# Multidimensional Welfare Aggregation*

CHRISTIAN LIST
Nuffield College, Oxford OX1 1NF, U.K.; E-mail christian.list@nuffield.oxford.ac.uk

Accepted subject to revision by *Public Choice*
Revised 3 December 2002

**Abstract.** Most accounts of welfare aggregation in the tradition of Arrow's (1951/1963) and Sen's (1970/1979) social-choice-theoretic frameworks represent the welfare of an individual in terms of a single welfare ordering or a single scalar-valued welfare function. I develop a multidimensional generalization of Arrow's and Sen's frameworks, representing individual welfare in terms of *multiple* personal welfare functions, corresponding to multiple 'dimensions' of welfare. I show that, as in the one-dimensional case, the existence of attractive aggregation procedures depends on certain informational assumptions, specifically about the measurability of welfare and its comparability not only across individuals but also across dimensions. I state several impossibility and possibility results. Under Arrow-type conditions, insufficient comparability across individuals leads to dictatorship of a single individual, while insufficient comparability across dimensions leads to dominance of a single dimension. Given sufficient comparability both across individuals and across dimensions, a range of possibilities emerges. I discuss the substantive implications of the results.

## 1. Introduction

The concern of this paper is the problem of aggregating the welfare of the individual members of a group into the corresponding welfare of the group *as a whole*. Individual welfare is typically assessed in terms of some normatively relevant evaluation standard. Examples of such evaluation standards, or 'currencies of welfare', are money, indices of resources, utility, or Rawlsian primary goods.

* Previous versions of this paper were presented at the seminar of the Philosophy Program of the Research School of Social Sciences, Australian National University, Canberra, in June 2000, at an NSF Workshop on Models of Individual and Public Choice, held at the University of California at Irvine, in July 2000, and at the Annual Meeting of the Public Choice Society in New Orleans, Louisiana, in March 1999. The author wishes to express his gratitude to the seminar and conference participants at these occasions and to A. B. Atkinson, Geoffrey Brennan, John Dryzek, Robert Goodin, Iain McLean, Kevin Roberts, John Weymark, and an anonymous reviewer for very helpful comments and discussion. Financial support of the (British) Economic and Social Research Council, the German National Merit Foundation and the Goodhart Fund is gratefully acknowledged.

By Arrow's theorem (1951/1963), there exists no procedure for aggregating individual welfare orderings over a set of alternatives into collective, or social, ones where the procedure satisfies a set of minimal conditions (transitivity of social orderings, universal domain, the weak Pareto principle, independence of irrelevant alternatives and non-dictatorship). Sen (1970/1979) has shown that Arrow's theorem depends crucially on the assumption that individual welfare is not interpersonally comparable: given interpersonal comparability of welfare levels or units, the impossibility result disappears. The question of whether or not interpersonally significant information is accessible depends on the chosen 'currency of welfare'. The question about the *significant* content, personal and interpersonal, of individual welfare information is called the question of *measurability and interpersonal comparability*.

Sen's (1970/1979) framework is more general than Arrow's in that it allows alternative assumptions on measurability and interpersonal comparability. But there is one assumption that both accounts share; namely the assumption that individual welfare can be expressed in terms of a single welfare ordering or a single scalar-valued welfare function. This assumption requires the existence of a single 'dimension' or a single 'currency' of welfare, in respect of which each individual's welfare over a set of alternatives can be assessed.

This paper provides a multidimensional generalization of Arrow's and Sen's frameworks. The welfare of each individual will be expressed in terms of *multiple* personal welfare functions, one for each relevant 'dimension'. Like a one-dimensional framework, the multidimensional framework raises the question of measurability and interpersonal comparability. Unlike a one-dimensional framework, it raises an additional question. Is it possible to compare an individual's welfare in one dimension with that same individual's welfare in another dimension? This, roughly, will be called the question of *interdimensional comparability*.[1]

I will first identify a multidimensional Arrow problem. In the absence of sufficient interpersonal comparability, any aggregation procedure satisfying Arrow-type conditions will make one individual dictatorial. And in the absence of sufficient interdimensional comparability, any such aggregation procedure will make one dimension dominant. An Arrow-type impossibility result can be avoided if sufficient comparability *both* across individuals and *and* across dimensions is admitted, and a rich set of possibilities emerges.

Although the literature contains some formally closely related results, multidimensionality is usually not developed as an interpretation. Such results are Roberts's (1995) results on the aggregation of multiple opinions about the welfare of a group of individuals into a single social ordering and Khmelnitskaya's (1999) and Khmelnitskaya and Weymark's (2000) results on social welfare orderings for different scales of individual utility measurement in distinct population subgroups. Amongst the first papers on multidimensional welfare are Plott, Little and Parks (1975) and Sen (1980/81), the former providing an Arrow-style theorem for aggregation across multiple dimensions within a single individual. Kelsey (1987) provides a relevant reinterpretation of the literature on informational assumptions in social choice. Less closely related precursors are results by Fishburn (1971) and Batra and Pattanaik (1972) on multi-stage majority decisions, involving nested aggregation over nested subprofiles of a given profile of preference orderings across individuals. Finally, a companion paper to this paper is concerned with multidimensional preference aggregation without any forms of comparability, but with intradimensional single-peakedness (List, 2002).

The paper is in five sections. In section 2, I will briefly discuss the political theory background of multidimensional aggregation problems. In section 3, I will address the formalization of measurability, interpersonal comparability and interdimensional comparability. In section 4, I will state the results, and, in section 5, I will draw some conclusions. The derivation of the main impossibility result will be stated in an appendix.

## 2. Multidimensionality and Social Choice

Sen has argued that the choice of a normatively relevant evaluation standard, or 'currency of welfare', may be as consequential as the choice of an aggregation procedure itself. The *locus classicus* is his famous paper "Equality of What?" (Sen, 1982). People may agree that equality of some form matters in the design of an aggregation procedure, and yet disagree on the question "Equality of What?". Income egalitarians, marginal-utility egalitarians, total-utility egalitarians, and Rawlsian primary goods egalitarians may all claim to use 'egalitarian' aggregation procedures. But, in light of their different views on what 'currency' should be equalized, they derive fundamentally different conclusions as to what social arrangements should be pursued.

Let me illustrate how the choice of a ‘currency of welfare’ affects the recommended social arrangements. Income egalitarians seek to equalize income across individuals. This may lead to an unequal distribution of utility, as the capacity to convert income into utility may differ across individuals. For instance, someone with a certain medical condition may need more income to sustain a particular utility level than someone without that medical condition. Marginal-utility egalitarians seek to equalize marginal utility across individuals, which typically amounts to maximizing the sum-total of utility across individuals. This will not in general lead to an equal distribution of income, nor to an equal distribution of utility. Those individuals who are more efficient in converting income into utility may receive more income under this proposal, and those who are less efficient may receive less. Further, a distribution that equalizes marginal utility across individuals (and thereby maximizes total utility) is not typically one in which utility itself is equally distributed. Total-utility egalitarians, by contrast, seek to equalize utility itself. This will not in general lead to a distribution which *maximizes* total utility, nor to an equal distribution of income. Total-utility egalitarians endorse income inequalities between individuals if these income inequalities lead to a more equal distribution of utility. Under total-utility egalitarianism, someone with a medical condition as in the example above should receive more income than someone without that medical condition. However, total-utility egalitarians face the challenge of expensive tastes. If a person has an expensive taste and requires caviar to achieve the same utility level that others achieve by consuming bread, then total-utility egalitarianism seems to entail that that person ought to receive additional resources in order to afford caviar. Many find this conclusion unacceptable and therefore seek to identify a 'currency' that is less 'subjectivist' than utility (so that it is immune to the problem of expensive tastes), but more welfare-relevant than income (so that it is sensitive to special needs, like medical ones). Whether a 'currency' with the desired properties exists is a matter of philosophical debate, but Rawlsian primary goods, as discussed below, are sometimes held to be a plausible such currency.

