in the latter.

If the SWFL is meant to be used as a utility-based tool for social evaluation, the ethical observer should make sure that all the individuals potentially concerned with the decisions are represented in the unanimous appraisal process; they should also be reasonably cognizant of the issues at stake. Moreover, all individual preferences had best be self-centered; for instance, one cannot consider as socially acceptable the unequal division of the fruits of a common property if all the owners have a valid title to an equal share and one of them has either altruistic or anti-social preferences, so that an equal split happens to be Pareto-dominated from the positive view point. From this simple example, we remark that although social choice theory may operate formally from given individual preferences, it is sometimes felt appropriate to prune or lander them informally prior to aggregating them into a social preference. Thus, a utility-based interpretation of the domain of definition of an SWFL can be maintained at the cost of reducing its scope.

We turn next to the alternative interpretation of our informational basis: what happens if an ethical observer adopts a non-welfarist theory of the good and the corresponding individual evaluation vector U_x dominates some competitor U_y? It seems quite sensible to apply a Pareto-like dominance principle, and to declare that xP_Uy. However, one could perhaps make an exception when every individual preference relation (possibly after laundering) points in the opposite direction. In this conflicting case, one can surmise that the ethical observer should give in, out of respect for the exercise of the free will at least when there is no conflict among individuals. In effect, this could be achieved at the informal preparation stage of the aggregation exercise by reinterpreting the set of possible alternatives X as the preference-based Pareto set, so that the SWFL would be defined for a fixed preference profile at least if it is to be utility-based. In another approach, suggested by the work of Kelsey (1987), the SWFL should be enlarged to accommodate two criteria per individual instead of one: both a non-welfarist criterion and some preference representation would be formally taken into account by the ethical observer.

Be this as it may, one can also imagine to obtain a complete social ranking on the basis of a sequential reasoning of lexicographic nature, whether the informational setup is utility-based or not. For instance, an ethical observer having to choose between any two alternatives x and y could begin to reason as any utilitarian: he or she would add up individual evaluations for each alternative. If one of them dominates, it is selected, and the process stops. If utilitarian sums turn out to be equal, then one can invoke another set of arguments, possibly based on individual evaluation counts too, but also possibly on other variables. If individual evaluation vectors happen to coincide fully, one can also apply the same idea and rely on other variables to justify a strict social ranking. In a remarkable paper, Roberts (1980a) managed to articulate fruitfully such a sequential reasoning: his approach is based on an SWO R^*, defined over $\mathcal{H}(X, D) = \Re^N, \Re^N_+$ or $\Re^N_{++}$ as in the previous section, so that his construct is related to formal welfarism; however, it does not require Pareto indifference, so that we have to content ourselves with weak recoverability. As a substitute for Pareto indifference, the weak Pareto principle is an essential input in this theory; and either minimal invariance axiom $\text{Inv}(a+u_i)$ or $\text{Inv}(bu_i)$ is also invoked. We shall say that F displays *weak-welfarism* (short for *formal* weak-welfarism) if and only if we can associate with F an SWO R^* with the following property:

$$\forall\, u, v \in \mathcal{H}(X, D), \forall\, x, y \in X, \forall\, U \in \mathcal{D}, \langle u = U_x \text{ and } v = U_y \rangle \Rightarrow \langle uP^*v \Rightarrow xP_U y \rangle.$$

As the last statement implies, $xP_U y$ is not consistent with vP^*u, whereas nothing can be inferred from uI^*v, and we say that R_U is only *weakly recoverable* from R^*. In keeping with our definition of weak welfarism, the trivial SWO, according to which all individual evaluation vectors are socially indifferent, can be associated with every SWFL. By itself, the weakly welfarist structure seems thus at first unpromising. In particular, starred SWO properties can no longer be mechanically associated with parent SWFL axioms as under formal welfarism.

Yet, as the following theorem suggests, the prospects for weak welfarism are better than one might think. Its key proposition is essentially based on Theorem I in Roberts (1980a):

Theorem 3.14 (Weak welfarism) *If F has $\mathcal{H}(X, \mathcal{D}) = \Re^N$ (resp. $\mathcal{H}(X, \mathcal{D}) = \Re^N_{++}$) and if it satisfies Binary independence, Weak Pareto and Inv$(a + U_i)$ (resp. Inv(bU_i)), there exists a unique SWO R^* satisfying Continuity*, Weak Pareto*, Inv*$(a + u_i)$ (resp. Inv*(bu_i)) and weak-welfarism associated with F.*

Two remarks are in order:

1. Originally, weak welfarism was derived by introducing a technical condition of "weak-continuity" [see Roberts (1980a), and the correction by Hammond (1999)]. Since these proposed weak-continuity conditions are implied by each of the more transparent invariance conditions we require F to satisfy alternatively, we have simply articulated them in the above presentation.

2. Under the assumptions of the theorem, weak welfarism captures only the first and foremost step in the lexicographic sequential reasoning we sketched earlier in this section: however, the loss is limited due to Weak Pareto*. Uniqueness further dissipates ambiguity. Moreover, as the ones mentioned in the theorem, *all* our invariance axioms can be translated without being weakened.

Proof: Omitted, except for the uniqueness and invariance statements. To establish uniqueness, we reduce absurdum the claim that, under the assumptions of the theorem, F is consistent with two distinct SWOs, say R_1^* and R_2^2, satisfying C^*, WP^* and weak-welfarism. We observe first that we cannot have both uP_1^*v and vP_2^*u, by definition. Suppose we have both uP_1^*v and vI_2^*u. Then we can choose some w close enough to u and such that $w \ll u$ and vP_2^*w, by

Weak Pareto* and transitivity, whereas wP_1^*v, by Continuity*. Thus, we obtain the announced contradiction.

We show next that R^* also satisfies Inv*$(a + u_i)$ in three steps:

Step 1. It does satisfy a property weaker than Inv*$(a + u_i)$, viz. for any $u, v \in \Re^N$ and any $a \in \Re$, uP^*v implies $(a+u_1, \cdots, a+u_n)R^*(a+v_1, \cdots, a+v_n)$. This is almost immediate by weak recoverability and because $H(X, \mathcal{D}) = \Re^N$ since, whenever uP^*v for any $u, v \in \Re^N$, we may find V in $\mathcal{D}$ and $w, z \in X$ such that $u = V_w$, $v = V_z$ and $wP_V z$. Then, by $\text{Inv}(a + U_i) : \forall V' \in \mathcal{D}, \forall a \in \Re, R_{V'} = R_V$ if $\forall (x, i) \in X \times N, V'(x, i) = a + V(x, i)$, so that $(a + u_1, \cdots, a + u_n)R^*(a + v_1, \cdots, a + v_n)$ because F displays weak welfarism.

Step 2. We show that $uI^*v \Rightarrow (a + u_1, \cdots, a + u_n)I^*(a + v_1, \cdots, a + v_n)$ for every $a \in \Re$, and any $u, v \in \Re^N$. Let uI^*v and consider the dominating sequence $(u^k)_{k=1}^{\infty} \subset \Re^N$, $\forall k$, $u^k \gg u$ (resp. $u^k \gg v$), converging pointwise towards u (resp. v). By Weak Pareto* and transitivity together with Step 1, we obtain: $\forall k, u^k P^* u, u^k P^* v$ and $\forall a \in \Re$, $(a + u_1^k, \cdots, a + u_n^k)R^*(a + v_1, \cdots, a + v_n)$, (resp. $(a + u_1^k, \cdots, a + u_n^k)R^*(a + u_1, \cdots, a + u_n)$). By Continuity*, we obtain $(a+u_1, \cdots, a+u_n)R^*(a+v_1, \cdots, a+v_n)$, resp. $(a+v_1, \cdots, a+v_n)R^*(a + u_1, \cdots, a + u_n)$. The result follows.

Step 3. Since every invariance transformation is invertible, the result of step two can be reversed: $(a+u_1, \cdots, a+u_n)I^*(a+v_1, \cdots, a+v_n) \Rightarrow uI^*v$; this is inconsistent with uP^*v. Thus, $uP^*v \Rightarrow (a + u_1, \cdots, a + u_n)P^*(a + v_1, \cdots, a + v_n)$ and Inv*$(a + u_1)$ is established. The proof pertaining to Inv*(bu_i) is almost the same. ■

For instance, the maximin SWO, biz. $\forall\, u, v \in \Re^N$, $uR^*v \Leftrightarrow \min_i u_i \geq \min_i v_i$, is the unique SWO associated with the *maximin* SWFL satisfying *formal welfarism*, and it can fruitfully be associated with both the *maximin* SWFL and the *lexicon* SWFL under *weak welfarism*, wheres the *lexicon* SWO is the unique formally welfarist (with full recoverability) associate of the *lexicon* SWFL. The maximin SWO satisfies (Weak Pareto* and Continuity*, whereas the lexicon SWO satisfies Strict Pareto* but fails to be continuous.

Let us consider next Separability*. Even though the assumptions of the last theorem are satisfied, they do not suffice to obtain it, unless it is implied by another property, such as an invariance axiom, and it must be replaced by a weaker version which proves rather unwieldy.[14] r instance, the maximin SWO does not satisfy Separability* even though the lexicon SWO passes the test.

In conclusion, let us recapitulate some remarks about the properties of the Pareto axioms. Even though their direct influence is very limited since it is felt only in the absence of social conflict, the very incomplete social ranking they induce in isolation does display a great deal of informational parsimony. This has two aspects: once established, the Pareto ranking of any two alternatives is fully independent of the individual evaluations of the remaining alternatives, and moreover, the social ranking is simply based on the rankings underlying the individual evaluations: it is consistent with the strongest invariance principle, i.e. $\mathrm{Inv}(\varphi_i(U_i))$. As the analysis of (weak) welfarism shows, they have remarkably profound implications when they are associated with appropriate information filtering axioms.

3.5 Equity axioms

In a spirit of impartiality or equity, the Pareto-inspired principles may be directly extended to the cases where individual names have been permuted in one of the vectors U_x, U_y. This is known as the Suppes (1966) Grading principle[15] We proceed by stating its weak form. To simplify, we shall represent a permutation of the players by a $n \times n$ permutation matrix π where each element is either a 0 or a 1, and each line and each column contains a single 1.

Weak Suppes dominance (WS):

For all permutations π, $\forall\, U \in \mathcal{D}, \forall\, x, y \in X, \quad xP_U y$ if $U_x \gg \pi U_y$.

Much as the equity content of Weak Pareto may be reinforced to generate Weak Suppes domi-

[14]By an argument analogous to the one used in Step 1 of the proof of the last theorem, the following SWO property is implied if SE and weak welfarism hold: $\forall\, u, u', v' \in \Re^N, uP^*v \Rightarrow u'R^*v'$, if $\exists,\ M \subset N$ such that, $\forall\, i \in M, u_i v_i$ and $u'_i = v'_i$ whereas, $\forall\, j \in N \setminus M$, $u_j = u'_j$ and $v_j = v'_j$.

[15]See also Suppes (1957).

nance, we can reinforce Pareto indifference to obtain

Suppes indifference (SI):

For all permutations π, $\forall U \in \mathcal{D}, \forall x, y \in X, \quad xI_U y$ if $U_x = \pi U_y$.

The former axiom which implies Weak Pareto, induces a less incomplete social ordering than the latter since it can solve some social conflicts. Both *incomplete* social rankings satisfy Binary independence equally well, but Weak Suppes dominance proves much more demanding informationally. On the other hand, Suppes indifference which implies Pareto indifference, retains some bite within the welfarist structure; it is also much more demanding than the latter from the invariance viewpoint, at least when Weak Pareto is also imposed. These observations are made precise in the following

Theorem 3.15 *Suppose F has universal domain $\mathcal{D} = U$. Then, if it satisfies Weak Suppes dominance, it contradicts both Inv$(a_i + b_i U_i)$ and Inv$(\varphi_i(U_i))$. The same conclusion is valid if $|X| > 3$ and both Suppes Indifference and Weak Pareto hold.*

Proof: We prove first the statement pertaining to WS. Actually, we can rely on positive evaluation counts. Suppose $N = \{\ell, g, k\}$. Consider any $U \in \mathcal{U}$, any $x, y \in X$ such that $U(y,\ell) > U(x,g) > U(y,k) > U(x,k) > U(y,g) > U(x,\ell) > 0$. By WS, we get $yP_U x$. We proceed by constructing a profile V, which is informationally equivalent to U by Inv$(a_i + b_i U_i)$, and yet which is chosen so that $V(x,g) > V(y,\ell) > V(x,\ell) > V(y,k)U(y,k) > V(x,k) = U(x,k) > V(y,g) > 0$. First of all, we let $V(x,\ell) = U(x,\ell) + U(y,k), V(y,\ell) = U(y,\ell) + U(y,k)$ and $V(x,g) = bU(x,g)$, where $b > (U(y,\ell) + U(y,k))/U(x,g)$. By construction, $V(x,\ell) > U(y,k), V(y,\ell) > V(x,\ell)$ and $V(x,g) > V(y,\ell)$, and these three inequalities hold for any arbitrary small pair $U(y,v), U(x,\ell)$, with $U(y,g) > U(x,\ell)$, in particular if $U(y,g) < U(x,k)/b$. Indeed, if the latter inequality holds, $U(x,k) > V(y,g) = bU(y,g)$. By WS, we obtain $xP_V y$ despite $yP_U x$, contradicting both Inv$(a_i + b_i U_i)$ and Inv$(\varphi_i(U_i))$. If there are other individuals in N, we simply clone individual k. If N is a pair, we eliminate k.

To prove the statement involving both SI and WP, we go back to the case pertaining to three individuals, with the pair of profiles U and V constructed as above. Moreover, by UD, we can choose $z \in X$ so that $U(y,\ell) > U(z,\ell) = U(x,g) > U(y,k) > U(z,k) = U(x,k) > U(y,g) > U(z,g) = U(x,\ell)$. By SI, we have zI_Ux, whereas we have yP_Uz by WP, so that yP_Ux. We can pick V and a fourth element in X, apply UD, SI and WP again to conclude xP_Vy. Eventually, we invoke $\mathrm{Inv}(a_i + b_iU_i)$ and we have the same contradiction again. ■

Other axioms are inter profile statements, and they have no meaning out of the SWFL context since they relate the respective images of two distinct profiles; they tell us whether the social ranking ought to change and, if yes, in what direction whenever some profile undergoes a specific change. This may be exemplified with help of the Anonymity axiom: if two profiles are the same once the individual evaluation functions have been permuted in one of them, impartiality would recommend that they both be assigned an identical social ranking by the SWFL.

Anonymity (A):

For all permutations π, $\forall\, U, V \in \mathcal{D}$, $\;R_V = R_U$ if $U_x = \pi V_x$, $\forall\, x \in X$.

Interested readers will easily check our earlier examples: they will find out that most of them satisfy Anonymity. In association with any Pareto axiom, Anonymity may thus be consistent with the strongest invariance axiom. This is in contrast with its welfarist version; indeed, both Anonymity and Suppes indifference have the same welfarist translation (among others, a proof can be found in d'Aspremont and Gevers (1977)):

Anonymity (A*):*

For all permutations π, $\forall\, u, v \in \Re^N$, $\;uI^*v$ if $v = \pi\, u$.