Sen uses the term *competitive plurality* to refer to this diversity in views on what the relevant 'currency of welfare' is: different proposals stand as rivals to each other. According to the classical assumption of one-dimensionality, one 'currency of welfare' is to be selected as the relevant one from amongst this competitive plurality.

Against this assumption, Sen (e.g. 1985, 1987), Walzer (1983) and others have argued that the relevant information for many problems of welfare evaluation is multidimensional. The effects of alternative social arrangements on an individual's income, nutrition and shelter, health, educational prospects, social status, and so on, may all be relevant. But it may be impossible to represent this information in terms of a *single* one-dimensional 'currency of welfare'. The use of *multiple* evaluation standards may therefore be warranted. Sen introduces the term *constitutive plurality* to refer to this view, that the relevant evaluation standard is *internally diverse* (e.g. 1987, pp. 2-3).

Although Sen and Walzer both defend constitutively plural conceptions of welfare, their accounts are quite different. Sen's account is motivated by the Aristotelian *essentialist* view that several *human functionings* form an essential part of a 'good life' and are thus relevant dimensions for welfare evaluation (see also Nussbaum, 1992). Such functionings may range from being well nourished and being free from avoidable disease to being able to participate in social life and having self-respect. Walzer's plurality of *spheres of goods*, by contrast, is motivated by a *communitarian* and *contextualist* theory of the good. The relevant spheres are determined by context-specific social meanings in the relevant society. Examples of such spheres might be health, education, employment, political influence, and money. Sen acknowledges the need for aggregation and cross-dimensional indexing (e.g. 1991; 1997, section A.7.3), whereas Walzer defends the mutual separateness, and arguably incomparability, of different dimensions. In particular, Walzer requires that different spheres of goods be kept separate and that different (dimension/goods-specific) principles of justice apply to different such spheres.

The evaluation standard of Rawls's *Theory of Justice* (1971), an index of *primary goods*, is also constitutively plural, including rights, liberties and opportunities, income and wealth, and the social bases of self-respect (Rawls, 1971, pp. 60 - 65). However, Rawls makes the tacit assumption that summarizing a plurality of primary goods into a single index (which is ordinally measurable and interpersonally comparable) is possible, but is silent on how to construct such an index. Formally, Rawls's social-choice-theoretic proposals, particularly the *difference principle*, are presented in a one-dimensional form. A number of social choice theorists have addressed the problem of indexing primary goods and derived some Arrow-type impossibility results, all based on the assumption that the comparability between different primary goods is limited (for instance, Plott, 1978; Gibbard, 1979; Blair, 1988; but see Sen, 1991).

At a practical level, the Human Development Index (HDI), as employed by the United Nations Development Programme, is also a constitutively plural 'currency of welfare', albeit an aggregate one. The Human Development Index combines three components, each of which is itself an aggregate measure across people in a country or region. The three components are longevity, knowledge, and standard of living. Longevity is measured in terms of life expectancy at birth. Knowledge is measured in terms of a combination of the adult literacy rate and enrolment in primary, secondary, and tertiary education. Standard of living is measured in terms of GDP per capita (UNDP, 2002).

The present approach allows to formalize a constitutively plural conception of welfare at a micro-level: it allows the separate representation of the effects of each alternative on each individual in each one of multiple relevant dimensions.

## 3. Measurability, Interpersonal Comparability and Interdimensional Comparability

Let $N = \{1, 2, ..., n\}$ be a set of individuals, and $X = \{x_1, x_2, ...\}$ a set of alternatives. Suppose, further, that there are $k$ relevant dimensions, contained in $K = \{1, 2, ..., k\}$. We will assume $n > 1$, $|X| > 2$, and, unless stated otherwise, $k > 1$.[2]

To each individual $i \in N$, there corresponds a $k$-tuple $W_i := <W_{ij}>_{j \in K} = <W_{i1}, W_{i2}, ..., W_{ik}>$ of personal welfare functions, containing one $W_{ij} : X \rightarrow \mathbf{R}$ for each dimension $j \in K$. For each individual $i \in N$ and each dimension $j \in K$, the function $W_{ij}$ assigns to each $x \in X$ a real number $W_{ij}(x)$. Here $W_{ij}(x)$ represents the welfare of individual $i$ in dimension $j$ under alternative $x$.

For example, let $X$ be a set of employment policy alternatives, and let $1, 2, 3 \in K$ be the dimensions of income, health and social status. Then $W_{i1}$, $W_{i2}$ and $W_{i3}$ represent individual $i$'s welfare over the alternatives in $X$ from the perspectives of income, health and social status.

A *profile of k-tuples of personal welfare functions* is an $n$-tuple $\{W_i\}_{i \in N} = \{W_1, W_2, ..., W_n\}$ of such $k$-tuples, containing one $k$-tuple $W_i$ for each individual $i \in N$. As a notational convention, we use {}-brackets to denote $n$-tuples across individuals and <>-brackets to denote $k$-tuples across dimensions.

A *multidimensional social welfare functional* (*MSWFL*) is a function $F$ which maps each profile of $k$-tuples of personal welfare functions (in a given domain), $\{W_i\}_{i \in N}$, to a corresponding social ordering $R = F(\{W_i\}_{i \in N})$ on $X$, where $R$ is reflexive, transitive and connected. $R$ induces a strong ordering $P$ and an indifference relation $I$, defined as follows: for all $x_1, x_2 \in X$,

$x_1 P x_2$ if and only if $x_1 R x_2$ and not $x_2 R x_1$;
$x_1 I x_2$ if and only if $x_1 R x_2$ and $x_2 R x_1$.

A one-dimensional *social welfare functional* (*SWFL*) is simply an MSWFL for the special case $k$=1.

How can we formalize assumptions on measurability, interpersonal comparability and interdimensional comparability of welfare? The formalization to be developed is directly analogous to Sen's well known method of formalizing assumptions on measurability and interpersonal comparability in a one-dimensional social choice framework (e.g. Sen, 1970/1979; Sen, 1982, ch. 11; see also List, 2001). We first explain the idea behind the formalization and then proceed to stating the formalization properly. Let us consider statements of the following forms:

**Level Comparisons (LC).** Alternative $x_1$ from the perspective of individual $i_1$ in dimension $j_1$ is at least as good as alternative $x_2$ from the perspective of individual $i_2$ in dimension $j_2$; formally $W_{i_1 j_1}(x_1) \geq W_{i_2 j_2}(x_2)$.

**Unit Comparisons (UC).** The ratio of [individual $i_1$'s gain/loss in dimension $j_1$ as a result of switching from alternative $x_1$ to $x_2$] to [individual $i_2$'s gain/loss in dimension $j_2$ as a result of switching from alternative $y_1$ to $y_2$] is $\lambda$, where $\lambda$ is a real number; formally $(W_{i_1 j_1}(x_2) - W_{i_1 j_1}(x_1))/(W_{i_2 j_2}(y_2) - W_{i_2 j_2}(y_1)) = \lambda$.

The key idea is that different assumptions on measurability, interpersonal comparability and interdimensional comparability of welfare imply different conditions under which (LC)- and (UC)-statements are meaningful. Specifically, we have the following:

- *Ordinal measurability* (O) implies that (LC)-statements are meaningful when $i_1=i_2$ and $j_1=j_2$.
- *Cardinal measurability* (C) implies that (LC)-statements and (UC)-statements are meaningful when $i_1=i_2$ and $j_1=j_2$.
- *Interpersonal level comparability* ($L_{per}$) implies that (LC)-statements are meaningful when $j_1=j_2$ but $i_1 \neq i_2$.
- *Interpersonal unit comparability* ($U_{per}$) implies that (UC)-statements are meaningful when $j_1=j_2$ but $i_1 \neq i_2$.
- *Interdimensional level comparability* ($L_{dim}$) implies that (LC)-statements are meaningful when $i_1=i_2$ but $j_1 \neq j_2$.
- *Interdimensional unit comparability* ($U_{dim}$) implies that (UC)-statements are meaningful when $i_1=i_2$ but $j_1 \neq j_2$.

We use the labels ($N_{per}$) and ($N_{dim}$) to refer to, respectively, no interpersonal comparability and no interdimensional comparability.

Once we have assigned a profile of *k*-tuples of personal welfare functions $\{W_i\}_{i \in N}$ to a set of individuals, we can of course make (LC)- and (UC)-statements *relative to that profile*. However, whether such statements are *meaningful* depends on *how unique* the profile $\{W_i\}_{i \in N}$ is. Suppose, for instance, that each $W_{ij}$ in the profile $\{W_i\}_{i \in N}$ is unique only up to a positive monotonic transformation, possibly a different transformation for different individuals and dimensions. Then interpersonal or interdimensional level or unit comparisons are not well-defined and thus not meaningful, as these comparisons are not in general invariant under the specified transformations.

Suppose, more generally, we specify the class of transformations $\Phi$ up to which each profile $\{W_i\}_{i \in N}$ is unique. We can then ask whether each kind of (LC)- or (UC)-statements is invariant under all the transformations in $\Phi$. A particular kind of (LC)- or (UC)-statement is said to be meaningful if and only if it is invariant under all transformations in $\Phi$.

We state assumptions on measurability, interpersonal comparability and interdimensional comparability by specifying the class of transformations $\Phi$ up to which each profile $\{W_i\}_{i \in N}$ is taken to be unique. The smaller this class of transformations, the more information is contained in a profile. Table 1 lists several alternative such classes of transformations.[3]

| The class of transformations $\Phi$ is the class of all $n$-tuples of $k$-tuples of transformations, $\{<\phi_{ij}>_{j\in K}\}_{i\in N}$, with the following properties: | | |
|---|---|---|
| | **O**<br>**Ordinal Measurability**<br>Each $\phi_{ij} : \mathbf{R} \rightarrow \mathbf{R}$ is a positive monotonic transformation, and ... | **C**<br>**Cardinal Measurability**<br>Each $\phi_{ij} : \mathbf{R} \rightarrow \mathbf{R}$ is a positive affine transformation, and ... |
| $\mathbf{N_{dim}N_{per}}$ | no further assumption required | no further assumption required |
| $\mathbf{U_{dim}N_{per}}$ | not applicable | for each $i\in N$, $\exists a_{i1}, a_{i2}, ..., a_{ik}\in \mathbf{R}$ such that $\phi_{i1}+a_{i1} = \phi_{i2}+a_{i2} = ... = \phi_{ik}+a_{ik}$ |
| $\mathbf{L_{dim}N_{per}}$ | for each $i\in N$, $\phi_{i1} = \phi_{i2} = ... = \phi_{ik}$ | not applicable |
| $\mathbf{L_{dim}U_{dim}N_{per}}$ | not applicable | for each $i\in N$, $\phi_{i1} = \phi_{i2} = ... = \phi_{ik}$ |
| $\mathbf{N_{dim}U_{per}}$ | not applicable | for each $j\in K$, $\exists a_{1j}, a_{2j}, ..., a_{nj}\in \mathbf{R}$ such that $\phi_{1j}+a_{1j} = \phi_{2j}+a_{2j} = ... = \phi_{nj}+a_{nj}$ |
| $\mathbf{N_{dim}L_{per}}$ | for each $j\in K$, $\phi_{1j} = \phi_{2j} = ... = \phi_{nj}$ | not applicable |
| $\mathbf{N_{dim}L_{per}U_{per}}$ | not applicable | for each $j\in K$, $\phi_{1j} = \phi_{2j} = ... = \phi_{nj}$ |
| $\mathbf{L_{dim}L_{per}}$ | all $\phi_{ij}$ are identical | not applicable |
| $\mathbf{U_{dim}U_{per}}$ | not applicable | $\exists a_{11}, ..., a_{1k}, a_{21}, ..., a_{2k}, a_{n1}, ..., a_{nk}\in \mathbf{R}$ such that $\phi_{11}+a_{11} =...= \phi_{1k}+a_{1k} = \phi_{21}+a_{21} =...= \phi_{2k}+a_{2k} = \phi_{n1}+a_{n1} =...= \phi_{nk}+a_{nk}$ |
| $\mathbf{U_{dim}L_{per}U_{per}}$ | not applicable | for each $j\in K$, $\phi_{1j} = \phi_{2j} = ... = \phi_{nj}$<br>and $\exists a_1, a_2, ..., a_k\in \mathbf{R}$ such that,<br>for each $i\in N$, $\phi_{i1}+a_1 = \phi_{i2}+a_2 = ... = \phi_{ik}+a_k$ |
| $\mathbf{L_{dim}U_{dim}U_{per}}$ | not applicable | for each $i\in N$, $\phi_{i1} = \phi_{i2} = ... = \phi_{ik}$<br>and $\exists a_1, a_2, ..., a_n\in \mathbf{R}$ such that,<br>for each $j\in K$, $\phi_{1j}+a_1 = \phi_{2j}+a_2 = ... = \phi_{nj}+a_n$ |
| $\mathbf{L_{dim}U_{dim}L_{per}U_{per}}$ | not applicable | all $\phi_{ij}$ are identical |

**Table 1**

An example shows how to read table 1. Consider $OL_{dim}N_{per}$, the assumption corresponding to the row labelled $L_{dim}N_{per}$ and the column labelled O. Then $OL_{dim}N_{per}$ is the assumption of ordinal measurability, interdimensional level comparability, but no interpersonal comparability. According to $OL_{dim}N_{per}$, $\Phi$ is the class of all $n$-tuples of $k$-tuples of transformations, $\{<\phi_{ij}>_{j\in K}\}_{i\in N}$ such that each $\phi_{ij} : \mathbf{R} \rightarrow \mathbf{R}$ is a positive monotonic transformation, and, for each $i\in N$, $\phi_{i1} = \phi_{i2} = ... = \phi_{ik}$.