Requiring an SWO to satisfy both Anonymity* and $\mathrm{Inv}^*(a_i + b_iu_i)$ leads to the most unpalatable consequences. This is a direct implication of a theorem by Krause (1995) summarized in the next section. We provide here a simple direct proof:

Theorem 3.16 *Suppose R^*, an SWO defined over $\Re^N$, satisfies both Anonymity* and Inv*$(a_i + b_iu_i)$ (resp. Inv*$(\varphi_i(u_i))$). Then all individual evaluation vectors must be socially indifferent.*

Proof: Let us pick any $u, v \in \Re^N$. We also avail ourselves of $w, s, r, a, b \in \Re^N$, such that $w_1 < \min\{u_1, u_2, \cdots, u_n, v_1, v_2, \cdots, v_n\} \leq \max\{u_1, u_2, \cdots, u_n, v_1, v_2, \cdots, v_n\} < w_2 < w_3 < \cdots < w_n$

and $s = (w_n, w_1, w_2, \cdots, n_{n-1})$, whereas $\forall\, i \in N, r_i = a_i + b_i w_i$ and $u_i = a_i + b_i s_i$. Furthermore, a is such that $r = w$; i.e. $a = w - bw$. By substitution, we obtain $u = w + b(s - w)$, so that, by definition of s, we can write: $u_1 = w_1 + b_1(w_n - w_1), u_2 = w_2 + b_2(w_1 - w_2), u_3 = w_3 + b_3(w_2 - w_3)$, etc. By A^*, wI^*s. By Inv*$(a_i + b_i u_i)$, uI^*w, provided $\forall\, i \in N$, $b_i > 0$. The latter condition can always be met if $u_1 > w_1$, $u_2 < w_2, u_3 < w_3$, etc. Notice that we have defined w so that the same reasoning may be applied to prove that vI^*w. By transitivity, we conclude that uI^*v. ■

Another intraprofile property, which is implied by Suppes indifference, simply requires that there be a very minimal symmetry in the treatment of any two individuals[16]. It sets a limit on the influence any individual can exert on the social ranking when he/she has a single opponent. As such, it does not rely on any interpersonal comparison.

Minimal individual symmetry (MIS) :

For any two individuals $i, j \in N$,

$\exists\, U \in \mathcal{D}, \exists\, x, y \in X$ such that

$U(x,i) > U(y,i), U(x,j) < U(y,j), U(x,\ell) = U(y,\ell), \forall\, \ell \notin \{i,j\},$ and $xI_U y$.

This requirement is satisfied by both the Borda method and the method of majority voting. If formal welfarism holds, we obtain the following version, which is expectedly much weaker than Anonymity*:

Minimal individual symmetry (MIS*)* :

For any two individuals $i, j \in N$,

$\exists\, u, v \in \Re^N$ such that $u_i >_i, u_j < v_j, u_\ell = v_\ell, \forall\, \ell \notin \{i,j\},$ and uI^*v.

We turn next to three more demanding equity conditions. The SWO form of the well-known Pigou-Dalton principle comes first.

Pigou-Dalton principle* :

$\forall\, i, j \in N, \forall\, u, v \in \Re^N, \forall\, \varepsilon \in \Re_{++},$

uR^*u if $(1) v_j \geq v_i, v_j = u_j - \varepsilon, v_i = u_i + \varepsilon$ *and* $(2)\ \forall\, \ell \notin \{i,j\}, v_\ell = u_\ell$.

[16]This condition was introduced in d'Aspremont (1985), where it is called Weak Anonymity.

A transfer from a relatively better-off individual to a relatively worse-off, without reversing their ranking, is weakly improving socially. A strict version of this axiom can also be found in the literature. If the SWO R^* is represented by a social evaluation function W from $\Re^N$ to $\Re$ (i.e., $uR^*v \Leftrightarrow W(u) \geq W(v)$), then our version of the Pigou-Dalton principle is equivalent to *Schur-concavity*: viz. for any $n \times n$ doubly stochastic matrix σ, for any $u, v \in \Re^N$, vR^*u if $v = \sigma u$, whereas Convexity* of R^*, the last axiom of Subsection 3.3.3, is equivalent to quasi-concavity of its representation, and this implies Schur-convavity if R^* satisfies also both Weak Pareto* and Anonymity*. See Moulin (1988).

Our last two equity axioms must be compared to the Pigou-Dalton* principle, as their opening statement is identical. The first one is known as Hammond's equity axiom [see Hammond (1976a, 1979)]:

$$\begin{array}{l} \textit{Hammond's equity* principle}: \\ \forall\, i, j \in N, \forall\, u, v \in \Re^N, \forall\, \varepsilon, \eta \in \Re_{++}, \\ vR^*u \text{ if } (1) v_j \geq v_i, v_j = u_j - \eta, v_i = u_i + \varepsilon \; \textit{and} \; (2)\; \forall\, \ell \notin \{i, j\}, v_\ell = u_\ell. \end{array}$$

We remark that the triggering condition implies $u_j > v_j < v_i < u_i$, a slight change from the previous axiom; yet, it is of much wider scope in the new axiom, since there is no requirement such as $\eta/\varepsilon = 1$. *A priori*, the latter ratio is not bounded above (ε can be arbitrarily small), and this new feature is controversial, despite the fact that the axiom conclusion is phrased only as a weak social preference relation. Our last axiom is introduced by Blackorby et al. (2002) under the name

$$\begin{array}{l} \textit{Incremental equity*}: \\ \forall\, i, j \in N, \forall\, u, v \in \Re^N, \forall\, \varepsilon \in \Re_{++}, \\ vI^*u \text{ if } (1) v_j \; = u_j - \varepsilon, v_i = u_i + \varepsilon \; \textit{and} \; (2)\; \forall\, \ell \notin \{i, j\}, v_\ell = u_\ell. \end{array}$$

Here also, we observe a widening of the triggering condition, in comparison with the Pigou-Dalton* principle; indeed, we have lost *any*reference to the relative welfare levels of i and j, and the size of the welfare transfer is unrestricted. On the other hand, the conclusion vI^*u is more demanding than the version we ascribe to Pigou and Dalton. In view of these remarks, we doubt that the above axiom will appeal to many as an equity property.

4 Independence and invariance-based characterizations

In order to uncover some ethical consequences of the axioms we have already introduced, it seems best to find our which social rankings they lead to. Our main purpose, in this survey, is to reveal the implications of declaring irrelevant various kinds of interpersonal comparisons. Ideally, we should investigate every invariance axiom we have defined. Our intention is not to be exhaustive but simply to illustrate the existing results. On occasion, we shall introduce new ones. For convenience, we shall cut this section in three parts based on as many clusters of axioms. Moreover, our theorems are often phrased in SWO language. Except for some special cases which are mentioned explicitly, formal welfarism is assumed to hold and the SWO pertains to the full domain $\Re^N$. In the next section, we also display some characterisations dispensing with invariance and/or independence axioms, at the cost of introducing other kinds of restrictions.

4.1 Restricting interpersonal level comparability

We begin this subsection with four invariance axioms, which we list in order of relative logical weakness: Inv*$(\varphi_i(u_i))$, Inv*$(a_i + b_i u_i)$, Inv*$(a_i + b u_i)$ and Inv*$(a_i + u_i)$. Each of the first two excludes interpersonal comparisons and leads to Arrow's negative conclusion, viz. weak dictatorship. From either of the last two, allowing for some interpersonal comparisons, we will derive something more palatable, the family of weak weighted utilitarian rules (clearly compatible with both of them). All three rules are weak in the sense that they do not define completely an SWO. Moreover, as we allow some weights to be nil in the weak utilitarian rule, weak dictatorship is a special case (where all weights are nil except one). We will indicate how these variations of classical results can be easily completed, in particular how pure utilitarianism can be obtained. We shall also innovate (1) by introducing a slight weakening of Inv*$(a_i + u_i)$, (2) by providing a new characterization of the "rank-weighted" utilitarian SWOs, among which the generalised Gini SWOs make up the most attractive subfamily; we do this with help of a new invariance axiom, obtained by restricting Inv*($a_i b u_i$) to the subset of well-ordered evaluation vectors $\mathcal{G}_N$.

As it turns out, to derive weak weighted-utilitarianism, we adapt the proof of Blackwell and Girshick (1954), Theorem 4.3.1, based on a supporting hyperplane argument. Instead on Inv*($a_i + u_i$) (their condition L_3), we use the weakened invariance condition

*Weak Invariance with respect to individual changes of origin** $(WInv*(a_i + u_i))$:
$\forall\, a \in \Re^N, \forall\, u, v \in \Re^N,$
$uP^*v \Rightarrow (a_1 + u_1, \cdots, a_n + u_n)R^*(a_1 + v_1, \cdots, a_n + v_n)).$

Moreover, we replace their dominance condition L_2 (i.e., $u \geq v \Rightarrow uR^*v$) by Weak Pareto* (Section 3.4). The result is strengthened if we add Minimal Individual Symmetry* or Anonymity* (Section 3.5).

Theorem 4.1 (Weak weighted utilitarianism) *If an SWO R^* satisfies Weak Pareto* and $WInv^*(a_i + u_i)$ (resp. Inv*($a_i + u_i$) or Inv*($a_i + bu_i$)), then there exists $\lambda \in \Re^N_+ \setminus \{0\}$ such that $\forall\, u, v \in \Re^N$*

$$\sum_{i \in N} \lambda_i u_i > \sum_{i \in N} \lambda_i v_i \Rightarrow uP^*v.$$

Moreover, if we add Minimal individual symmetry (resp. Anonymity*), we must have every component of λ strictly positive (resp. strictly positive and equal).*

Proof: Let $\mathcal{P} \equiv \{p \in \Re^N : p \gg 0\}$, $\mathcal{S} \equiv \{s \in \Re^N : sR^*e\}$, with $e = (0, \cdots, 0)$, and $\mathcal{Q} \equiv \{q \in \Re^N : q = s + p, s \in \mathcal{S}, p \in \mathcal{P}\}$. By WP*, for any $s \in \mathcal{S}$ and $p \in \mathcal{P}$, $(s + p)P^*sR^*e$, so that $P \subset \mathcal{Q} \subset \mathcal{S}$. Since $\mathcal{Q} = \cup_{s \in S}(\mathcal{P} + s)$, it is open in $\Re^N$ with e as a boundary point. Thus, if we can show that $\mathcal{Q}$ is convex, there is [see e.g., Theorem 2.2.1 in Blackwell and Girshick (1954)]a supporting hyperplane to $\mathcal{Q}$ through e: $\exists\, \lambda \in \Re^N - \{e\}$ such that, for all $q \in \mathcal{Q}$, $\sum_{i=1}^n \lambda_i q_i > 0$. Moreover, $\lambda_i \geq 0$ for each i, since $\mathcal{P} \subset \mathcal{Q}$. Finally, for any $u, v \in \Re^N$, uR^*v implies (by WP*) that $(u + \theta p)P^*v$, for any $p \in \mathcal{P}$ and $\theta \in (0, 1)$, and (by Winv*($a_i + u_i$)) that $[(u - v - +\theta p]R^*e$ so that $[(u - v + \theta)p] \in \mathcal{Q}$. Therefore, uR^*v implies $\sum_{i=1}^n \lambda_i(u_i - v_i + p_i) > 0$, for any $p \in \mathcal{P}$, and hence $\sum_{i=1}^{n} \lambda_i(u_i - v_i) \geq 0$. This is the result. To end the proof, it just remains to show that $\mathcal{Q}$ is convex.

For $s, s' \in \mathcal{S}, p, p' \in \mathcal{P}$, we have $(s+p), (s'+p') \in \mathcal{Q}$, implying $(s+p)P^*e$ and, by WInv*$(a_i + u_i)$, $(s+p+s') \in \mathcal{S}$ (since $(s+p+s')R^*s'R^*e$) so that $(s+p+s'+p') \in \mathcal{Q}$. Therefore $\mathcal{Q}$ is closed under addition, and to show that it is convex it is enough to show that $\mu q \in \mathcal{Q}$ whenever $q \in \mathcal{Q}$ and $\mu > 0$.

We first show that, for any $s \in \mathcal{S}$, $p \in \mathcal{P}$, for all positive integers k, m, and any $\theta \in (0,1), [(k/m)(s+\theta p)] \in \mathcal{S}$. Suppose not. Then, for some positive integers k, m and some $theta' \in (0,1), eP^*[(k/m)(s+\theta p)]$ for all positive $\theta \leq \theta'$ (using WP*). By WInv*$(a_i + u_i)$, we get $[-(k/m)(s+\theta p)]R^*e$, for all positive $\theta \leq \theta'$ and then: $[-k(/m)(s+\theta' p)]P^*[(k/m)(s+\theta')]$, implying $eR^*[2(k/m)(s+\theta' p)]$ and $eP^*[2(K/m)(s+\theta'' p)]$ for $0 < \theta'' < \theta'$; also $[-(k/m(s+\theta'' p)]P^*[2(k/m)(s+\theta'' p)]$, implying $eR^*[3(k/m)(s+\theta'' p)]$ and $eP^*[3(K/m)(s+\theta''' p)]$ for $0 < \theta''' < \theta''$; and so on, until we get $eP^*[m(k/m(s+\theta p)]$, for some positive $\overline{\theta}$ (and all $\theta \in (0, \overline{\theta})$ by WP*). However, for any $\theta^0 \in (0, \overline{\theta}), (s+\theta^0 p)P^*e$ for $\theta^2 \in (\theta^1, \overline{\theta})$, again a contradiction. And so on, for any value of k. Therefore $[k/m)(s+\theta p)] \in \mathcal{S}$ for all positive integers k and m, and any $\theta \in (0,1)$.

Now, choose any $q = s + p$ (with $s \in \mathcal{S}, p \in \mathcal{P}$) and any $\mu > 0$. We can find κ a positive rational close to μ, and some $p' \in \mathcal{P}$, such that $\mu(s+p) = \mu(s+\theta p + (1-\theta)p) = \kappa(s+\theta p) + p'$, with $\kappa(s+\theta p) \in \mathcal{S}$, by the above. Hence, $\mu(s+p) \in \mathcal{Q}$, so that $\mathcal{Q}$ is convex. The last clause of the theorem proceeds from a reductio ad absurdum. ∎

This is a derivation of weak weighted utilitarianism, which does not characterize completely an SWO, because it says nothing about the social ranking of any two alternatives whose numerical utilitarian evaluation is the same. In order to obtain a complete characterization of weighted utilitarianism as an SWO, (1) we use Continuity* (see Subsection 3.3.3) to grot indifference hyperplanes, and (2) we replace Weak Pareto* by Strict Pareto* in order to ensure that all weights be positive; (3) alternatively, to obtain the same result, we maintain Weak Pareto* and we add Minimal individual symmetry* or we can rely on the latter two axioms and restrict ourselves to the more demanding invariance axiom, viz. Inv*$(a_i + bu_i)$, Weak Pareto* and Minimal individual symmetry* while we drop Continuity*.

Theorem 4.2 (Weighted utilitarianism)

1. *Suppose an SWO R^* satisfies Continuity*, Weak Pareto* and any element of the triple $\{WInv^*(a_i + u_i), Inv^*(a_i + u_i), Inv^*(a_i + bu_i)\}$. Then, R^* is a member of the weighted utilitarian family: there exists $\lambda \in \Re^N_+ \setminus \{0\}$, and $\forall\, u, v \in \Re^N$,*

$$uR^*v \Leftrightarrow \sum_{i \in N} \lambda_i u_i \geq \sum_{i \in N} \lambda_i v_i.$$

2. *Moreover, if either Strict Pareto* is substituted or Minimal individual symmetry* is added, we must have every component of λ strictly positive.*

3. *The latter result holds also if R^* satisfies only Weak Pareto*, Inv$*(a_i + bu_i)$ and Minimal individual symmetry*.*

Proof:

1. By the above theorem, we know that there exists $\lambda \in \Re^N_+ \setminus \{0\}$, and $\forall\, u, v \in Re^N$, $\sum_{i \in N} \lambda_i u_i > \sum_{i \in N} \lambda_i v_i$ implies uP^*v. We first prove that $\sum_{i \in N} \lambda_i u_i = \sum_{i \in N} \lambda_i v_i$ implies uI^*v, whenever C^* holds. Suppose, instead, $\sum_{i \in N} \lambda_i u_i = \sum_{i \in N} \lambda_i v_i$ and vP^*u for some $u, v \in \Re^N$. Since the set $\{v' \in \Re^N \mid v'P^*u\}$ is open by C^*, there is some $v' \in \Re^N$ in a neighbourhood of u such that $v'P^*u$ and $\sum_{i \in N} \lambda_i u_i > \sum_{i \in N} \lambda_i v'_i$, a contradiction.

2. $\lambda \gg 0$ results from SP* or MIS*, as in Theorem 4.1.