Now suppose that each $\{W_i\}_{i\in N}$ is unique up to the transformations in $\Phi$. We then require that a MSWFL be invariant under these transformations. The idea behind this requirement is this. Suppose $\{W_i\}_{i\in N}$ can be transformed into $\{W^*_i\}_{i\in N}$ by some transformation in $\Phi$. Then $\{W_i\}_{i\in N}$ and $\{W^*_i\}_{i\in N}$ are taken to contain exactly the same relevant information. Therefore our MSWFL should map $\{W_i\}_{i\in N}$ and $\{W^*_i\}_{i\in N}$ to the same social ordering. Formally, we can state this requirement as follows.

**INVARIANCE ASSUMPTION WITH RESPECT TO $\Phi$.** For any $\{W_i\}_{i\in N}$ and $\{W^*_i\}_{i\in N}$ in the domain of $F$, if there exists $\{<\phi_{ij}>_{j\in K}\}_{i\in N} \in \Phi$ such that, for each $i\in N$ and each $j\in K$, $W^*_{ij} = \phi_{ij}(W_{ij})$, then $F(\{W_i\}_{i\in N}) = F(\{W^*_i\}_{i\in N})$.

Thus each assumption in table 1 defines a specific class of transformations $\Phi$, and we can consider the corresponding invariance assumption with respect to $\Phi$. We use the name of each assumption in round brackets, for example ($OL_{dim}N_{per}$), to denote the corresponding invariance assumption.[4]

Let me make one final remark. Suppose $\Phi$ and $\Psi$ are two classes of transformations such that $\Psi\subseteq\Phi$: for example, $\Phi$ corresponds to $OL_{dim}N_{per}$, and $\Psi$ corresponds to $CL_{dim}U_{dim}N_{per}$. Then any MSWFL satisfying the invariance assumption with respect to $\Phi$ will also satisfy the invariance assumption with respect to $\Psi$.

## 4. Results

Before we can present impossibility and possibility results, we need to state multidimensional generalizations of Arrow's conditions. The generalizaton of universal domain, the weak Pareto principle and independence of irrelevant alternatives is straightforward. Given any profile of $k$-tuples of personal welfare functions, $\{W_i\}_{n\in N}$, we define $R := F(\{W_i\}_{n\in N})$.

**UNIVERSAL DOMAIN (U).** The domain of $F$ is the set of all logically possible profiles of $k$-tuples of personal welfare functions.

**WEAK PARETO PRINCIPLE (P).** Let $\{W_i\}_{i\in N}$ be any profile in the domain of $F$. For any $x_1, x_2\in X$, we have $x_1Px_2$ whenever, for all $i\in N$ and all $j\in K$, $W_{ij}(x_1)>W_{ij}(x_2)$.

**INDEPENDENCE OF IRRELEVANT ALTERNATIVES (I).** Let $\{W_i\}_{i\in N}$ and $\{W^*_i\}_{i\in N}$ be any profiles in the domain of $F$. Suppose $x_1, x_2\in X$ such that, for all $i\in N$ and all $j\in K$, $W_{ij}(x_1)=W^*_{ij}(x_1)$ and $W_{ij}(x_2)=W^*_{ij}(x_2)$. Then $x_1Rx_2$ if and only if $x_1R^*x_2$.

The generalization of non-dictatorship is less straightforward (see also List, 2002). In the multidimensional framework, non-dictatorship corresponds to two conditions: non-dictatorship and non-dominance. Non-dictatorship requires the non-existence of a fixed single individual whose $k$-tuple of personal welfare functions always determines the social ordering. Non-dominance requires the non-existence of a fixed single dimension such that the personal welfare functions (across individuals) *in that dimension* always determine the social ordering.

**NON-DICTATORSHIP (D).** There does not exist an individual $i \in N$ (a *dictator* for $F$) such that, for all profiles $\{W_i\}_{i \in N}$ in the domain of $F$ and all $x_1, x_2 \in X$, $f(<W_{ij}(x_1)>_{j \in K}) > f(<W_{ij}(x_2)>_{j \in K})$ implies $x_1 P x_2$, where $f: \mathbf{R}^k \to \mathbf{R}$ is a strictly increasing function.

**NON-DOMINANCE (DOM).** There does not exist a dimension $j \in K$ (a *dominant dimension* for $F$) such that, for all profiles $\{W_i\}_{i \in N}$ in the domain of $F$ and all $x_1, x_2 \in X$, $f(\{W_{ij}(x_1)\}_{i \in N}) > f(\{W_{ij}(x_2)\}_{i \in N})$ implies $x_1 P x_2$, where $f: \mathbf{R}^n \to \mathbf{R}$ is a strictly increasing function.

Non-double-dictatorship, finally, requires the non-existence of a fixed single individual and a fixed single dimension such that this individual's personal welfare function in that dimension always determines the social ordering.

**NON-DOUBLE-DICTATORSHIP (DD).** There does not exist an individual $i \in N$ and a dimension $j \in K$ such that, for all profiles $\{W_i\}_{i \in N}$ in the domain of $F$ and all $x_1, x_2 \in X$, $W_{ij}(x_1) > W_{ij}(x_2)$ implies $x_1 P x_2$.

Under all these generalized non-dictatorship conditions, personal welfare functions other than the dictatorial or dominant ones can act at most as tie-breakers, namely in those cases in which the dictatorial or dominant functions reflect indifference.

In the absence of both interpersonal comparability and interdimensional comparability, Arrow's theorem (in the version of Sen, 1970/1979; see also List, 2002) implies the following result:

**Theorem 1.** There exists no MSWFL satisfying ($ON_{dim}N_{per}$) or ($CN_{dim}N_{per}$), and (U), (P), (I) and (DD).

A multidimensional aggregation problem with $n$ individuals and $k$ dimensions, with ($ON_{dim}N_{per}$) or ($CN_{dim}N_{per}$), is equivalent to a one-dimensional aggregation problem with $nk$ individuals and without interpersonal comparability. Theorem 1 thus follows directly from Arrow's theorem.

### 4.1. *Lexicographic Dictatorships and Lexicographic Hierarchies of Dimensions*

Is interpersonal comparability, in analogy to the one-dimensional case, sufficient for avoiding the multidimensional Arrow problem? The answer to this question is negative. If we assume ordinal or even cardinal measurability, with interpersonal comparability of both levels and units, but without interdimensional comparability, any MSWFL which satisfies (U), (P) and (I) still violates (DOM).

**Theorem 2.** There exists no MSWFL satisfying ($ON_{dim}L_{per}$) or ($CN_{dim}L_{per}U_{per}$), and (U), (P), (I) and (DOM).

The result can be seen as a reinterpretation of results by Roberts (1995, theorem 6) and Khmelnitskaya and Weymark (2000, theorem 2). See appendix 1.

Similarly, if we assume ordinal or even cardinal measurability, with interdimensional comparability of both levels and units, but without interpersonal comparability, any MSWFL satisfying (U), (P) and (I) still violates (D).

**Theorem 3.** There exists no MSWFL satisfying ($OL_{dim}N_{per}$) or ($CL_{dim}U_{dim}N_{per}$), and (U), (P), (I) and (D).

See also appendix 1. Given comparability across individuals but not across dimensions (i.e. ($ON_{dim}L_{per}$) or ($CN_{dim}U_{per}$)), there are MSWFLs satisfying (U), (P), (I) and (D), but not (DOM). Examples are suitable *lexicographic hierarchies of dimensions*, as defined below. Given comparability across dimensions but not across individuals (i.e. ($OL_{dim}N_{per}$) or ($CU_{dim}N_{per}$)), there are MSWFLs satisfying (U), (P), (I) and (DOM), but not (D). Examples are suitable *lexicographic dictatorships*, also defined below.

A MSWFL $F$ is a *lexicographic hierarchy of dimensions* if there exist $k$ strictly increasing functions $f_1, f_2, ..., f_k$: $\mathbf{R}^n \to \mathbf{R}$ (possibly different), one for each dimension in $K$, and a permutation $\sigma$ of $K$ such that, for any $\{W_i\}_{i\in N}$ and any $x_1, x_2 \in X$,

$x_1 P x_2$ if and only if

$$f_{\sigma(j)}(\{W_{i\sigma(j)}(x_1)\}_{i\in N}) > f_{\sigma(j)}(\{W_{i\sigma(j)}(x_2)\}_{i\in N}) \quad \text{for some } j\in K$$

and $$f_{\sigma(h)}(\{W_{i\sigma(h)}(x_1)\}_{i\in N}) = f_{\sigma(h)}(\{W_{i\sigma(h)}(x_2)\}_{i\in N}) \quad \text{for all } h<j.$$

Under this definition, the dimensions are ranked in a fixed hierarchy of importance. For each dimension $j$, there exists a function $f_j$ (to be called an *intradimensional aggregation function*) for aggregating the dimension-$j$-specific welfare information for each alternative $x$ across individuals into a single aggregate figure for that dimension. The alternatives in $X$ are then ranked, lexically, according to the aggregate figures for the dimensions which are first, second, third, and so on in the hierarchy. As noted above, a lexicographic hierarchy of dimensions need not violate (D). Given sufficient interpersonal comparability, any non-dictatorial one-dimensional SWFL can be the basis for the intradimensional aggregation functions $f_1, f_2, ..., f_k$. Prominent examples are:

- if (at least) ($CN_{dim}U_{per}$) is satisfied:
  a *'utilitarian' intradimensional aggregation function*:
  $f_j(\{W_{ij}(x)\}_{i\in N}) := \lambda_1 W_{1j}(x) + \lambda_2 W_{2j}(x) + ... + \lambda_n W_{nj}(x)$,
  where $\lambda_1, \lambda_2, ..., \lambda_n \geq 0$;
- if (at least) ($ON_{dim}L_{per}$) is satisfied:
  a *'maximin' intradimensional aggregation function*:
  $f_j(\{W_{ij}(x)\}_{i\in N}) := min(W_{1j}(x), W_{2j}(x), ... ,W_{nj}(x))$.

In a lexicographic hierarchy of dimensions, different intradimensional aggregation functions can be chosen for different dimensions. Given interpersonal comparability within each dimension, but no interdimensional comparability, a lexicographic hierarchy of dimensions may in some cases be an attractive MSWFL. In *Theory of Justice*, Rawls contrasts the 'special conception' with the 'general conception' of justice. The ‘special conception’ is intended to apply to socio-economically well developed societies, whereas the ‘general conception’ is intended to apply to socio-economically less developed societies. The 'general conception' requires that "the difference principle [be] applied to *all* primary goods including liberty and opportunity" (Rawls, 1971, p. 83; my italics). This presumably involves aggregation, in accordance with (DOM), across *all* primary goods. According to theorem 2, if the different dimensions are constituted by different primary goods and if these different primary goods are mutually incommensurable (as captured by ($ON_{dim}L_{per}$) or ($CN_{dim}L_{per}U_{per}$)), then Rawls’s ‘general conception’ of justice leads to an impossibility result. The 'special conception' of justice, on the other hand, assigns lexical priority to some primary goods (e.g. liberty and opportunity) over others. A lexicographic hierarchy of dimensions captures precisely this idea.

A MSWFL $F$ is a *lexicographic dictatorship* if there exist $n$ strictly increasing functions $f_1, f_2, ..., f_n$: $\mathbf{R}^k \rightarrow \mathbf{R}$ (possibly different), one for each individual in $N$, and a permutation $\sigma$ of $N$ such that, for any $\{W_i\}_{i\in N}$ and any $x_1, x_2 \in X$,

$x_1 P x_2$ if and only if

$$f_{\sigma(i)}(<W_{\sigma(i)j}(x_1)>_{j\in K}) > f_{\sigma(i)}(<W_{\sigma(i)j}(x_2)>_{j\in K}) \quad \text{for some } i\in N$$

$$\text{and} \quad f_{\sigma(h)}(<W_{\sigma(h)j}(x_1)>_{j\in K}) = f_{\sigma(h)}(<W_{\sigma(h)j}(x_2)>_{j\in K}) \quad \text{for all } h<i.$$

In a lexicographic dictatorship, the individuals are arranged in a fixed hierarchy of decisiveness. For each individual $i$, there exists a function $f_i$ (to be called an *intrapersonal aggregation function*) for aggregating this individual's $k$-tuple of welfare information for each alternative $x$ into a single aggregate figure for that individual. The alternatives in $X$ are then ranked, lexically, according to the aggregate figures for the individuals who are first, second, third, and so on in the hierarchy. A lexicographic dictatorship need not violate (DOM). Given sufficient interdimensional comparability, intrapersonal aggregation functions that focus on more than one dimension are available, e.g.

- if (at least) ($CU_{dim}N_{per}$) is satisfied:
  a *'utilitarian' intrapersonal aggregation function*:
  $f_i(<W_{ij}(x)>_{j\in K}) := \lambda_1 W_{i1}(x) + \lambda_2 W_{i2}(x) + ... + \lambda_k W_{ik}(x)$,
  where $\lambda_1, \lambda_2, ..., \lambda_k \geq 0$;
- if (at least) ($OL_{dim}N_{per}$) is satisfied:
  *a 'maximin' intrapersonal aggregation function*:
  $f_i(<W_{ij}(x)>_{j\in K}) := min(W_{i1}(x), W_{i2}(x),... ,W_{ik}(x))$.

In a lexicographic dictatorship, different intrapersonal aggregation functions can be chosen for different individuals. However, unless there are particular reasons why certain individuals should be given lexical priority over others, lexicographic dictatorships are hardly attractive from a liberal-egalitarian viewpoint.