3. Using Inv*$(a_i + bu_i)$ and MIS* we may show now that, for any $u, v \in \Re^N$, $uI * v$ if and only if $\sum_{i \in N} \lambda_i u_i = \sum_{i \in N} \lambda_i v_i$. We know that, for all $u, v \in \Re^N$, uR^*v implies $\sum_{i \in N} \lambda_i u_i \geq \sum_{i \in N} \lambda_i v_i$, and, hence, uI^*v implies $\sum_{i \in N} \lambda_i u_i = \sum_{i \in N} \lambda_i v_i$. The converse of this last implication can be proved by recurrence. Assume that, for an integer m, $1 \leq m < n$, and for any $u, v \in \Re^N$ satisfying $u_i = v_i$ for all $i > m$, $\sum_{i \in N} \lambda_i u_i = \sum_{i \in N} \lambda_i v_i \Rightarrow uI^*v$ (of course this holds trivially for $m = 1$, assuming w.l.o.g. that $\lambda_1 > 0$). We need to show that

the same implication holds if we take any $u', v' \in \Re^N$, with (whenever $m+1 < n$)$u'_i = v'_i$ for all $i > m+1$. So, suppose $\sum_{i \in N} \lambda_i w'_i = 0$, for $w' \equiv u' - v' \neq e$. We want $w' I^* e$. Observe that, by Inv*($a_i + bu_i$), the set $\{w \in \Re^N \mid wI^*e\}$ is convex[17] Then using MIS* for appropriate pairs of individuals and taking convex combinations, we can construct $u'', v'' \in \Re^N$ such that $u'' \neq v''$ and $u'' I^* v''$, for, by Inv*($a_i + bu_i$), $w'' \equiv (u'e - v'')I^*e$, and hence $\sum_{i \in N} \lambda_i w''_i = 0$, and such that $\text{sgn}(w''_{m+1}) = \text{sgn}(w'_{m+1})$ and $w''_i = 0$ for all $i > m+1$ (whenever $m+1 < n$). If we let $b > 0$ be such that $bw''_{m+1} = w'_{m+1}$, then $\sum_{i \in N} \lambda_i (w'_i - bw''_i) = 0$, and $w'_i - bw''_i = 0$ for all $i > m$, implying $(w' - bw'')I^*e$, or, using Inv*($a_i + bu_i$) twice, $w' I^* bw'' I^* be = e$. Therefore, for any $u, v \in \Re^N$, uI^*v if and only if $\sum_{i \in N} \lambda_i u_i = \sum_{i \in N} \lambda_i v_i$. Again, $\lambda \gg 0$ results immediately from MIS*. ■

By taking advantage of weak welfarism, we can also derive from Theorem 4.1 the SWFL version[18] of Arrow's (1963) General Possibility Theorem.

Theorem 4.3 (Weak dictatorship) *If an SWFL F satisfies Domain universality, Binary independence, Weak Pareto and Inv($\varphi_i(U_i)$) (resp. Inv($a_i + b_i U_i$)), then it is weakly dictatorial: i.e., there exists $i \in N$, such that, $\forall\, x, y \in X, \forall\, U \in \mathcal{U}$,*

$$U(x,i) > U(y,i) \Rightarrow xP_U y.$$

Proof: By weak welfarism, there is an SWO R^* satisfying WP* and Inv*($\varphi_i(u_i)$) (resp. Inv*($a_i + b_i u_i$)), and hence WInv*($a_i + u_i$). Then, by Theorem 4.1, R^* is weak weighted utilitarianism for some non-negative vector of weights $\lambda \neq 0$.

Now, take $x, y \in X$, and $U \in \mathcal{U}$ such that $U(x,i) > U(y,i)$, for some i, $U(y,j) > U(x,j)$, for all $j \neq i$, and $\sum_{j \in N} \lambda_j (U(x,j) - U(y,j)) > 0$. Hence, letting $u \equiv U_x$ and $v \equiv U_y$, we get uP^*v

[17] By Inv*($a_i + bu_i$), for wI^*e and $w'I^*e$, and for $\theta \in [0,1]$, $\theta w I^* e$ and $(1-\theta)w' I^* e$, so that $(\theta w + (1-\theta)w')I^*(1-\theta)w' I^* e$.

[18] This is slightly less general than Arrow's original results in terms of social welfare functions, defined on profiles of individual *preference*orderings, since not all preference orderings are representable by utility functions. Arrow's Independence of irrelevant alternatives can be seen as the conjunction of BIN (see Subsection 3.3.1) and Inv($\varphi_i(U_i)$).

and, by Inv*($\varphi_i(u_i)$) (resp. Inv*($a_i + bu_i$)): $(u_1, \beta u - 2? \cdots, \beta u_n) P^* (v_1, \beta v_2, \cdots, \beta v_n)$, hence $\lambda_i(u_i - v_i) \geq \beta \sum_{j \neq i} \lambda_j(v_j - u_j) > 0$, for any $\beta > 0$. This is impossible, unless $\lambda_j = 0$, $\forall\, j \neq i$. Weak dictatorship follows. ■

As suggested by Luce and Raiffa (1957 p. 344), such a result (and the argument of Theorem 4.1), can be used repeatedly to get weak lexicographic dictatorship, i.e., there exist a subset of N and a permutation of its members such that the social ranking always mimics the strict preference of the first one, and in cases he or she is indifferent, moves on to mimic the strict preference of the following member, and so forth until the subset is exhausted. By dropping[19] Weak Pareto (while assuming Pareto indifference), Krause (1995) provides a neat generalization of Theorem 4.3: the set N is exogenously partitioned in three subsets, one of which consists of dummies who are denied any influence on the social ranking, so that society is indifferent if and only if all the remaining individuals are indifferent. The strict preference of the latter is taken into account in lexicographic fashion, the order of priority being given exogenously. Yet, the nature of the influence exerted by each member of the two remaining subsets depends on the one they belong to. Society mimics the strict preference of members of one of them, whereas it reverses the strict preference of members of the other one. Adding Anonymity as another requirement in the same context precipitates social triviality, as we already showed in Theorem 3.16.

In contrast, availing ourselves of Anonymity*, Weak Pareto* and Inv*($a_i + u_i$), we can characterize pure utilitarianism with help of a beautifully simple proof argument due to Milnor (1954)[20], to recall

Theorem 4.4 (Pure utilitarianism) *An SWO R^* is pure utilitarianism if and only if it satisfies Weak Pareto*, Anonymity* and Inv*($a_i + bu_i$) (resp. Inv*($a_i + u_i$)).*

Proof: Necessity is clear. To prove sufficiency, suppose first that two vectors u and v in $\Re^N$ add to the same amount. By A^*, we can permute their components to get two new vectors in

[19] A seminal paper along this line is Wilson (1972).

[20] As shown in d'Aspremont and Gevers (1977).

increasing order, but mutually ranked as u and v. Considering successively, in these new vectors, each pair of corresponding components and subtracting from each the minimal one, we get again two new vectors which by Inv*$(a_i + u_i)$ (or Inv*$(a_i + bu_i)$) are again mutually ranked as u and v. Repeating these two operations at most n times, we finally get two vectors with all components equal to zero but still mutually ranked as u and v. Therefore u and v should be indifferent. Second, to prove that $\sum_{i\in N} u_i > \sum_{i\in N} v_i \Rightarrow uP^*v$, we define $\forall\, i \in N, w_i = u_i - \delta$, where $\delta = (1/n)(\sum_{i\in N} u_i - \sum_{i\in N} v_i)$, and, to get uP^*v, we simply combine the preceding argument to obtain $wI * v$ with WP* to obtain uP^*w. ■

Denicolò (1999) points out that the above result and its proof remain valid if we assume R^* to be a reflexive binary relation defined on $\Re^N$ that need not be transitive. The underlying structure must be more general than the SWFL concept, since its image set does not consist only of orderings of X. What amounts to a slight strengthening of A^* within this more general context is however required, to wit: For every permutation π of N, $\forall\, t, u, v \in \Re^N$ such that $v = \pi u$, we must have $uR^*t \Leftrightarrow vR^*t$.

Alternative characterisations of pure utilitarianism can also be obtained straightforwardly from Theorem 4.2.

Corollary 4.5 *An SWO R^* is purely utilitarian if and only if it satisfies Weak Pareto*, Anonymity*, WInv*$(a_i + u_i)$ and Continuity*.*

Other implications of Anonymity can be derived with help of an argument inspired by the proof of Theorem 4.1, provided we restrict invariance to the well-ordered space $\mathcal{G}_N$ (defined in Section 2.1) in imitation of Weymark (1981)[21]. This is

$\mathcal{G}_N$*-Invariance* with respect to common rescaling* :

and ordered individual changes of origin $(\mathcal{G}_N - Inv * (a_i + bu_i))$:

$\forall\, b > 0, \forall\, uv, a \in \mathcal{G}_n,\ uR^*v \Leftrightarrow (a + bu)R^*(a + bv).$

[21] Weymark (1981) characterizes the generalized Gini absolute inequality indices. See also Bossert (1990). Notice that our restricted axiom is too weak to be helpful in Theorem 4.4: Milnor's argument would not go through.

We then get the SWO which is the formally welfarist associate of the rule defined in Subsection 2.2.5.

Theorem 4.6 (Weighted rank utilitarianism) *If an SWO R^* satisfies Continuity*, $\mathcal{G}_N - Inv^*(a_i + bu_i)$, Weak Pareto* and Anonymity*, then, there exists $\lambda \in \Re^N_+ \setminus \{0\}$ such that $\forall\, u, v \in \Re^N$,*

$$\sum_{k\in N} \lambda_k u_{i(k)} \geq \sum_{k\in N} \lambda_k v_{i(k)} \Leftrightarrow uR^* v.$$

Proof: Using the notation of Theorem 4.1, we define the sets $\mathcal{P} \equiv \{p \in \Re^N \mid p \gg 0, p_1 < p_2 < \cdots > p_n\}$, $\mathcal{S}(u) \equiv \{s \in \mathcal{G}_N \mid sR^* u\}$, and $\mathcal{Q}(u) \equiv \{q \in \mathcal{G}_N \mid q = s + p, s \in \mathcal{S}(u), p \in \mathcal{P}\}$, for any $u \in \mathcal{G}_N$. By WP*, $(s+p)P^* sR^* u$, so that $\mathcal{Q}(u) \subset \mathcal{S}(u)$. We remark that $\mathcal{S}(u)$ is convex by $\mathcal{G}_N$$-$Inv*$(a_i + bu_i)$; indeed, for sR^*u and $s'R^*u$, and for $\theta \in [0,1]$, $\theta s R^* \theta u$ and $(1-\theta)s' R^* (1-\theta)u$, so that $(\theta s + (1-\theta)s')R^*(\theta u + (1-\theta)s')$, and $(\theta u + (1-\theta)s')R^*(\theta u + (1-\theta)u)$. Therefore, $(\theta s + (1-\theta)s')R^* u$. Clearly u is a boundary point of $\mathcal{Q}(u)$, and, since $\mathcal{P}$ is open in $\Re^N$ and convex trivially, $\mathcal{Q}(u)$ is open in $\Re^N$ and convex in view of the last remark. Therefore, there is a supporting hyperplane to $\mathcal{Q}(u)$ through u: there exists $\lambda^u \in \Re^N_+ \setminus \{0\}$, and $\forall\, q \in \mathcal{Q}(u)$, $\sum_{i\in N} \lambda^u_i q_i > \sum_{i\in N} \lambda^u_i u_i$. Also, $\lambda^u_i \geq 0$, $\forall\, i \in N$, since otherwise taking $q \in \mathcal{Q}(u)$, with q_i large enough, would violate the inequality. Now, by contraposition, $\forall\, v \in \mathcal{G}_n$, if $\sum_{i\in N} \lambda^u_i v_i < \sum_{i\in N} \lambda^u_i u_i$, then $v \notin \mathcal{Q}(u)$. Moreover, we cannot have vR^*u. Otherwise, we could find $p \in \mathcal{P}$ such that $(v+p) \in \mathcal{Q}(u)$ and $\sum_{i\in N} \lambda^u_i (v_i + p_i) < \sum_{i\in N} \lambda^u_i u_i$, a contradiction.

Assuming w.l.o.g. that $\sum_{i\in N} \lambda^u_i = 1$, for any $u \in \mathcal{G}_N$, we may show that, for $u = (\overline{u}, \cdots, \overline{u})$ and any $v \in \mathcal{G}_N$,

$$\left\langle \sum_{i\in N} \lambda^u_i v_i = \sum_{i\in N} \lambda^u_i u_i = \overline{u} \right\rangle \Rightarrow vI^* u.$$

Indeed, by C^*, we cannot have vP^*u, since the for $p \gg 0$ small enough, we would have $(v-p)R^*u$ implying that $v \in \mathcal{Q}(u)$, in contradiction with $\sum_{i\in N} \lambda^u_i v_i = \sum_{i\in N} \lambda^u_i u_i$. Moreover, if uP^*v, then, for $p \gg 0$ small enough, we would have $(u-p)R^*v$ implying that $u \in \mathcal{Q}(v)$, and hence

$$\overline{u} = \sum_{i\in N} \lambda^v_i u_i > \sum_{i\in N} \lambda^v_i v_i = \sum_{i\in} \lambda^u_i v_i = \sum_{i\in N} \lambda^u_i u_i = \overline{u},$$

again a contradiction. Therefore: vI^*u.

Finally, to show that all $\lambda^u s$ are equal, it is enough to take $u = (\overline{u}, \cdots, \overline{u})$ and $v = (\overline{v}, \cdots, \overline{v})$ in $\mathcal{G}_N$ and show that their associated weights, λ^u and λ^v respectively are equal. Suppose $\lambda^u \neq \lambda^v$, and assume w.l.o.g. that both $\overline{u}$ and $\overline{v}$ are positive (otherwise, by Gs_N−Inv*$(a_i + bu_i)$, we would translate the two hyperplanes $H^u = \{w \in \mathcal{G}_n : \sum_{i \in N} \lambda_i^u w_i = \overline{u}\}$ and $H^v = \{w \in \mathcal{G}_N : \sum_{i \in N} \lambda_i^u w_i = \overline{w}\}$ by adding $a \gg 0$ to both $\overline{u}$ and $\overline{v}$, while preserving the social welfare ordering between all pairs of points in $H^u \cup H^v$). Define the hyperplane $H^{bv} = \{w \in \mathcal{G}_N : \sum_{i \in N} \lambda_i^u w_i = b\overline{v}\}$, with $b \gg 0$ in such that $\overline{u} = b\overline{v}$. By $\mathcal{G}_N-$ Inv*$(a_i + bu_i)$ again, for any $w, w' \in H^v$, $wR^*w' \Leftrightarrow bwR^*bw'$ and bw, $bw' \in H^{bv}$. Then, since $\lambda^u \neq \lambda^v$, there is some $w^0 \in H^u \setminus H^{bv}$ and some $w^1 \in H^{bv}$ such that, say, $\sum_{i \in N} \lambda_i^u w_i^0 < \sum_{i \in N} \lambda_i^v w_i^1$ and $w^1 \gg w^0$. By WP*, this is in contradiction with $w^0 I^* u I^* w^1$. ■

Substituting SP* to WP*, all weights become positive. As we shall see later, we can characterize the generalized Gini family of SWOs by adding the Pigou-Dalton* principle to our list of axioms.

4.2 Full comparability: from cardinal to ordinal measurability

We proceed by studying the implications of invariance axioms allowing at least to compare evaluation levels interpersonally, and not weaker than the one based on common positive affine transformation. In other words, we dal not only with the polar cases Inv$(\varphi(U_i))$ and Inv$(a+bU_i)$, but also with their intermediary links Inv$(\varphi(\Delta U))$ and Inv$(\{a_i + bU_i\} \& \varphi(U_i))$. We adopt again the language of social welfare orderings to simplify. Application to SWFLs proves straightforward, when formal welfarism holds. Under weak welfarism, care may be required at least when Separability holds. The extent to which Separability may be imposed may vary: this provides a criterion for dividing our presentation.