4.2. *Stronger Possibility Results*

Comparability *both* across individuals *and* across dimensions is required if we regard *both* (D) *and* (DOM) and the other Arrow-type conditions as indispensable. To state stronger possibility results, we must distinguish between different ways of aggregating across individuals and dimensions. Aggregation may take place *either* in two steps *or* in one step. Two-step aggregation can mean one of two things: (1) for each individual, the $k$-tuple of welfare information for each alternative is first aggregated across dimensions into an aggregate figure for that individual, and these aggregate figures are then aggregated across individuals; (2) for each separate dimension, the dimension-specific welfare information for each alternative is first aggregated across individuals into an aggregate figure for that dimension, and these aggregate figures are then aggregated across dimensions. One-step aggregation means that the entire profile of $k$-dimensional personal welfare functions is aggregated directly into an overall social ordering.[5]

*Multidimensional Utilitarian Rules*

**Proposition 4.** There exist MSWFLs satisfying ($CU_{dim}U_{per}$), (U), (P), (I), (D) and (DOM).

Given interdimensional comparability and interpersonal comparability of *units*, suitable multidimensional utilitarian rules satisfy (U), (P), (I), (D) and (DOM). All multidimensional utilitarian rules are expressible as two-step aggregation mechanisms. We will distinguish between those that prioritise intrapersonal aggregation and those that prioritise intradimensional aggregation.

A MSWFL $F$ is a *multidimensional utilitarian rule that prioritises intrapersonal aggregation* if there exist

(i) $n$ strictly increasing intrapersonal aggregation functions $f_1, f_2, ..., f_n$: $\mathbf{R}^k \to \mathbf{R}$;

(ii) $\alpha_1, \alpha_2, ..., \alpha_n \geq 0$ (with $\sum \alpha_{i \in N} = 1$) (weights corresponding to the $n$ individuals)

such that, for any $\{W_i\}_{i \in N}$ and any $x_1, x_2 \in X$,

$$x_1 R x_2 \text{ if and only if } \sum_{i \in N} \alpha_i f_i(<W_{ij}(x_1)>_{j \in K}) \geq \sum_{i \in N} \alpha_i f_i(<W_{ij}(x_2)>_{j \in K}).$$

Possible choices for $f_1, f_2, ..., f_n$ are, for example, 'utilitarian' intrapersonal aggregation functions or 'maximin' intrapersonal aggregation functions (requiring at least ($CL_{dim}U_{dim}U_{per}$)) as defined above.

A SWFL $F$ is a *multidimensional utilitarian rule that prioritises intradimensional aggregation* if there exist

(i) $k$ strictly increasing intradimensional aggregation functions $f_1, f_2, ..., f_k$: $\mathbf{R}^n \to \mathbf{R}$;

(ii) $\beta_1, \beta_2, ..., \beta_k \geq 0$ (with $\sum \beta_{j \in K} = 1$) (weights corresponding to the $k$ dimensions)

such that, for any $\{W_i\}_{i \in N}$ and any $x_1, x_2 \in X$,

$$x_1 R x_2 \text{ if and only if } \sum_{j \in K} \beta_j f_j(\{W_{ij}(x_1)\}_{i \in N}) \geq \sum_{j \in K} \beta_j f_j(\{W_{ij}(x_2)\}_{i \in N}).$$

Possible choices for $f_1, f_2, ..., f_k$ are, for example, 'utilitarian' intradimensional aggregation functions or 'maximin' intradimensional aggregation functions (requiring at least ($CU_{dim}L_{per}U_{per}$)) as defined above.

Multidimensional utilitarian rules are versions of the classical utilitarian principle: make social choices so as to maximize the sum-total of welfare. In the case of rules that prioritise intrapersonal aggregation, an aggregate figure is first determined for each individual (by intrapersonal aggregation), and these aggregate figures are then added up across individuals. In the case of rules that prioritise intradimensional aggregation, an aggregate figure is first determined for each dimension (by intradimensional aggregation), and these aggregate figures are then added up across dimensions.

The concept of two-step aggregation combined with different types of intrapersonal or intradimensional aggregation functions (e.g. utilitarian ones as well as maximin ones) greatly extends the variety of possible utilitarian rules as compared with the one-dimensional case. Suppose, for example, we have interdimensional comparability not only of units but also of levels. We might then use maximin intrapersonal aggregation functions for each individual as the basis for utilitarian welfare summation across individuals. For each individual and each alternative, we find the dimension in which the individual is worst off and define the individual's aggregate figure to be his or her welfare in that dimension. The sum-total of welfare is then determined by summation of these aggregate figures across individuals. To consider a different example, suppose we have interpersonal comparability not only of units but also of levels. We might then use maximin intradimensional aggregation functions as the basis for utilitarian welfare summation across dimensions: for each dimension and each alternative, we identify the individual who is worst off in that dimension and define the dimension's aggregate figure to be the welfare of the worst-off individual in that dimension. The sum-total of welfare is then determined by summation of these aggregate figures across dimensions. Other combinations are possible.

The first of the two rules captures the view that the overall welfare of an individual should be identified with the lowest component of this individual's welfare *k*-tuple, but that subsequent aggregation across individuals should take the form of utilitarian summation. The second rule captures the view that – in a Rawlsian spirit – the overall 'social' welfare in each dimension should be identified with the welfare of the worst-off individual in that dimension, but that – diverging from Rawls – subsequent aggregation across dimensions should take the form of utilitarian summation.

*Multidimensional Leximin Rules*

**Proposition 5.** There exist MSWFLs satisfying ($OL_{dim}L_{per}$), (U), (P), (I), (D) and (DOM).

Given interdimensional comparability and interpersonal comparability of *levels*, suitable multidimensional positional rules, in particular leximin rules, satisfy (U), (P), (I), (D) and (DOM).[6] We will distinguish between three types of multidimensional leximin rules; two-step rules that prioritise intrapersonal aggregation, two-step rules that prioritise intradimensional aggregation, and one-step rules.

A MSWFL $F$ is a *multidimensional leximin rule that prioritises intrapersonal aggregation* if there exist

(i) $n$ strictly increasing intrapersonal aggregation functions $f_1, f_2, ..., f_n$: $R^k \to R$; and

(ii) for each $x \in X$, a permutation $i \mapsto [i]$ of $N$ (depending on $x$) such that

$f_{[1]}(<W_{[1]j}(x)>_{j\in K}) \leq f_{[2]}(<W_{[2]j}(x)>_{j\in K}) \leq ... \leq f_{[n]}(<W_{[n]j}(x)>_{j\in K})$;[7]

such that, for any $\{W_i\}_{i\in N}$ and any $x_1, x_2 \in X$,

$x_1 P x_2$ if and only if

$$f_{[i]}(<W_{[i]j}(x_1)>_{j\in K}) > f_{[i]}(<W_{[i]j}(x_2)>_{j\in K}) \quad \text{for some } i\in N,$$

$$\text{and} \quad f_{[h]}(<W_{[h]j}(x_1)>_{j\in K}) = f_{[h]}(<W_{[h]j}(x_2)>_{j\in K}) \quad \text{for all } h<i.$$

Possible choices for $f_1, f_2, ..., f_n$ are, for example, 'maximin' intrapersonal aggregation functions or 'utilitarian' intrapersonal aggregation functions (requiring at least ($CL_{dim}U_{dim}L_{per}U_{per}$)) as defined above.