4.2.1 No separability

We start with Inv*$(a + bu_i)$ and report first about its implications when the SWO satisfies Weak Pareto* and Continuity*. In order to gain some intuition, we assume initially that N reduces to the pair $\{i, j\}$. The family of SWOs we are led to describe is known as the min-of-means family: it encompasses the Generalised Gini family of orderings. We first split $\Re^2$ by means of the main diagonal to obtain two rank-ordered subsets, viz. $\{u \in \Re^2 \mid u_i \geq u_j\}$ and $\{u \in \Re^2 \mid u_i \leq u_j\}$: for each of them, we are given a pair of non-negative individual weights and we apply weighted utilitarianism separately; social indifference half lines meet on the 45-degree line. This geometric sketch translates formally as follows:

Theorem 4.7 *If $N = \{i, j\}$ and R^* satisfies Inv*($a+bu_i$), Weak Pareto* and Continuity*, there exist two ordered pairs $(\lambda_i, \lambda_j), (\mu_i, \mu_j) \in \Re^2_+$ summing to one and such that, either $\forall\, u, v \in \Re^2$, $uR^*v \Leftrightarrow \min\{\lambda u, \mu u\} \geq \min\{\lambda v, \mu v\}$, or $\forall\, u, v \in \Re^2, uR^*v \Leftrightarrow \max\{\lambda u, \mu u\} \geq \max\{\lambda v, \mu v\}$.*

A proof may be found in Bossert and Weymark (2000), Theorem 5 or in Deschamps and Gevers (1977). This straightforward result proved difficult to generalise: a beautiful solution is provided by Gilboa and Schmeidler (1989) and adapted to our context by Ben-Porath et al. (1997). Notice that the expression $\min\{\lambda u, \mu u\}$ (resp. $\max\{\lambda u, \mu u\}$) used in Theorem 4.7 can equivalently be replaced by $\min_{\alpha \in [0,1]}\{\alpha\lambda u + (1-\alpha)\mu u\}$ (resp. $\max_{\alpha \in [0,1]}\{\alpha\lambda u + (1_\alpha)\mu u\}$), and this seemingly more complicated presentation involving a line segment is in effect the non degenerate two-dimensional version of the compact and convex set which is the key element of the SWO representation they provide. Their theorem characterizes the *min-of-means* family of SWOs. Each member of the family may be defined as follows:

> *Min-of-means SWO* :
> An SWO belongs to the min-of-means family if and only if there exists some compact and convex subset of weights Λ of the $(n-1)$-dimensional simplex,
> (i.e. $\Lambda \subseteq \{\lambda \in \Re^N_+ \mid \sum_i \lambda_i = 1\}$),
> such that $\forall\, u, v \in \Re^N$, $uR^*v \Leftrightarrow \min_{\lambda \in \Lambda}\{\lambda u\} \geq \min_{\lambda \in \Lambda}\{\lambda v\}$.

If the SWO is continuous and satisfies Convexity*, they establish that its relation with Λ is one-to-one. Continuity* is necessary, as our discussion of weak welfarism makes clear. Convexity* is also required, in view of Theorem 4.7, where the same set Λ is relied on in two contrasting representations, one function being concave and the other convex.

If the SWO satisfies Anonymity*, its representation is based on a correspondence defined on the set consisting of the rank-ordered subsets of the SWO domain and linking each of them to one or more weighting schemes, i.e. points of the $(n-1)$- dimensional simplex, located in the same rank-ordered subset of the latter, so that their union makes up a symmetric Λ. In case this statement requires clarification, we remind the reader that each dimension of the simplex is associated with a fixed individual, just as it holds true to the full SWO domain. Instances in three dimensions may be illuminating. Under Anonymity*, Λ is symmetric with respect to the three 45° lines. For example, the maximin SWO is paired with the full simplex of dimension two; it is a member of the generalized Gini family, which as a rule calls for an hexagonal figure within the interior of the simplex. If the SWO reflects weighted utilitarianism, it is paired with a mere single point in the simplex. All the above comments are intended to convey some intuition about the following[22]:

Theorem 4.8 (Min-of means SWO) *A SWO R^* defined on $\Re^N$ satisfies Inv*$(a+bu_i)$, Weak Pareto*, Continuity*, and Convexity* if and only if it is a member of min-of-means-family.*

The implications of this representation theorem[23] for the SWFL context are obtained at once, provided formal welfarism holds:

Theorem 4.9 (Min-of-means SWFL) *Suppose F has universal domain; then, the two following statements are equivalent: (1) F satisfies Binary independence, Weak Pareto Continuity, Inv$(a+bU_i)$, and Convexity, and (2) there exists a compact and convex set: $\Lambda subseteq \{\lambda \in \Re^N_+, \sum_i \lambda_i = 1\}$ such that $\forall\, U \in U, \forall\, x, y \in X$,*

$$\min_{\lambda\in\Lambda}\{\lambda U_x\} \geq \min_{\lambda\in\Lambda}\{\lambda U_y\} \Leftrightarrow xR_U y.$$

[22]In there 1997 account of this theorem (p. 199), the three authors add unnecessarily the following clause to Weak Pareto*: $\forall\, u, v \in \Re^N, u \geq v \Rightarrow uR^*v$; indeed, this is implies by Weak Pareto* and Continuity* taken together. On the other hand, our Convexity* axiom is stronger than the related axiom they rely on.

[23]Each *representation theorem* specifies as narrowly as possible the utility representations of each element of a set of orderings displaying common characteristics. In particular, representation theorems prove very useful in formally welfarist contexts.

In view of Roberts's theorem on weak welfarism (our Theorem 3.14), we can drop Continuity from the above list and exhibit another characterisation result based on Theorem 4.8, provided Convexity of F implies Convexity* of its associated SWO under weak welfarism. To check this, we consider any $u, v \in \Re^N$ such that uR^*v, and any $\alpha \in \Re_{++}$, $\alpha < 1$. Then, we can find $U \in \mathcal{D}$ and $x, y, z \in X$ such that $u = U_x$, $v = U_y$ and $\alpha u + (1_\alpha)v = U_z$, and we observe that $U_z R^* v$ by Convexity of F; indeed, $\forall U \in \mathcal{U}$, $\forall x, y \in X, xR_U y$ implies $U_x R^* U_y$; hence, we can apply Roberts' theorem:

Theorem 4.10 (Weak Min-of-means SWFL) *Suppose F satisfies Domain universality, Binary independence, Weak Pareto, Inv$(a + bU_i)$, and Convexity; then, there exists a compact and convex set $\Lambda \subseteq \{\lambda \in \Re^N_+ \mid \sum_i \lambda_i = 1\}$ such that $\forall U \in \mathcal{D}$, $\forall x, y \in X$,*

$$\min_{\lambda \in \Lambda}\{\lambda U_x\} > \min_{\lambda \in \Lambda}\{\lambda U_y\} \Rightarrow xP_U y.$$

However interesting these results may be, they uncover an embarrassment of riches and call for the addition of new axioms.

4.2.2 A modicum of Separability

Our next theorems introduce more or less demanding versions of Separability and/or more restrictive invariance axioms. One of them, proved by Gevers (1979) and further strengthened by Ebert (1987), is based on the mixed-invariance axiom Inv*$(\{a_i + bu_i\}\&\varphi(u_i))$, which is consistent with a very limited amount of separability.

Theorem 4.11 *Suppose an SWO R^* satisfies Inv*$(\{a_i + bu_i\}\&\varphi(u_i))$, Strict Pareto* and Anonymity*; then, there exists a subset $M \subseteq \{1, 2, \cdots, n\}$, such that $\forall u, v \in \Re^N$,*

$$\sum_{j \in M} u_{i(j)} > \sum_{j \in M} u_{i(j)} \Rightarrow uP * v.$$

We shall omit the lengthy proof. By adding $\mathcal{G}_N$-Minimal individual symmetry* to our list of axioms, we eliminate some uninteresting SWOs.

G-*Minimal individual symmetry**:

For any two individuals $i, j \in N$, $\exists\, u, v \in \mathcal{G}_N$ such that $u_i < v_i, u_j < v_j$,

$u_\ell = v_\ell$, $\forall\, \ell \notin \{i, j\}$, *and* uI^*v.

This new property *not* weaker than the original statement, even though it is obtained from it by substituting $\mathcal{G}_N$ for $\Re^N$. For instance, the lexicon SWO satisfies Minimal individual symmetry*, whereas it contradicts the restricted version.

Corollary 4.12 *If an SWO R^* satisfies Inv*($\{a_i + bu_i\}\&\varphi(u_i)$), Strict Pareto*, Anonymity* and $\mathcal{G}_N$-minimal individual symmetry*, it is weakly pure utilitarian.*

Alternativeley, we can add Continuity*, which turns out to have stronger marginal implications, to wit:

Corollary 4.13 *An SWO R^* satisfies Inv*($\{a_i+bu_i\}\&\varphi(u_i)$), Strict Pareto*, Anonymity* and Continuity* if and only if it is purely utilitarian.*

It is also instructive to go back to the weaker and more natural invariance axiom Inv*($a+bu_i$), and to study the family of anonymous SWOs satisfying $\mathcal{G}_N$-Separability*, a property defined in Subsection 3.3.3. Indeed, by restricting the unconditional property to well-ordered evaluation vectors under these circumstances, we obtain a slightly modified version of an important result due to Ebert (1988b):

Theorem 4.14 *If an SWO R^* satisfies Inv*($a+bu_i$), Strict Pareto*, Anonymity*, Continuity* and $\mathcal{G}_N$-Separability*, then, there exists $(\lambda_j)_{j=1}^n \in \Re^N_{++}$ such that, $\forall\, u, v \in \Re^N$,*

$$uR^*v \Leftrightarrow \sum_{j=1}^n \lambda u_{i(j)} \geq \sum_{j=1}^n \lambda_j v_{i(j)}.$$

Remark 1. As he applies Debreu's (1960) classical theorem on the representation of orderings involving at least three independent factors, Ebert requires $n \geq 3$. This is unnecessary in view of Theorem 4.7.

Remark 2. This family of SWOs intersects the generalized Gini family described in example 2.2.5.

4.2.3 Full Separability

We proceed with a result due to Deschamps and Gevers (1977,1978), which is based on the full strength one can impart to both Separability and Pareto dominance. It relies neither on Continuity, nor on Roberts' theorem, as the latter seems little helpful under Separability; yet the theorem provides a joint derivation of weak weighted utilitarian rules, together with lexicon, and its inequitable mirror image, leximax. This result (the proof of which is too long even to be sketched here), leads to corollaries characterizing very antagonistic social welfare orderings viz. weighted utilitarianism and pure utilitarianism on one hand, lexicon on the other.

Theorem 4.15 *For $n > 2$, an SWO R^* satisfying Strict Pareto*, Minimal Individual Symmetry*, Inv*$(a + bu_i)$, (resp. Inv*$(\{a_i + bu_i\}\&\varphi(u_i))$ and Separability* is either leximin, leximax or weak weighted utilitarianism (with all weights positive).*

Proof: This follows directly from the theorem of Deschamps and Gevers (1977), which is proved for Inv*$(a + bu_i)$. We can also observe that the original result goes through if we rely on the stronger mixed-invariance axiom Inv*$(\{a_i + bu_i\}\&\varphi(u_i))$ since none of the conclusions are contradicted. The really new feature is the use of MIS*. Indeed, as stated without MIS*, the original version of the theorem only requires a non-empty set of at least three undominated individuals. An individual i is undominated if and only if for all $j \neq i$, there is some $u, v \in \Re^N$ such that $u_h = v_h$ for all $h, i \neq h \neq j, u_j > v_j$ and vR^*u (with SP* one should have $v_i > u_i$). This clause is required because the Separability axiom loses its bite and reduces to Strict Pareto* if $n = 2$, so that we are back to Theorem 4.2. Clearly, MIS* implies that all individuals are undominated. ■

A first characterizationn that one can deduce from Theorem 4.15 is that of lexicon. To get rid of utilitarianism one can further strengthen the mixed-invariance requirement by considering cardinal interpersonal comparisons as irrelevant, and rely on Inv*$(\varphi(u_i))$, but one can also use another axiom lying between the latter and Inv*$(a + bu_i)$, viz. Inv*$(\varphi(\Delta u))$. To eliminate leximax, any innocuous equity condition can be introduced, such as

Minimal equity (ME*)* :

For some $i, j \in N$, there exist $u, v \in \Re^N$ such that

$u_h = v_h, \forall\, h, i \neq h \neq j, v_j > u_j > u_i > v_i$ and uR^*v.

We then get

Theorem 4.16 *For $n > 2$, an SWO R^* satisfying Strict Pareto*, Minimal Individual Symmetry*, Minimal Equity*, Inv*($\varphi(\Delta u)$) (resp. Inv*($\varphi(u_i)$)) and Separability* is leximin.*

Proof: We simply prove the part relying on Inv*($\varphi(\Delta u)$), by going back to the previous theorem, and by showing that the stronger invariance axiom contradicts utilitarianism. The proof is based on two steps.

Step one: Suppose first that at least two individuals i, j have unequal utilitarian weights denoted α and β, respectively, and such that $\alpha/\beta > 1$. We consider three evaluation vectors u^0, u^1, v defined as follows: $\forall\, k \in N, i \neq k \neq j$, $u^0_k = u^1_k = 0 = v_i = v_j = v_k$; $u^0_1 = -\varepsilon^0 < 0, u^0_j = \eta^0 > 0$, $u^1_i = -\varepsilon^1 < 0, u^1_j = \eta^1 > 0$, where we can always choose the last four numbers so that $\eta^1/\varepsilon^1 > \alpha/\beta > \eta^0/\varepsilon^0 > 1$. By construction, we have $\alpha\varepsilon^0 > \beta\eta^0$ and $\alpha\varepsilon^1 < \beta\eta^1$, so that utilitarianism implies $u^1 P^* v P^* u^0$, which contradicts Inv*($\varphi(\Delta u)$).

Step two: Suppose nedt that all weights are equal. Let $n = 3$ and consider the three following evaluation vectors: $u^0 = (1, 2, 8.5)$, $u^1 = (1, 2, 85 + \varepsilon)$, where $\varepsilon > 1$, and $v = (4, 4, 4)$; we notice again that utilitarianism implies $u^1 P^* v P^* u^0$, which contradicts Inv*($\varphi(\Delta u)$). If there are more individuals, we simply add indifferent people with constant evaluation count 4. ■

Our proof is obtained by adapting an argument used by Bossert (1991) in the context of a weaker theorem.

Let us turn next to characterisations of utilitarianism. Weak weighted utilitarianism immediately follows from adding $\mathcal{G}_N$-Minimal individual symmetry* to the set of axioms used in Theorem 4.15. Continuity also contradicts both lexicon and leximax [see, e.g., Maskin (1978), Moulin (1988)], so that the following consequence can be easily establishes (see the argument in Theorem 4.2).

Corollary 4.14 *For $n > 2$, the SWO R^* satisfies Strict Pareto*, Minimal individual symmetry*, Inv*$(a + bu_i)$ (resp. Inv*$(\{a_i + bu_i\}\&\varphi(u_i))$), Separabilitiy* and Continuity* if and only if it is weighted utilitarianism (with all weights positive).*

Another way to eliminate both lexicon and leximax in Theorem 4.15 is to strengthen the invariance axiom Inv*$(a+bu_i)$ to Inv*$(a+bu_i)$, which implies SE*. Only weak weighted utilitarianism (with positive weights) remains. This result can be obtained using still other arguments [Maskin (1978), d'Aspremont (1985)].

4.3 Homothetic vs. translatable social welfare functionals

The last group of results will be concerned with ratio-scale invariance. As we have seen for SWFLs, statements that are compatible with such invariance axioms are meaningful if the domain is restricted to be $\mathcal{D}^*$, a domain which is homogeneous in sign. To illustrate such axioms in the welfarist framework, we shall restrict ourselves to the domain of positive utilities $\Re^N_{++}$. This will have the additional advantage of simplifying the proofs. For the first result we may even use Blackwell and Girshick (1954) again.

Theorem 4.17 (Weighted utilitarianism vs. Nash) *Suppose an SWO R^* satisfies Continuity* and Strict Pareto*. If Inv*$(a_i + u_i)$ holds, then it is weighted utilitarianism; if Inv*$(b_i u_i)$ holds and $\mathcal{H}(X, \mathcal{D}) = \Re^N_{++}$, then we get the Nash bargaining solution with status-quo point normalised to zero, i.e. there exists $(\lambda_j)_{j=1}^n \in R^N_{++}$, such that $\forall\, u, v \in \Re^N_{++}$,*

$$uR^*v \Leftrightarrow \prod_{i=1}^n u_i^{\lambda_i} \geq \prod_{i=1}^n v_i^{\lambda_i}.$$

With Anonymity in addition, we get respectively pure utilitarianism and the symmetric Nash solution.*

Proof: From Theorem 4.2 we know that, with Inv*$(a_i + u_i)$, C* and SP*, we get a weighted utilitarian SWO R^*. Using this fact, another SWO $\hat{R}^*$ can now be defined on $\Re^N_{++}$ by putting, $\forall\, u, v \in \Re^N$,

$$(e^{u_1}, \cdots, e^{u_n})\hat{R}^*(e^{v_1}, \cdots, e^{v_n}) \Leftrightarrow uR^*v.$$

Clearly $\hat{R}^*$ also satisfies C* and SP*, and Inv(a_i+u_i) for R^* translates into Inv*$(b_i u_i)$ for $\hat{R}^*$ (taking $b_i = e^{a_i}$). Finally, for any $u, v \in \Re^N_{++}$, $u\hat{R}^*v$, being equivalent to $(\ln u_1, \cdots, \ln u_n)R^*(\ln v_1, \cdots, \ln v_n)$, holds if and only if $\sum_{i\in N} \lambda_i \ln u_i \geq \sum_{i\in N} \lambda_i \ln v_i$ or, equivalently, $\prod_{i=1}^n u_i^{\lambda_i} \geq \prod_{i=1}^n v_i^{\lambda_i}$. With A^*, all the weights should be equal. ■

If we allow for interpersonal comparability, we enlarge considerably the admissible class of social welfare orderings. For example, with C*, SP*, A* and Inv*(bu_i) (resp. Inv*$(a + u_i)$), the class of admissible SWOs are all this representable by a social evaluation function defined on $\Re^N_{++}$ which is continuous, strictly increasing in each argument, symmetric and nomothetic (resp. translatable). However, the addition of separability makes it possible to get functional-firm characterizations, as is well known from the work of Blackorby and Donaldson (1980, 1982). An interesting intermediate case is given by imposing the weaker separability condition $\mathcal{G}_N$- Separability* defined in Subsection 3.3.3. In fact, having a large class of SWOs gives the possibility of introducing additional equity conditions, more demanding than simple anonymity. In the final result that we state here to illustrate this possibility, combining results in Ebert (1988a,b), we use the Pigou-Dalton* principle defined in Section 3.5:

Theorem 4.18 *Suppose $n > 2$, and the SWO R^* satisfies Continuity*, $\mathcal{G}_N$-Separability*, Anonymity*, Strict Pareto*, and the Pigou-Dalton* principle.*

1. *If Inv*$(a+u_i)$ *holds, then, for some $\beta \in (0,1)$ and some $(\lambda_j)_{j=1}^n \in \Re^N_{++}$ with $\lambda_j \geq \lambda j + 1$,*

$$\forall\, u, v \in \Re^N_{++}, uR^*v \Leftrightarrow \frac{1}{\ln \beta} \ln \sum_{j=1}^n \lambda_j e^{(\ln \beta)u_{i(j)}} \leq \frac{1}{\ln \beta} \ln \sum_{j=1}^n \lambda_j e^{(\ln \beta)v_{i(j)}},$$

$$or, \forall\, u, v \in \Re^N_{++}, uR^*v \Leftrightarrow \sum_{j=1}^n \lambda_j u_{i(j)} \geq \sum_{j=1}^n \lambda_j v_{i(j)}.$$

2. *If Inv*(bu_i) *holds, then, for some $\beta \in (0,1)$ and some $(\lambda_j)_{j=1}^n \in \Re^N_{++}$ with $\lambda_j \geq \lambda_{j+1}$,*

$$\forall\, u, v \in \Re^N_{++}, uR^*v \Leftrightarrow \left(\sum_{j=1}^n \lambda_j u_{i(j)}^\beta\right)^{1/\beta} \geq \left(\sum_{j=1}^n \lambda_j v_{i(j)}^\beta\right)^{1/\beta},$$

$$or, \forall\, u, v \in \Re^N_{++}, uR^*v \Leftrightarrow \prod_{j=1}^n u_{i(j)}^{\lambda_j} \geq \prod_{j=1}^n v_{i(j)}^{\lambda_j}$$

3. *If Inv*$(a + bu_i)$ holds, then we get (only) the generalized Gini social evaluation function: for some $(\lambda_j)_{j=1}^n \in \Re_{++}^N$ with $\lambda_j \geq \lambda_{j+1}$,*

$$\forall\, u, v \in \Re_{++}^N, uR^* v \Leftrightarrow \sum_{j=1}^n \lambda_j u_{i(j)} \geq \sum_{j=1}^n \lambda_j v_{i(j)}.$$

This is one characterization of the generalized Gini social evaluation function. Others follow readily from Theorems 4.6 and 4.14, provided the Pigou-Dalton* principle is added.

5 Discarding neutrality or invariance

Some interesting SWFL characterizations do not rely on Invariance, and some others do away with Strong neutrality. We have assembled below some of those we are aware of. Harsanyi (1955, 1977) and Hammond (1976a) pioneered this approach. Their results, which are not based on invariance properties, deal respectively with utilitarianism and the lexicon principle. For easy comparison, we present them in the SWFL framework, even though neither contribution was using it originally. As we shall indicate briefly, some very recent work has developed in either case the consequences of leaving aside Strong neutrality, while reintroducing at least implicitly the invariance properties which are familiar in noncooperative game theory, viz. Inv$(\varphi_i(U_i))$ or Inv$(a_i + b_i U_i)$. Our presentation falls in two natural subsections: in the first, we deal with utilitarianism as a tool for social evaluation under risk and uncertainty and we comment extensively on the latter context; in Subsection 5.2 we deal with the lexicon principle, in both abstract and more structured economic environments. In the latter case, the limitations of formal welfarism are put in evidence.

5.1 Uncertainty and risk: from Harsanyi to Relative Utilitarianism

As we stand on the threshold of this subsection, we mention the invariance-free approach to utilitarianism which is developed by Blackorby et al. (2002). In contrast with Harsanyi, it does not rely on risk and uncertainty, but it makes essential use of the Incremental equity axiom (defined at the end of Section 3.5), the cogency of which is not self-evident.

Theorem 5.19 (Pure utilitarianism) *An SWO satisfies both Strict Pareto* and the Incremental equity* axiom if and only if it is purely utilitarian.*

Requiring the SWO to be purely utilitarian in each case where there are only two non indifferent individuals is likely to appear objectionable if this is proposed as an independent axiom. Yet, as d'Aspremont (1985, lemma 3.3.1) shows, it may replace both Incremental equity* and Strict Pareto* to characterize pure utilitarianism. Even though this result cannot be more convincing than the axiom it is based on, it may prove useful as an intermediate lemma if it can be derived from more palatable axioms.

As we wish to relate some of Harsanyi's classical contributions and SWFL theory, we have to introduce uncertainty and risk. Indeed, explicit consideration of these features raises specific difficulties when evaluating social decisions. We shall be inexorably brief on technical matters and concentrate on interpretation; we shall emphasize, in particular, an example originally proposed by Diamond (1967) and further developed by Ben-Porath et al. (1997).

In order to add the dimensionn of state contingency to our aggregation problem, we shall define a set S of states of the world, the realization of any of which is regarded as a possibility, whether the evaluation of social decisions is to be made by an individual or by the social observer of by both. The states are assumed mutually exclusive and such that only one of them will eventually obtain.

The evaluate takes place or the social decision is selected *before* the true state is observed, and it can take into account the relative likelihood of the states. Furthermore, people's beliefs about the latter may legitimately diverge in the context of uncertainty; in this case, it is assumed that an n-tuple of not necessarily distinct probability distributions $(p_s^i)_{s\in S}^{i\in N}$ can be elicited from the individuals, while the ethical observer's beliefs may be captured by yet another probability distribution $(p_s)_{s\in S}$. In the context of risk, we assume that everyone entertains the same beliefs, which can be captured by the same probability distribution[24].

Under uncertainty or risk, the consequence of any social decision may vary according to the state. As we define it, a consequence involves a full description of all aspects which are

[24] For more on this issue, see Mongin and d'Aspremont (1998).

expected to matter in the eyes of any individual or in those of the ethical observer, except for the evaluation process itself. In the sequel, we shall denote by $\mathcal{C}$ the set of all relevant consequences, whereas c_s^x will designate the particular consequence contingent on the realization of s if a particular social decision $x \in X$ is chosen.

In this setup, the ethical observer, required to establish a social ranking over x, is supposed to be informed of as many individual evaluation profiles as there are states in S. The generic element of the s-conditional profile is given by $V(c_s^x, i)$, where $x \in X$ and $i \in N$. Having s as a direct argument would prove redundant if we define consequences in a comprehensive manner as we do. The observer's goal is to aggregate with respect to both individuals and states. Before we embark on this double exercise, let us point out that we shall single out expected utility maximization as the main individual decision criterion under uncertainty and risk[25]. Even though it may lack empirical support, it is considered by many authors, including Harsanyi, as the hallmark of rationality. Its mathematical closeness to utilitarianism is interpreted by Mirrlees (1982) in a natural way: under uncertainty and risk, the individual decision-maker imagines by anticipation the experience of as many possible selves as there are states. It is therefore tempting to combine both aggregation principles and to assume that inequality aversion and risk aversion are two faces of the same attitude, as Harsanyi suggests.

Before we embark on the technical presentation, we would like to convey some intuition about the problem at stake and discuss Diamond's (1967) example, as developed by Ben-Porath et al. (1997). It combines a simple pure distribution problem with the possibility to toss a fair coin, the sides of which are respectively denoted h and t. To simplify, we omit any reference to C. Society consists of two perfect twins ℓ and m, with the same risk aversion and the same level of initial income which we use as origin for individual utilities. The ethical observer is to rank three social decisions y, z and w. Decision y consists of allocating an extra unit of utility to ℓ, without transferring anything to m, independently of the toss outcome. Decision z involves such

[25] It has invaluable advantages when decisions must be taken sequentially while information evolves between decisions, a situation best described by a decision tree. See for instance Hammond (1988,1998) and Sarin and Wakker (1994).

a dependence since it consists of a fair lottery: ℓ gets the extra unit of utility and m receives nothing if h shows up, whereas the reverse takes place if t shows up. Decision w is such that one unit of utility is allocated to both ℓ and m if t obtains, whereas both receive nothing otherwise. This description is summarised in the following table, where $x \mid s$ means "decision x, given state s", and the numbers in the various cells indicate the relevant values of $V(c_s^x, i)$:

	ℓ	m		ℓ	m
$y \mid h$	1	0	$y \mid t$	1	0
$z \mid h$	1	0	$z \mid t$	0	1
$w \mid h$	1	1	$w \mid t$	0	0

We shall define yet another reference decision when we need it. How are we to rank socially y, z and w? It may depend on the order of aggregation. We distinguish two methods:

1. In the *ex post* approach, aggregation over individuals is done first, conditionally on each state in turn. This is a problem SWFL theory is meant to help solving. If the $|S|$ conditional social rankings can be represented by as many social evaluation functions, the intermediate output is a $|S|$-tuple of evaluation indicators of each decision through its consequences associated with the states. When both $|X|$ and $|S|$ are finite the typical element of the corresponding $|X| \times |S|$ matrix may be denoted $W(x, s)$. For instance, in case the maximin principle is applied at the first round of aggregation, we let $W(x, s) = \min_{i \in N}\{V(c_s^x, i)\}$, whereas we define $W(x, s) = \sum_{i \in N} V(c_s^x, i)$ if the aggregation principle is purely utilitarian. The ethical observer's own probability distribution is relevant in the next step of this approach, when the social ranking of decisions is obtained by aggregating over the set of states.

   In our example, decision z, given h, is the same as $y \mid h$, with respect to both income and individual utilities, so that we declare them socially indifferent, given h. Assuming that the relevant social orderings are anonymous, we observe that $z \mid t$ must be also socially indifferent with $y \mid t$. As we proceed to the next aggregation stage, we may conclude that y and z are socially indifferent for lack of a reason pointing one way or another.

When comparing $z \mid h$ with $w \mid h$, the latter dominates, whereas the opposite is true if t obtains. Suppose we sum utility numbers at the first level of aggregation to evaluate social decisions conditionally on states and we rely on expected social utilities at the second stage; then z and w must be socially indifferent. This seems to conform to Hammond's (1982) version of utilitarianism which relies on the ex post approach; individual utilities are initially summed, and their degree of concavity is chosen to reflect the ethical observer's aversion for income inequality, independently of the individuals' attitudes towards risk. In this particular instance, however, many people prefer w to z because equality is achieved in the former case both ex ante and ex post; this ranking is imp lied if the slightest degree of social aversion towards inequality in *utilities* is applied at the first stage of aggregation, as is consistent with the generalized Gini SWFLs among others, even tough expected social welfare is maximized at the ensuing stage.

2. In the *ex ante* approach, aggregation over states is done first for each $i \in N$ in turn. It relies on the individual's own method and beliefs, as the latter are captured by $(p^i_s)_{s \in S}$, at least when the individual's learning process is deemed reasonable. If the latter condition fails, it is natural to rely on the ethical observer's own method and/or beliefs. The intermediate output is a list of n individual evaluation functions defined over the set of social decisions. The typical element of the corresponding $|X| \times n$ matrix may be denoted $U(x, i)$. For instance, in case an individual's risk aversion is extreme, as perhaps Rawls would think appropriate in the original position, the maximin principle is applied at the first level of aggregation, and we let $U(x, i) = \min_{s \in S}\{V(c^x_s, i)\}$, whereas we define $U(x, i) = \sum_{s \in S} p^i_s V(c^x_s, i)$ if aggregation over states relies on the expected utility principle, as Harsanyi recommends. This approach is relevant in contexts where the SWFL is utility-based and the ethical observer disregards what is to take place after the state of the world is disclosed, in particular, as Kolm (1998) explains, when individuals are held responsible for the risk they incur. The intermediate output is used as an input for the ensuing interpersonal aggregation stage, at which SWFL theory can be invoked. Let us turn next to our example and rearrange our date in more appropriate fashion; here, $x \mid i$

means that the corresponding row displays the relevant values of $V(c_s^x, i)$ for each state in turn.

	h	t		h	t
$y \mid \ell$	1	1	$y \mid m$	0	0
$z \mid \ell$	1	0	$z \mid m$	0	1
$w \mid \ell$	1	0	$w \mid m$	1	0

At the first round of aggregation, we must have $U(z,\ell) = U(w,\ell)$ as well as $U(z,m) = U(w,m)$, because the coin is fair. As both individuals are indifferent between w and z, society may also be declared indifferent between w and z at the next aggregation stage. On the other hand, it is natural to assume that $U(y,\ell) = 1$ and $U(y,m) = 0$. If we denote by η the common value of $U(z,\ell) = U(z,m)$, we must have $0 \leq \eta \leq 1$. For instance, if individuals are expected utility maximizers, $\eta = \frac{1}{2}$, whereas $\eta = 0$ if they endorse the maximin principle. To make more transparent the choice between y and z as we move to interpersonal aggregation, we may assume that there exists some $v \in X$ of which *both* ℓ and m derive a sure utility level ε and which is socially indifferent to y, so that comparing socially y with z amounts to Pareto-comparing ε with η, i.e. social acceptance of inequality in sure utilities with private acceptance of utility risk. Harsanyi argues that the same expected utility ought to be applied both in ordinary life and in the original position, so that $\varepsilon = \eta = \frac{1}{2}$. This makes for simpler decision criteria since there is no need to distinguish the ex ante approach from the ex post method. Yet, as Diamond pointed out, many people's ethical intuitions indicate a strict preference for z over y because their aversion to utility inequality among individuals is greater than the individuals' aversion to utility risk. The same conclusion is reached by an interesting alternative theory developed by Epstein and Segal (1992), which is also based on the ex ante approach; it assumes expected utility at the private level and implies quadratic aggregation of individual expected utilities, so that $\varepsilon < \eta = \frac{1}{2}$. The generalized Gini SWFLs could be used to the same effect instead of the quadratic, as Ben-Porath et al. (1977) suggest.

Unless one is persuaded by Harsanyi's arguments, the ex ante approach ma thus clash with the ex post method. One way of solving the clash would be to add a time dimension, allowing for the fact that each individual can only entertain expectations at the first period, whereas he or she experiences the true state of the world at the following period. This would generally warrant yet another round of aggregation, even though social evaluation takes place before the true state of the world is known. As Ben-Porath et al. indicate, an averaging process would seem reasonable and it would have the distinct advantage to be consistent with expected utility or welfare maximization at both the individual level (in the ex ante approach) and the social level (in the ex post method), while it concludes that w is ranked socially above z, and the latter is ranked socially above y, an ordering which fits many people's ethical intuition and which no single aggregation method seems capable of delivering in isolation.

We introduce next the formal apparatus we rely upon to link Harsanyi's classical contributions to SWFL theory. For this purpose[26], we shall restrict ourselves to risky situations and treat any decision $x \in X$ as a lottery with prizes in C. Thus, we want to associate any $c \in C$ with its probability denoted p_c^x, given x. Let us collect all $c_s^x \in C$ such that $c_s^x = c$ as well as the subset of states involved; if this set is empty, we define $p_c^x = 0$; if it is not empty, we define p_c^x as the sum of all probabilities p_s associated with the relevant states. In conclusion, we shall simply assume that the set of alternatives X is a convex subset of some linear space.

The domain restriction we just described is not the only one we rely on. Next, we restrict each R, be it *individual* or *social*, to be a *von Neumann and Morgenstern (VNM) preference ordering*, i.e. to satisfy two additional properties:

VNM-continuity of R :
$\forall\, x, x', x'' \in X,$ the sets
$\{\lambda \in [0,1] x'' R [\lambda x + (1-\lambda) x']\}$ and $\{\lambda \in [0,1] \mid [\lambda x + (1-\lambda) x'] R x''\}$
are closed in [0,1].

[26]The following developments are based on d'Aspremont and Mongin (1997). See also Mongin (1994).

VNM-independence of R :

$\forall\, x, x', x'' \in X, \forall\, \lambda \in]0,1],$

$xRx' \Leftrightarrow [\lambda x + (1-\lambda)x'']R[\lambda x' + (1-\lambda)x''].$

A VNM preference ordering R can always be represented by a function u defined on X. Moreover, in this framework, every utility representation of R is either *mixture-preserving*, i.e. $\forall\, x, y \in X, \forall\, \lambda \in [0,1], u(\lambda x+(1-\lambda)y) = \lambda u(x)+(1-\lambda)u(y)$, or a monotone transformation of a mixture-preserving utility function. In comparison with the abstract framework, in which uncertainty and risk remain implicit, we are here narrowing down both the SWFL's domain and its image set. It is the price to be paid for getting rid of any invariance requirement.

To insure that the full domain of the corresponding SWO (viz. $\Re^N$) is attainable, we shall introduce yet another domain assumption under the name of "independent prospects axiom", as it is called by Weymark (1993) and implicitly used by Harsanyi (1955), and which applies to a *single* profile of VNM preferences $(R_1, \cdots, R_i, \cdots, R_n)$ defined on X, which, as we have seen, is itself assumed to be a convex subset of some linear space.

Independent prospects:

For every $i \in N$, there are $x, y \in X$ such that xP_iy and xI_iy for all $j \in N \setminus \{i\}$.

A domain $\mathcal{D}$, consisting of all VNM-utility representations of a single profile of VNM-preference orderings satisfying independent prospects, will be called a *Harsanyi domain*. This is enough to ensure Domain attainability[27]

We can now state Harsanyi's aggregation theorem (1955, 1977), which eschews Strong neutrality, but does not dispense with Pareto indifference. We have the following:

Theorem 5.2 *Suppose that F is defined on a Harsany i domain $\mathcal{D}$ and satisfies Pareto indifference, and that, for $\mathcal{U} \in \mathcal{D}$, R_U is a VNM preference ordering. Then, there is some $\lambda^U \in \Re^N$,*

[27]Independent prospects for a given profile of VNM rankings is equivalent to having affine independence of any profile of VNM utility representations of the given rankings. Hence the range of any such utility profile has full dimension. See Fishburn (1984) and Weymark (1993).

$\lambda^U \neq 0$, *such that*

$$\forall\, x, y \in X,\;\; xR_U y \Leftrightarrow \sum_{i=1}^{n} \lambda_i^U U(x,i) \geq \sum_{i=1}^{n} \lambda_i^U U(y,i).$$

Proof: By Theorem 3.3, we may define an ordering R_U^* on $U(X)$, a convex set with nonempty interior. Now, since R_U is VNM on X, R_U^* is VNM on $U(X)$. Indeed, for VNM-independence of R_U^*, we need to show that: $\forall\, u, v, w \in U(X)$, $\forall\, \alpha \in]0,1]$,

$$uR_U^* v \Leftrightarrow [\alpha\, u + (1-\alpha)w] R_U^* [\alpha\, v + (1-\alpha)w].$$

By definition of $U(X)$, there are x, y and z in X such that $U(x) = u$, $U(y) = v$, $U(z) = w$ and, by VNM-independence of R_U,

$$xR_U y \Leftrightarrow [\alpha\, x + (1-\alpha)z] R_U [\alpha y + (1-\alpha)z],$$

so that

$$U_x R_U^* U_y \Leftrightarrow U_{\alpha x+(1-\alpha)z} R_U^* U_{\alpha y+(1-\alpha)z}.$$

Since $\mathcal{D}$ is a Harsanyi domain, each U_i is mixture-preserving, hence

$$\begin{aligned} U_{\alpha x+(1-\alpha)z} &= \alpha U_x + (1-\alpha)U_z, \\ U_{\alpha y+(1-\alpha)z} &= \alpha U_y + (1-\alpha)U_z, \end{aligned}$$

and R_U^* is VNM-independent. To derive the VNM-continuity of R_U^*, a similar argument can be used. R_U^* being a VNM ranking of the convex set $U(X) \subset \Re^N$, it has a VNM utility representation W. This mixture-preserving function is affine[28] on $U(X)$, i.e. for all $u \in U(X), W(u) = \sum_{i \in N} \beta_i u_i + \gamma$, for some vector $(\beta_1, \cdots, \beta_n)$ and some scalar γ. The result follows. ■

Because in this theorem the weights depend on the chosen profile U, it should not be taken as a characterisation of weighted utilitarianism [as remarked by Sen [1986a)], but simply as a representation theorem [see also Blackorby et al. (1990)]. Moreover, the weights could be negative or nil. However, strengthening Pareto indifference twice, into strong neutrality and

[28] For the equivalence of mixture-preserving and affine functions on convex sets, see, e.g., Coulhon and Mongin (1989).

into strong Pareto, is enough to get weighted utilitarianism with all weights positive, since, then, there is a well-defined SWO R^* on $\Re^N$ which is a VNM ranking and which coincides with any R^*_U on its domain of definition.

Corollary 5.3 (Weighted utilitarianism.) *Suppose that the SWFL F is defined on a Harsanyi domain $\mathcal{D}$, that it satisfies Strong neutrality and Strong Pareto, and that, for every $U \in \mathcal{D}$, R_U is a VNM preference ordering. Then, F is Weighted utilitarianism, with all weights positive.*

Of course, in order to characterize pure utilitarianism instead of weighted utilitarianism, it is enough to supplement our set of axioms with an anonymity requirement. This enables us to prove the following version of Harsanyi's aggregation theorem.

Theorem 5.4 (Pure utilitarianism.) *Suppose that the SWFL F is defined on a Harsanyi domain $\mathcal{D}$, that it satisfies Strong neutrality, Strong Pareto and Anonymity and that, for every $U \in \mathcal{D}$, R_U is a VNM preference ordering. Then, F is Pure utilitarianism.*

Proof: From $\mathcal{D}$ being a Harsanyi domain and Strong neutrality, formal welfarism follows. Strong Pareto and Anonymity of F implies respectively SP* and A* for the SWO R^*. Also, using the same argument as above, R^* satisfies VNM independence. This is enough to get Inv*$(a_i + bu_i)$, and hence Pure utilitarianism. Indeed, take any vector $a = (a_1, \cdots, a_n)$ and $b > 0$. If $b < 1$, we can simply put $w = a/(1-b)$ and $\lambda = b$, then apply VNM independence. If $b > 1$, clearly $uR^*v \Leftrightarrow \frac{1}{2b}(2bu)R^*\frac{1}{2b}(2bv)$, which by VNM indepence is equivalent to $(2bu)R^*(2bv)$ [letting $w \equiv 0$ and $\lambda = 1/(2b)$]. To get the conclusion, let $\lambda = \frac{1}{2}$ and $w = 2a$, and apply VNM-indepence again. ■

To prevent any misinterpretation of the last two characterisation statements when comparing them to the results presented in the previous section, we stress again that: (1) a specific form of continuity of the social ranking is implied by the definition of a VNM ordering; (2) no invariance axiom is relied on in the theorem statement, although the proof is based on the equivalence of VNM independence of R^* with Inv*(a_i+bU_i); and (3) the definition of the Harsanyi domain rests in turn on a rich domain of individual VNM utilities representating a *single* profile of individual

VNM preference relations. Yet, this has only technical significance, because we require the SWFL to be strongly neutral. If we maintain this requirement, we do not alter the substance of the results even though we adopt a domain of profiles involving many alternative VNM preference relations for every individual.

We proceed with a couple of remarks meant to clarify the link between the strongly neutral version of pure utilitarianism presented in Theorem 5.4 and relative utilitarianism, which is to come next. Choosing interpersonally comparable preference representations in order to evaluate adequately the alternatives in a specific social conflict is a momentous and thorny task, but the formal Harsanyan model we have been describing provides the ethical observer with no clue towards solving this problem, a feature shared by the bulk of SWFL theory. Indeed, suppose two individuals have the same VNM preference ordering and unequal scales are used to represent their common ordering in the profile under the ethical observer's consideration; then the associated social ranking may be strongly influenced by the specific scales chosen a priori. In brief, the formally welfarist SWFL characterized in Theorem 5.4 does not satisfy Inv$(a_i + b_i U_i)$. Relative utilitarianism, introduced in Section 2.2.4, is immune from this criticism because it involves a Kaplan normalization of *VNM utilities* in case every individual VNM preference ordering has both a maximal and a minimal element. If an anonymous aggregation method involves a process of Kaplan normalization as an essential intermediate stage, it must treat symmetrically any two individuals having the same VNM preference relation. This is in contrast with Theorem 5.4.

Dhillon (1998) and Dhillon and Mertens (1999) recently extended in a natural way the Arrovian definition of social welfare function by incorporating as basic datum a universal domain of individual VNM *preferences* and by maintaining the Harsanyi requirement that the social ranking be VNM. In this framework, they propose a set of interesting axioms and obtain two alternative characterisations of *relative utilitarianism.* It is instructive, if not economical, to recast their construct in the SWFL framework. Due to the normalization process involved in the SWFL, the final result is not influenced by the choice made a priori among the many distinct VNM-utility representations of any individual VNM *preference* profile. In other words, the SWFL must satisfy Inv$(a_i + b_i U_i)$. Moreover, the SWFL version of Independence requirement, which proves too weak for precipitating formal welfarism even though Pareto indifference does hold. Two new pieces of notation prove useful to introduce it. As before, X consists of a full set of lotteries based on diskless social decisions. The latter make up the finite subset A of degenerate lotteries and they could be interpreted as prizes. Moreover, if $A' \subset A$, we shall denote $\Xi(A')$ the full set of lotteries based on A'. Dhillon's axiom is

Independence of redundant alternatives:
$\forall U, V \in \mathcal{D}, \forall A' \subset A, \forall\, x, y \in \Xi(A')$,
$xR_U y \Leftrightarrow xR_V y$ if
(1) $\forall\, a \in A', U_a = V_a$ and (2) $\forall\, z \in X, \exists\, z' \in \Xi(A')$ such that $U_z = V_{z'}$.

Dhillon (1998) also relies on an interesting axiom called Extended Pareto applied to a variable population context; the version we offer here to facilitate comparisons is adapted for constant population and it is related to Separability, because it treats totally indifferent individuals as though they were absent:

Extended Pareto:
Let $U, V, W \in \mathcal{D}, x, y \in X,$ and $L, M \subset N$ be such that :
(1) $N = L \cup M, L \cap M = \emptyset$,
(2) $\forall\, i \in L, U_i = V_i$ and W_i is trivial, and
(3) $\forall\, i \in M, U_i = W_i$ and $V_i$is trivial;
then $xR_U y$ if both $xR_V y$ and $xR_W y$, and $xP_U y$ if both $xP_V y$ and $xP_W y$.

Our translation of Theorem 2 in Dhillon (1998) comes next:

Theorem 5.5 (Relative utilitarianism) *For a fixed and finite set A such that $|A| \geq 4$ and for all N such that $n \geq 3$ and for all profiles having at least 3 linearly independent individual utility functions, an SWFL satisfies relative utilitarianism if and only if it satisfies the following set of axioms: Strong Pareto, Extended Pareto, Inv$(a_i + b_i U_i)$, Anonymity and Independence of redundant alternatives.*

Another interesting characterisation of relative utilitarianism is provided by Dhillon and Mertens (1999). They show in particular that a pair of new axioms can replace Extended Pareto in the last statement of Theorem 5.5. The first one is called Consistency:

Consistency:

$\forall U, V \in \mathcal{D}, \forall x, y \in X, xR_U y \Rightarrow xR_V y$ if there exists $i \in N$ such that

(1)$V(x,i) = V(y,i)$

(2)$\forall j \in N \setminus \{i\}, U_j = V_j,$ and

(3)U_i is trivial.

Due to the binaries involved in the last clause, Consistency is neither implied by Inv$(a_i+b_iU_i)$, nor by Separability. The remaining axiom has a distinctly technical flavor; it is a complementary continuity requirement:

DM-Continuity of F :

Consider any sequence $(U^\ell)_{\ell=1}^\infty \subset \mathcal{D}$ such that

$\forall i \in N \setminus \{n\}, \forall a \in A \setminus \{a_0\},$

$U^\ell(a,i) = U^*(a,i)$ for $\ell = 1, 2, \cdots,$ whereas $U^\ell(a_0, n)$ converges to $U^*(a_0, n)$.

If the corresponding social ranking converges to $\lim R_{U^\ell}$,

then either R_{U^*} is trivial or $R_{U^*} = \lim R_{U^\ell}$.

Relative utilitarianism is worth comparing to the methods of voting based on such scoring functions as the Borda standardisation described in example 2.2.7: the latter also rest on a purely utilitarian formula applied to intermediate individual scores. Young (1975) provides a remarkable characterisation (which is also based on another version of Extended Pareto). These methods do not require asset of alternatives endowed with a specific structure such as convexity. Computing individual scores in an ordinal procedure, so that Inv$\phi_i(U_i)$ is satisfied, while Arrovian Independence is violated and formal welfarism does not hold. The SWFLs based on the summation of individual scores may have merits as voting rules, but they seem even less fit for social evaluation than relative utilitarianism, which does not require that intervals between different levels of preferences be equalised across individuals, but requires only equalised maxima and minima. This has some relation to, but is less demanding than, requiring a "fundamental preference" as in Harsanyi's (1953, 1977) Impartial Observer theorem on utilitarianism. Yet, as our discussion of Diamond's (1967) example suggests, the social ranking implied by relative utilitarianism is not always in agreement with common ethical intuitions.

To conclude this subsection, we mention again the elegant contribution by Epstein and Segal (1992). After dissociating the *ex post* aggregation method from the *ex ante* approach, they rely on the latter to separate the social aversion to inequality from individual risk aversion, and they obtain a characterisation of a family of SWOs involving quadratic aggregation of individual expected utilities. To address the Diamond (1967) problem, they introduce an axiom of preference for randomization as follows: consider any pair x, y of socially indifferent lotteries (or social decisions), such that two individuals have conflicting strict preferences about them; then, both lotteries are socially dominated by the two-stage lottery based on the toss of a fair coin to decide which of x and y will be chosen eventually. This axiom does not allow for flat line segments within social-indifference loci; it is thus inconsistent with utilitarianism.

5.2 On some egalitarian social rankings

Instead of the three characterisations of the lexicon principle presented in Section 4, all of which invoke invariance properties, we turn next to alternative derivations. They are obtained by dropping invariance and separability completely, and by substituting a much stronger equity requirement. Hammond's equity axiom will be used first:

Theorem 5.6 *Any SWO R^* is the unique formally welfarist associate of the lexicon principle if and only if it satisfies Strict Pareto*, Anonymity*, and Hammond's equity* principle.*

If we strengthen Hammond's equity* principle, we obtain the lexicon principle applied to cases where only two persons are not indifferent. This property is contagious, and it can be extended to social conflicts involving gradually more and more individuals: as a matter of fact, this extension is carried out when the standard strategy of proof is followed. For details, we refer the reader to Bossert and Weymark (2000), who also offer a nice account of the history of this result. Hammond's (1976a) role seems essential, but he recognizes his debt to Strasnick (1976). Moreover, he did not use an SWFL framework, and the generalized Arrovian social welfare function he defined for his purpose hid the fact that no invariance axiom was in effect necessary. Tungodden (1999) showed recently that any member of a set of alternative properties can be substituted for Hammond's equity axiom in the last theorem. They have in common the notion of a deprived group of individuals associated with every pair of evaluation vectors. Moreover, the social ranking of this pair cannot contradict bluntly the strict preference relation of the least favored non-indifferent individual provided he/she belongs to the deprived group. Among the many deprivation criteria mentioned in the specialised literature, we shall consider only a one-parameter family, which reminds one of a popular poverty definition. It will be convenient to work with $\mathcal{G}_N$, the set of well-ordered evaluation vectors, within which the identity of the k-worst off individuals does not change. Given some number $\alpha \in (0,1]$, we shall say that $i \in N$ belongs to the α-deprived group pertaining to $\{u,v\} \subset \mathcal{G}_N$ if and only if $u_i < \alpha \sum_{j\in N}(u_j/n)$ or $v_i < \alpha \sum_{j\in N}(v_j/n)$.

Theorem 5.7 *Suppose R^* is an SWO defined on $\Re^N$. Then, R^* is the lexicon SWO if it satisfies Strict Pareto*, Anonymity* and $\exists\, \alpha \in (0,1]$ for which, $\forall\, u,v \in \mathcal{G}_N$, vR^*u whenever $v_k > u_k$ for some member k of the α-deprived group pertaining to $\{u,v\}$, whereas $v_i = u_i$ for any individual i, $i < k$.*

The new equity axiom introduced as the last theorem condition seems to restrict to the relevant deprived group the veto power granted to the last-favored non-indifferent individual. Yet, as Tungodden shows, this property is contagious and can be extended to all individuals. He also proves that more general families of definition of the deprived group could be used in Theorem 5.6 without altering the substance of the result. On the other hand, Tungodden remarks that the theorem does not go through if a given percentile (e.g., the median) is substituted for the mean in the definition of the α-deprived group.

In a highly original paper, Barberà and Jackson (1988) provide among other things an alternative characterisation of the leximin principle. They rely on an SWO framework, which is extended to accommodate a variable set of individuals. Our generic notation N thus designates henceforth any subset of the set of positive integers, the latter being interpreted as the set of potential agents, and some axioms are designed as robustness properties of the social ranking R^*_N when N undergoes certain alterations. The goal of the exercise is to characterize social rankings which apply only to societies having a constant population; we are not pursuing here an optimal population theory. Barberà and Jackson's lexicon characterisation eschews all invariance axioms. As to equity properties, they are not so demanding as Hammond or Tungodden, since they rely only on Anonymity* and a version of Convexity*. What is more debatable in our context is their axiom of *Independence of duplicated individuals*, to wit: For any N, for any $i, j \in N$, for any $u, v \in \Re^N$, such that $u_i = u_j$ and $v_i = v_j$, $uP^*_N v$ if and only if $u_{-i}P^*_{N\setminus\{i\}}v_{-i}$, where the subscripts mean that individual i is no longer a member of the economy. As the authors write, "... the axiom constitutes a strong value judgement: that it is not the number of individuals in a welfare group, but the level of welfare within the groups, that should count in comparing social states". The axiom would of course be highly appropriate for the theory of individual decision under complete ignorance.

To conclude this section, we sample some recent work by Fleurbaey and Maniquet (2000b) who adopt a structure much more specific than those we have dealt with so far. Indeed, they study the set of Edgeworth boxes with a constant number ℓ of private goods. An economy e consists of three elements: (1) A set N of individuals defined as in the previous paragraph; (2) a profile U_N consisting of an $|N|$-tuple of continuous, strictly increasing and quasi-concave self-oriented individual utility functions which are defined over the conventional individual consumption set $\Re^\ell_+$, and (3) an ℓ-tuple $\omega \in \Re^\ell_{++}$ designating the quantity of each good available as social endowment. By letting these three elements vary to the largest extent consistent with their definition, we obtain a rich domain $\mathcal{E}$ of economies, to which we devote our normative investigation. We are interested in the social evaluation of a set of alternatives that we define as follows: $X_N = \Re^{\ell|N|}_+$ i.e. the $|N|$-fold cartesian product of $\Re^\ell_+$. We want to study social mappings which associate with every $e \in \mathcal{E}$ a social preference ordering R_e over the corresponding X_N. To complete our notation, we let x_N designate a typical element of X_N, whereas $x_{i,N}$ stands for a generic component of x_N.

As the number $|N|$ of individuals populating an economy $e \in \mathcal{E}$ may be any positive integer, we must stretch the SWFL definition to accommodate the social mappings we just defined. Moreover, even if we restrict our attention to some specific N, we do not face an ordinary SWFL. Indeed, even though we want to obtain a social ranking of the same set of alternatives, viz. X_N, the social endowment ω is allowed to influence R_e. Eventually, we obtain an SWFL by restricting attention to the subset of elements of $\mathcal{E}$ generated by a given pair (N, ω).

The family of social rankings we want to study combines a utility-based version of Rawls's difference principle with the concept of egalitarian equivalence introduced by Pazner and Schmeidler (1978). Therefore, we call it for short the *RPS (i.e. Rawls-Pazner-Schmeidler) family.* It is convenient to proceed in two steps to define it:

1. We associate with every individual component U_i of the profile U_N of any $e \in \mathcal{E}$ a canonical utility representation of the preference relation underlying U_i. We denote it $W_{i,e} := \Re_+^\ell \rightarrow \Re_+$ and we define it implicitly by letting its image $w_i = W_{i,e}(x_i)$ be i's equivalent share in the social endowment vector ω, i.e. $U_i(x_i) = U_i(w_i, \omega)$.

2. For any $e \in \mathcal{E}$, any $x_N, x'_N \in X_N, x_N P_e x'_N$ if $\min_{i \in N}\{w_i\} > \min_{i \in N}\{w'_i\}$.

Our next task is to describe a set of axioms sufficient to precipitate a social ordering belonging to the RPS family. We start with two cross-economy robustness properties, and we introduce first the *Replication invariance* axiom, the technicalities of which will not retain us: it says that for any $e \in \mathcal{E}$, any $x_N, x'_N \in X_N, x_N R_e x'_N$ implies the same social ranking of any v-fold replication of either allocation, assuming that e itself is replicated v times. An analogous axiom was used by Debreu and Scarf (1963) in their study of convergence of the core to a competitive equilibrium.

The *e-Separability* axiom is slightly more delicate: it provides a material condition for deleting a subset $M \subset N$ from the a priori given economy $e \in \mathcal{E}$, assuming that both ω and the restricted profile $U_{N \setminus M}$ remain unaltered. We denote the reduced economy $g \in \mathcal{E}$. The axiom runs as follows: for any $e \in \mathcal{E}$, any $x_N, y_N \in X_N$, $x_N R_e y_N$ implies $x_{N \setminus M} R_g y_{N \setminus M}$ if $\forall\, i \in M$, $x_{i,N} = y_{i,N}$.

It seems worth pointing out that, despite the binary nature of its implication, the triggering condition is much narrower than its welfarist counterpart, because it is based on a maintained consumption level among the members of M. An analogous remark is also valid for our next axiom, which is meant to express an equity norm. It provides conditions under which utility transfers do not hurt society, whenever two individuals having the same preferences are initialy not on the same indifference curve.

Conditional Hammond equity:
For any $e \in \mathcal{E}$, any $x_N, y_N \in X_N$,
$x_N R_e y_N$ if there exist $j, k \in N$ such that
$U_k = U_j$ and $U_j(y_{j,N}) > U_j(x_{j,N}) > U_j(x_{k,N}) > U_j(y_{k,N}),$ whereas
$\forall\, i \in N, j \neq i \neq k, x_{i,N} = y_{i,N}$.

Another equity norm comes next; it is concerned with a highly specific allocation, so that an exact welfarist translation does not exist.

Equal split:
For any $e = \mathcal{E},$ any $x_N, y_N \in x_N$,
$x_N P_e y_N$ if $\forall\, i \in N, x_{i,N} = (\omega)/|N|$ and
there exists some $j \in N$ for which $U_j(x_{i,N}) > U_j(y_{j,N})$.

Our last theorem is a weak version of Theorem 3 by Fleurbaey and Maniquet (2000b). Our weakening is meant to facilitate the comparison with the bulk of SWFL theory:

Theorem 5.8 (RPS family) *A given member of the RPS family of social orderings is associated with every $e \in \mathcal{E}$ whenever the social mapping is required to satisfy the following set of axioms: Weak Pareto, Pareto Indifference, Inv($\varphi_i(U_i)$), Replication invariance, e-Separability, Equal split and Conditional Hammond equity.*

We remark that no version of interpersonal welfare comparability is used: two agents having the same self-oriented preferences are treated alike because there is no reason for doing otherwise within the limits of our model. In principle, the approach we just exemplified does not preclude the social ordering farm taking also into account idiosyncrasies for which individuals cannot be held responsible, such as metabolic peculiarities; see Fleurbaey (1995) and Vol. 2 of this Handbook. Moreover, the existence of an intermediate stage of canonical preference representation calls for two further comments: (1) it seems to take care of the expensive taste criticism adduced against utility-based social evaluation by Rawls among others, and (2) as Dhillon and Mertens (1999), it is inconsistent with the Arrovian independence property and formal welfarism does not hold.

6 Conclusion

The SWFL concept can be of help to the ethical observer aiming at an appropriate social evaluation. It helps organize the inner debate and ask the relevant questions: what is the set of issues, in what ways should society be concerned with individual consequences of its decisions, how should they be adequately represented by individual scores and how should one process this information to obtain an appropriate social ranking? Until recently, the SWFL literature has been mostly helpful in answering the latter query. It does indeed provide an enlightening analysis of various SWFLs as it compares their relative merits by means of stylized axiomatic properties. If the formally welfarist framework is accepted, a good deal of interpersonal comparability does seem required for obtaining social rankings which satisfy minimal equity requirements. Among formally welfarist SWFLs, the lexicon principle and pure utilitarianism might seem to be the candidates displaying the most attractive set of properties. In its own way, each one is moderately demanding in terms of interpersonal comparisons. Critics of the lexicon principle object to the absolute priority it gives to favoring the least advantaged individual, without regard for the number of losers and the average size of their loss. On the other hand, the examples adduced respectively by Diamond (1967) and by Pen-Porath et al. (1997) for social evaluation under risk and uncertainty are rather damaging for utilitarianism. Among the SWFLs which are immune from either criticism, the other members of the generalized Gini family seem least demanding in terms of interpersonal comparisons.

Once theoretical analysis has delivered its conclusions and an SWFL is chosen, the most difficult task is to select an adequate profile; indeed, the rest of the selection work becomes more or less mechanical and it can usually be fed into a computer, as it consists of looking for the set of best social decisions associated with any feasible subset of alternatives. If several evaluation profiles seem reasonable in consideration of the context at hand, the ethical observer faces the embarrassment of a multiplicity of candidate social rankings. Upon closer examination, ethical intuitions may be found in disagreement with some of them, and the latter can be discarded. The remaining social rankings can be intersected; this procedure is useful if the resulting incomplete relation has enough bite, as it has proved to be the case with Lorenz dominance in empirical analysis of the pure distribution problem. See, for instance, Shorrocks (1983). Failing this, one could submit the social rankings to another round of aggregation. However, as Roberts (1995) and Suzumura (1996) show, such a procedure is plagued with Arrovian difficulties.

Choosing on *a priori* grounds interpersonally comparable individual evaluation counts is probably the least consensual element in the SWFL approach, and several authors have chosen to rely only on what seems to be the weakest possible comparability assumption. Following their approach, some individual characteristics are singled out as essentially mattering for equity judgments, so that it is socially desirable that two individuals displaying the same relevant characteristics be treated similarly. This equal-treatment principle must prevail even though the rest of the characteristics are individually differentiated; the latter are considered as ethically negligible or secondary. As Kolm (1996b) insists, there is an ethical drive towards equality if any reason that might justify inequality is lacking. If individual preferences are singled out as an essential element for elaborating equity judgments, a non comparability invariance axiom is always adhered to if only implicitly, as in positive economic theory. To recall, this can be consistent with cardinality as in Nash equilibrium in mixed strategies, an invariance axiom we have abbreviated as $\text{Inv}(a_i + b_i U_i)$, or it can be only ordinally measurable as in Nash equilibrium in pure strategies, an SWFL property we have denoted $\text{Inv}(\varphi(U_i))$. This approach is *inconsistent with formal welfarism*, because the filtering of information the latter stricture implies leaves out essential aspects of the problem at hand.

Moreover, once this extreme informational parsimony is abandoned, we realize that the abstract Arrovian framework and its natural universal domain do not seem conducive to any SWFL that is fit for social evaluation, even though it may be promising as a constitutional rule, as the Borda method of voting. Nowadays, a majority of researchers seem to have turned towards less sweeping research goals; they prefer to deal with the conditions of justice and equity in a variety of specific contexts, in order to take advantage of the extra structure they provide. Besides Roemer (1986, 1996) and Young (1995), we refer the reader to Moulin (2002), Thomson (2002) and Fleurbaey and Maniquet (2002). Following this approach, the feasible set is more narrowly specified, but on the other hand, it is allowed to vary. Together with X and U, the set of individuals N may also be treated as an independent variable. Very interesting solution concepts whose purpose is to recommend a set of socially best decisions for every economy in the domain under scrutiny have been studied in the literature. The proposed solution can always be interpreted as a social ranking consisting only of two indifference curves, but this usually leads to violations of the Pareto principle, and it appears to be too rough for second-best or reform problems, where the structure of the set of alternatives is likely to be too unwieldy for an axiomatic analysis.

A more promising research strategy may be to try to associate a set of social rankings to a set of solutions of the allocation problem. This approach was developed successfully by Young (1987) in a taxation problem where taxable income is treated as a fixed parameter that may vary across individuals. Young shows that two sets of properties are equivalent; one of them defines a family of solutions, whereas the other one pertains to social orderings whose subset of best elements always coincides with a solution in the family. Young (1987) obtains analogous results for bankruptcy and profit-sharing models.

As we have seen, Dhillon (1998) and Dhillon and Mertens (1999) dealt recently with a more general context: the complete domain of lotteries one can define over a give nest of pure abstract alternatives. On the other hand, Fleurbaey and Maniquet obtained characterisations for two classical economic domains: the set of pure exchange economies [Fleurbaey and Maniquet (2000b)], and the set of two-good artisan economies with linear production technologies [Fleurbaey and Maniquet (2000a)]. In every case, a numerical profile is obtained as an intermediate product, which is further processed as a utilitarian or a lexicon practitioner would do to deliver eventually an appropriate social ranking. One can thus speak of *interim* or *ex post* interpersonal comparability of the intermediate evaluation profile which is induced by the proposed procedure. Yet, to our knowledge, it has never been technically fruitful to separate into two stages the axiomatic derivation of the social ranking, because the intermediate profile appears as a pure by-product of the analysis. It seems unlikely that the arguments proving useful in formally welfarist characterisations may be of any help for normative analysis taking advantage of peculiarly structured economic environments. Due to its discarding of Arrovian Binary independence, the latter approach is technically very demanding and its practitioners may soon face at least temporary feasibility limits. Whether it will eventually succeed in superseding the abstract SWFL framework is for the future to decide. So far, empirical analysis seems to lean in its favor, but its support comes from a specific angle, which is more positive than normative, and it will not necessarily convince the ethical observer we have been alluding to since the beginning of this chapter.

References

Arrow, K.J. (1963), *Social Choice and Individual Values*, 2nd Edition (Wiley, New Yok); 1st Edition, 1951.

Arrow, K.J. (1977), "Extended sympathy and the possibility of social choice", *American Economic Economic Review Papers and Proceedings* 67: 219-225.

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

Barberà, S., and M.O. Jackson (1988), "Maximin, lexicon and the protective criterion: characterisations and comparisons", *Journal of Economic Theory* 46: 34–44.

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

Ben-Porath, E., I. Gilboa and D. Schmeidler (1997), "On the measurement of inequality under uncertainty", *Journal of Economic Theory* 75: 194–204.

Bentham, J. (1789), An Introduction to the Principles of Morals and Legislation (Payne, London). Reprinted in 1970 in: J.M. Burns and H.L.A. Hart, eds. (Athlone Press, London).

Bergson, A. (1954), "On the concept of social welfare", *Quarterly Journal of Economics* 68:233–252.

Blackorby, C., and D. Donaldson (1980), "A theoretical treatment of indices of absolute inequality", *International Economic Review* 21:107–136.

Blackorby, C., and D. Donaldson (1982), "Ratio-scale and translation-scale full interpersonal comparability without domain restrictions: admissible social evaluation functions", *International Economic Review* 23: 249–268.

Blackorby, C., D. Donaldson and J.A. Weymark (1984), "Social choice with interpersonal utility comparisons: a diagrammatic introduction", *International Economic Review* 25: 327–356.

Blackorby, C., D. Donaldson and J.A. Weymark (1990), "A welfarist proof of Arrow's theorem", *Recherches Economiques de Louvain* 56: 259–286.

Blackorby, C., W. Bossert and D. Donaldson (2002), "Utilitarianism and the theory of justice", in: K.J. Arrow, A.K. Sen and K. Suzumura, eds., *Handbook of Social Choice and Welfare*, Vol. 1 (Elsevier, Amsterdam) Chapter 11, this volume.

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

Blau, J.H. (1971), "Arrow's theorem with weak independence", *Economica* 38:413–420.

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

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

Bossert, W., and J.A. Weymark (2000), "Utility in social choice", in: S. Barberà, P. Hammond and C. Seidl, eds., *Handbook of Utility Theory, Vol. II, Applications and Extensions* (Kluwer, Dordrecht).

Broome, J. (1993), "A cause of preference is not an object of preference", *Social Choice and Welfare* 10: 57–68.

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

d'Aspremont, C. (1985), "Axioms for social welfare orderings", in: L. Hurwicz, D. Schmeidler and H. Sonnenschein, eds., *Social Goals and Social Organizations: Essays in Memory of Elisha Pazner* (Cambridge University Press, Cambridge) pp. 19–76.

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

d'Aspremont, C., and P. Mongin (1997), "A welfarist version of Harsanyi's aggregation theorem", CORE Discussion paper; also in: M. Salles and J.A. Weymark, eds., *Justice, Political Liberalism and Utilitarianism, Volume in Honor of John Harsanyi and John Rawls* (Cambridge University Press, Cambridge) forthcoming.

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

Debreu, G., and H. Scarf (1963), "A limit theorem on the core of an economy", *International Economic Review* 4: 235–246.

Denicolò, N. (1999), "A characterization of utilitarianism without the transitivity axiom", *Social Choice and Welfare* 16:273–278.

Deschamps, R., and L. Gevers (1977), "Separability, risk-bearing and social welfare judgements", *European Economic Review* 10: 77–94.

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

Dhillon, A. (1998), "Extended Pareto rules and relative utilitarianism", *Social Choice and Welfare* 15: 521–542.

Dhillon, A., and J.-F. Mertens (1999), "Relative utilitarianism", *Econometrica* 67:417–498.

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

Dixit, A.K. (1980), "Interpersonal comparisons and social welfare functions", Mimeo (University of Warwick).

Dutta, B. (2002), "Inequality, poverty and welfare", in: K.J. Arrow, A.K. Sen and K. Suzumura, eds., *Handbook of Social Choice and Welfare*, Vol. 1 (Elsevier, Amsterdam) Chapter 12, this volume.

Ebert, U. (1987), "A note on social welfare orderings", *European Economic Review* 31: 1145–1147.

Ebert, U. (1988a), "Rawls and Bentham reconciled", *Theory and Decision* 24: 215–223.

Ebert, U. (1988b), "Measurement of inequality: an attempt at unification and generalization", *Social Choice and Welfare* 5: 147–169.

Epstein, L.G., and U. Segal (1992), "Quadratic social welfare functions", *Journal of Political Economy* 100: 691–711.

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

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

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

Fleurbaey, M. (1995), "Three solutions of the compensation problem", *Journal of Economic Theory* 65: 505–521.

Fleurbaey, M., and F. Maniquet (1996), "Utilitarianism versus fairness in welfare economics", in: M. Salles and J.A. Weymark, eds., *Justice, Political Liberalism and Utilitarianism: Themes from Harsanyi and Rawls* (Cambridge University Press).

Fleurbaey, M., and F. Maniquet (2000a), "Fair social orderings with unequal production skills", Mimeo (Namur University, Belgium).

Fleurbaey, M., and F. Maniquet (2000b), "Fair social orderings in division economies", Mimeo (Namur University, Belgium).

Fleurbaey, M., and F. Maniquet (2002), "Compensation and responsability", in: K.J. Arrow, A.K. Sen and K. suzumura, eds., *Handbook of Social Choice and Welfare*, Vol. 2 (Elsevier, Amsterdam).

Gaertner, W., P.K. Pattanaik and K. Suzumura (1992), "Individual rights revisited", *Economica* 59: 161–177.

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

Gilboa, I., and D. Schmeidler (1989), "Maxmin expected utility with non-unique prior", *Journal of Mathematical Economics* 18: 141–153.

Hammond, P.J. (1976a), "Equity, Arrow's conditions, and Rawls' difference principle", *Econometrica* 44: 793–804.

Hammond, P.J. (1976b), "WHy ethical measures of inequality need interpersonal comparisons", *Theory and Decision* 7: 263–274.

Hammond, P.J. (1979), "Equity in two person situations: some consequences", *Econometrica* 47::1127–1135.

Hammond, P.J. (1982), "Utilitarianism, uncertainty and information", in: A.K. Sen and B. Williams, eds., *Utilitarianism and Beyond* (Cambridge University Press, Cambridge) pp. 85–102.

Hammond, P.J. (1988), "Consequentialist foundations for expected utility", *Theory and Decision* 25: 25–78.

Hammond, P.J. (1991), "Interpersonal comparisons of utility: Why and how they are and should be made", in: J. Elster and J.E. Roemer, eds., *Interpersonal Comparisons of Well-Being* (Cambridge University Press, Cambridge) pp. 200–254.

Hammond, P.J. (1998), “Interpersonally comparable utility”, in: S. Barberà, P. Hammond and C. Seidl, eds., *Handbook of Utility Theory*, Vol. 2 (Kluwer, Dordrecht).

Hammond, P.J. (1999), “Roberts’ weak welfarism theorem: a minor correction”, Department of Economics Working Paper 99-021 (Stanford University).

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

Harsanyi, J.C. (1955), “Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility”, *Journal of Political Economy* 63: 309–321.

Harsanyi, J.C. (1977), *Rational Behavior and Bargaining Equilibrium in Games and Social Situations* (Cambridge University Press, Cambridge).

Hicks, J.R. (1959), *Essays in World Economics* (Clarendon Press).

Kaneko, M. (1984), “On interpersonal utility comparisons”, *Social Choice and Welfare* 1: 165–175.

Kelsey, D. (1987), “The role of information in social welfare judgements”, *Oxford Economic Papers* 39; 301–317.

Kemp, M.C., and Y.-K. Ng (1976), “On the existence of social welfare functions, social orderings and social decision functions”, *Economica* 43: 59–66.

Kolm, S.-Ch. (1966), “The optimal production of social justice”, Proceedings of an IEA conference held in Biarritz; reprinted, 1969, in: H. Guitton and J. Margolis, eds., *Public Economics* (Macmillan, London) pp. 145–200.

Kolm, S.-Ch. (1972), *Justice et Equité* (CNRS, Paris).

Kolm, S.-Ch. (1996a), *Modern Theories of Justice* (MIT Press, Cambridge, Mass.)

Kolm, S.-Ch. (1996b), “Rational just social choice”, in: K.J. Arrow, A.K. Sen and K. Suzumura, eds., *Social Choice Re-Examiend*, Vol. II (Macmillan Press, London) pp. 167–196.

Kolm, S.-Ch. (1998), "Chance and justice: social policies and the Harsanyi-Vickrey-Rawls problem", *European Ecoomic Review* 42: 1393–1416.

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

Krause, U. (1995), "Essentially lexicographic aggregation", *Social Choice and Welfare* 12: 233–244.

Lockwood, B. (1984), "Social choice, interpersonal comparability and welfare economics", in: F. van der Ploeg, ed., *Mathematical Methods in Economics* (Wiley, New York) pp. 371–400.

Luce, R.D., and H. Raiffa (1957), *Games and Decisions: Introduction and Critical Survey* (Wiley, New York).

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

Milnor, J. (1954), "Games against nature", in: R.M. Thrall, C.H. Coombs and R.L. Davis, eds., *Decision Processes* (Wiley, New York) pp. 49–59.

Mirrlees, J.A. (1982), "The economic uses of utilitarianism", in: A.K Sen and B. Williams, eds., *Utilitarianism and Beyond* (Cambridge University Press, Cambridge) pp. 63–84.

Mongin, P. (1994), "Harsanyi's aggregation theorem: multi-profile version and unsettled questions", *Social Choice and Welfare* 11: 331–354.

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

Moulin, H. (1983), *The Strategy of Social Choice, Advanced Textbooks in Economics* (North-Holland, Amsterdam).

Moulin, H. (1988), *Axioms of COoperative Decision Making* (Cambridge University Press, Cambridge).

Moulin, H. (2002), "Axiomatic cost and surplus-sharing", in: K.J. Arrow, A.K. Sen and K. Suzumura, eds., *Handbook of Social Choice and Welfare*, Vol. 1 (Elsevier, Amsterdam) Chapter 6, this volume.

Nash, J.F. (1950), "The bargaining problem", *Econometrica* 18: 155–162.

Parks, R.P. (1976), “An impossibility theorem for fixed preferences: a dictatorial Bergson-Samuelson social welfare function”, *Review of Economic Studies* 43: 447–450.

Pazner, E.A., and D. Schmeidler (1978), “Egalitarian-equivalent allocations: a new concept of economic equity”, *Quarterly Journal of Economics* 92: 671–687.

Pigou, A.C. (1920), *The Economics of Welfare* (Macmillan, London); 4th Edition 1950.

Pollak, R.A. (1979), “Bergson-Samuelson social welfare functions and the theory of social choice”, *The Quarterly Journal of Economics* 93: 73–90.

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

Roberts, K.W.S. (1980a), “Interpersonal comparability and social choice theory”, *Review of Economic Studies* 47: 421–439.

Roberts, K.W.S. (1980b), “Social choice theory: the single-profile and multi-profile approaches”, *Review of Economic Studies* 47: 441–450.

Roberts, K.W.S. (1995), “Valued opinions or opinionated values: the double aggregation problem”, in: K. Basu, P. Pattanaik and K. Suzumura, eds., *Choice, Welfare, and Development* (Oxford University Press, Oxford) pp. 140–165.

Roemer, J.E. (1986), “The miscarriage of bargaining theory and distributive justice”, *Ethics* 97: 88–110.

Roemer, J.E. (1990), “Welfarism and axiomatic bargaining theory”, *Recherches Economiques de Louvain* 56: 287–301.

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

Sarin, R., and P. Wakker (1994), “Folding back in decision tree analysis”, *Management Science* 40: 625–628.

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

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), "Utilitarianism and welfarism", *Journal of Philosophy* 76: 463–489.

Sen, A.K. (1980), "Equality of what?", *The Tanner Lectures on Human Values*, Vol. 1 (Cambridge University Press, Cambridge); reprinted in A.K. Sen, 1982, *Choice, Welfare and Measurement* (Basil Blackwell, Oxford) pp. 353–369.

Sen, A.K. (1982), *Choice, Welfare and Measurement* (Basil Blackwell, Oxford).

Sen, A.K. (1986a), "Social choice theory", in: K.J. Arrow and M.D. Intriligator, eds., *Handbook of Mathematical Economics*, Vol. III (North-Holland, Amsterdam) pp. 1073–1181.

Sen, A.K. (1986b), "Information and invariance in normative choice", in: W. Heller, R. Starr and D. Starrett, eds., *Social Choice and Public Decision-Making, Essays in Honor of K.J. Arrow*, Vol. 1 (Cambridge University Press, Cambridge) pp. 29–55.

Sen, A.K. (1997), "Individual preference as the basis of social choice", in: K.J. Arrow, A.K. Sen and K. Suzumura, eds., *Social Choice Re-Examined*, Vol. I (Macmillan Press, London) pp. 15–38.

Sen, A.K., and B. Williams (1982), "Introduction", in A.K. Sen and B. Williams, eds., *Utilitarianism and Beyond* (Cambridge University Press, Cambridge) pp. 1–21.

Shapley, L.S. (1969), "Utility comparison and theory of games", in: *La Décision: Agrégation et Dynamique des Ordres de Préférences* (Editions du CNRS, Paris) pp. 251–263.

Shapley, L.S., and M. Shubik (1975), "Game theory in economics, Chapter 4, References and utility", Rand report R-904/4-NSF (Rand Corporation, Santa Monica).

Shorrocks, A.F. (1983), "Ranking income distributions", *Economica* 50: 3–17.

Shubik, M. (1982), *Game Theory in the Social Sciences* (MIT Press, Cambridge, Mass.).

Sprumont, Y. (1996), "Axiomatizing ordinal welfare egalitarianism when preferences may vary", *Journal of Economic Theory* 68: 77–110.

Strasnick, S. (1976), “Social choice and the derivation of Rawls’ difference principle”, *Journal of Philosophy* 73: 184–194.

Suppes, P. (1957), “Two formal models for moral principles”, Techn. Report No 15, Office of Naval Research Contract 225(17) (Applied Math. & Statistics Lab., Stanford University) pp. 17–18.

Suppes, P. (1966), “Some formal models of grading principles”, *Synthese* 6: 284–306.

Suzumura, K. (1996), “Interpersonal comparisons and the possibility of social choice”, in: K.J. Arrow, A.K. Sen and K. Suzumura, eds., *Social Choice Re-Examined*, Vol. II (Macmillan Press, London) pp. 202–229.

Suzumura, K. (2002), “Welfare, rights and social choice procedures”, in: K.J. Arrow, A.K Sen and K. Suzumura, eds., *Handbook of Social Choice and Welfare*, Vol. 2 (Elsevier, Amsterdam).

Thomson, W. (2002), “Fair allocations”, in: K.J. Arrow, A.K. Sen and K. Suzumura, eds., *Handbook of Social Choice and Welfare*, Vol. 2 (Elsevier, Amsterdam).

Tsui, K.-Y., and J.A. Weymark (1997), “Social welfare orderings for ratio-scale measurable utilities”, *Economic Theory* 10: 241–256.

Tungodden, B. (1999), “The distribution problem and Rawlsian reasoning”, *Social Choice and Welfare* 16: 599–614.

Vickrey, W.S. (1945), “Measuring marginal utility by the reactions to risk”, *Econometrica* 13: 319–333.

Weymark, J.A. (1981), “Generalized Gini inequality indices”, *Mathematical Social Sciences* 1: 409–430.

Weymark, J.A. (1993), “Harsanyi’s social aggregation theorem and the weak Pareto principle”, *Social Choice and Welfare* 10: 209–221.

Wilson, R.B. (1972), “Social choice theory without the Pareto principle”, *Journal of Economic Theory* 5: 478–486.

Yaari, M.H. (1978), “Rawls, Edgeworth, Shapley, Nash: theories of distributive justice re-examined”, *Journal of Economic Theory* 24: 1–30.

Young, H.P. (1975), “Social choice scoring functions”, *SIAM Journal of Applied Mathematics* 28: 824–838.

Young, H.P. (1987), “On dividing an amount according to individual claims or liabilities”, *Mathematics of Operations Research* 12: 398–414.

Young, H.P (1995), *Equity: in Theory and in Practice* (Princeton University Press, Princeton, NJ).