A MSWFL $F$ is a *multidimensional leximin rule that prioritises intradimensional aggregation* if there exist

(i) $k$ strictly increasing intradimensional aggregation functions $f_1, f_2, ..., f_k$: $R^n \to R$; and

(ii) for each $x \in X$, a permutation $j \mapsto [j]$ of $K$ (depending on $x$) such that

$f_{[1]}(\{W_{i[1]}(x)\}_{i\in N}) \leq f_{[2]}(\{W_{i[2]}(x)\}_{i\in N}) \leq ... \leq f_{[k]}(\{W_{i[k]}(x)\}_{i\in N})$;

such that, for any $\{W_i\}_{i\in N}$ and any $x_1, x_2 \in X$,

$x_1 P x_2$ if and only if

$$f_{[j]}(\{W_{i[j]}(x_1)\}_{i\in N}) > f_{[j]}(\{W_{i[j]}(x_2)\}_{i\in N}) \quad \text{for some } j\in K$$

$$\text{and} \quad f_{[h]}(\{W_{i[h]}(x_1)\}_{i\in N}) = f_{[h]}(\{W_{i[h]}(x_2)\}_{i\in N}) \quad \text{for all } h<j.$$

Possible choices for $f_1, f_2, ..., f_k$ are, for example, 'maximin' intradimensional aggregation functions or 'utilitarian' intradimensional aggregation functions (requiring at least ($CL_{dim}U_{dim}L_{per}U_{per}$)) as defined above.

A MSWFL $F$ is a *(one-step) multidimensional leximin rule* if there exists

(i) for each $x \in X$, a bijection $h \mapsto [h] = (i, j)$ from $\{1, 2, ..., nk\}$ to $N \times K$ (depending on $x$) such that $W_{[1]}(x) \leq W_{[2]}(x) \leq ... \leq W_{[nk]}$

such that, for any $\{W_i\}_{i\in N}$ and any $x_1, x_2 \in X$,

$x_1 P x_2$ if and only if

$$W_{[h]}(x_1) > W_{[h]}(x_2) \quad \text{for some } h \in \{1, 2, ..., nk\}$$

$$\text{and} \quad W_{[g]}(x_1) = W_{[g]}(x_2) \quad \text{for all } g<h.$$

Multidimensional leximin rules are versions of Rawls's (lexicographic) difference principle: make social choices so as to maximize the lowest welfare levels; if there are ties, maximize, in a lexicographic hierarchy, the second lowest, third lowest, ..., welfare levels. These welfare levels are overall personal welfare levels in the case of rules that prioritise intrapersonal aggregation and overall dimensional welfare levels in the case of rules that prioritise intradimensional aggregation.

Here the concepts of one-step and two-step aggregation combined with different types of intrapersonal or intradimensional aggregation functions (e.g. utilitarian ones as well as maximin ones) also extend the variety of possible leximin rules as compared with the one-dimensional case. Suppose, for example, we have interdimensional comparability of levels and units. We might then use utilitarian intrapersonal aggregation functions for each individual as the basis for leximin aggregation across individuals. For each individual and each alternative, we take the (possibly weighted) sum of this individual's welfare levels in all dimensions to be the aggregate figure for that individual. The leximin rule is then applied to these aggregate figures across individuals: make social choices so as to maximize the lowest such aggregate figures across individuals; if there are ties, maximize, in a lexicographic hierarchy, the second lowest, third lowest, ..., such aggregate figures. To consider a different example, suppose we have interpersonal comparability of levels and units. We might then use utilitarian intradimensional aggregation functions for each dimension as the basis for leximin aggregation across dimensions. For each dimension and each alternative, we take the (possibly weighted) sum of dimension-specific welfare levels across individuals to be the aggregate figure for that dimension. The leximin rule is now applied to these aggregate figures across dimensions: make social choices so as to maximize the lowest such aggregate figures across dimensions; if there are ties, maximize, in a lexicographic hierarchy, the second lowest, third lowest, …, such aggregate figures.

The first of the two rules captures the view that the welfare levels of an individual should be identified with the (possibly weighted) sum of the components of this individual's welfare $k$-tuple, but that subsequent aggregation across individuals should focus, lexically, on the lowest, second lowest, third lowest, …, such individual welfare levels. This is one version of the Rawlsian difference principle, where the primary goods indexing problem for each individual is solved by summation across different primary goods. The second rule captures the view that for each dimension the overall welfare in that dimension should be identified with the (possibly weighted) sum of dimension-specific welfare across individuals, but that subsequent aggregation across dimensions should focus, lexically, on the lowest, second lowest, third lowest, …, such dimensional welfare levels.

A one-step multidimensional leximin rule, finally, treats the multidimensional aggregation problem as if a one-dimensional leximin rule were applied to a society of $nk$ individuals, effectively considering each of the *nk* individual-dimension pairs as a distinct individual and focussing, lexically, on the lowest, second lowest, third lowest, …, welfare levels amongst these *nk* individual-dimension pairs.

## 5. Some Implications

What conclusions can be drawn from this? Table 2 summarizes the basic results of this paper:

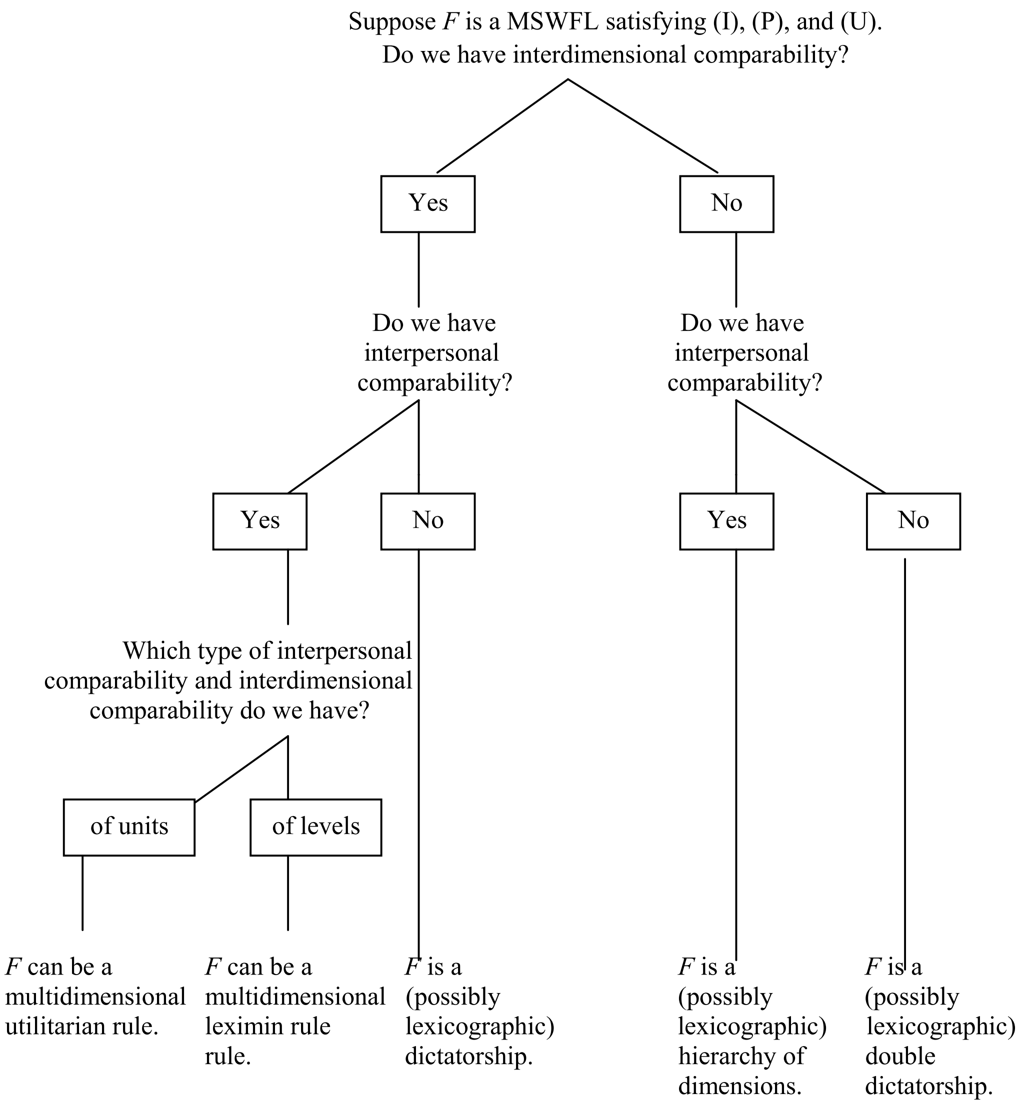


**Table 2**

In short, if we assume comparability across individuals, but not across dimensions, the only aggregation procedures satisfying Arrow-type conditions are (possibly lexicographic) hierarchies of dimensions. And if we assume comparability across dimensions, but not across individuals, the only aggregation procedures satisfying Arrow-type conditions are (possibly lexicographic) dictatorships of individuals. If we assume comparability across both individuals and dimensions, a range of aggregation procedures satisfying non-dictatorship and non-dominance emerge, such as multidimensional utilitarian rules and multidimensional leximin rules.

Is it plausible to assume interpersonal comparability and interdimensional comparability of welfare? Let me first turn to interpersonal comparability.

Arrow himself excludes interpersonal comparisons of welfare from his framework, holding "that interpersonal comparison of utilities has no meaning and, in fact, that there is no meaning relevant to welfare comparisons in the measurability of individual utility" (Arrow, 1951/1963, p. 9).

In response to this view, two points must be noted. First, it is not obvious that interpersonal comparisons of utility are indeed as meaningless as Arrow claims they are, but even if interpersonal comparisons of utility were *empirically* meaningless, this would not entail that there is no other *non-empirical* but nonetheless normatively significant and non-arbitrary way of making such comparisons.

Second, as noted above, the question of whether interpersonal comparisons are meaningful in a given sense depends crucially on what welfare evaluation standard we choose to compare. To give just two examples, interpersonal comparisons of *money* are unproblematic (leaving practical issues aside), and interpersonal comparisons of the *amount of health care* or *education* a person has access to may also be unproblematic (again leaving practical issues aside). This means that, even if we concede Arrow's view that "interpersonal comparison of *utilities* has no meaning" (my italics), this entails *not* that interpersonal comparisons are in principle impossible, but only that such comparisons are possible only for certain welfare evaluation standards *other than* utility. And several welfare evaluation standards with the required properties have been proposed, including Rawls's index of primary goods and Sen's functionings.

Comparability across dimensions is a more serious problem. An important intuition underlying many arguments for a constitutively plural conception of welfare is that welfare has several, possibly mutually incommensurable, aspects, and that those aspects cannot easily be combined into a single overall index. Accordingly, the results of the present paper raise a tension that needs to be taken seriously. In particular, they raise a tension between the following three claims:

(i) There exist several normatively relevant dimensions of welfare, which (for instance, as a result of their different social meanings) are not mutually commensurable.

(ii) There are decisions which affect different dimensions, but which come in 'packages', indivisible into separate dimension-specific sub-decisions.

(iii) In making decisions of the kinds described by (ii), there should be no dominant dimension that has lexical priority in determining the outcome.

It seems that a violation of at least one of (i), (ii), or (iii) is inevitable. *Prima facie*, Walzer's approach to constitutive plurality seems more vulnerable to this tension than Sen's or Rawls's. Walzer's approach requires that different dimensions (spheres of goods) be kept separate. As a result of this requirement, the approach cannot accommodate cross-dimensional indexing as easily as Sen's or Rawls's approaches. The most consistent Walzerian escape from the tension, presumably the one pursued in *Spheres of Justice*, would be to try to avoid claim (ii). Indeed, when Walzer's argues that different distributive principles should be applied to different dimensions (spheres of goods), on the basis of the different social meanings of these dimensions (goods), Walzer’s proposal might be interpreted as a version of the above discussed strategy of subdividing decisions into several dimension-specific sub-decisions and thus 'defining away' the multidimensional aggregation problem.

But, as we have indicated above, while lexicographic dictatorships are hardly attractive from a liberal-egalitarian viewpoint, there are situations in which lexicographic hierarchies of dimensions (i.e. violations of claim (iii)) are defensible. (Thus there is an asymmetry between the normative appeal of non-dictatorship and the normative appeal of non-dominance.) Rawls's 'special conception' of justice explicitly assigns lexical priority to some primary goods, such as liberty and opportunity, over others, thus in effect defining a lexicographic hierarchy of dimensions. *If* there is a tension between holding a constitutively plural conception of welfare and accepting interdimensional comparability, then Rawls's defence of a lexicographic hierarchy of different primary goods in the 'special conception' of justice is clearly the social-choice-theoretically most compelling solution to that tension.

## Appendix 1. Proof of Theorems 4 and 5

Both theorems can be deduced from theorem 6 in Roberts (1995). A formally closely related result is Khmelnitskaya (1999, theorem 2), also presented with a modified proof in Khmelnitskaya and Weymark (2000, theorem 2), as mentioned above. This result uses a continuity condition in addition to conditions equivalent to the ones used here. The result can also be restated for the present multidimensional framework.

"Theorem 6. Let there be n opinions about the well-being of r individuals. (In particular, n = r is permitted.) If f satisfies [U][8], I, P[9], and FC[10], then there exists an individual i and a strictly monotonic function W with r arguments such that, for all x, $y \in X$,

$$W(u(x,\cdot,i)) > W(u(y,\cdot,i)) \Rightarrow xPy;$$

i.e. individual i's opinions are dictatorial."

We will explicitly deduce only theorem 2 from Roberts's theorem (hereafter simply referred to as R); theorem 3 can be deduced analogously.

Assume, for a contradiction, that $F$ is a MSWFL satisfying ($CN_{dim}L_{per}U_{per}$), (U), (P), (I) and (DOM) (any MSWFL satisfying ($ON_{dim}L_{per}$) will also satisfy ($CN_{dim}L_{per}U_{per}$)). We identify n in R (the number of opinions) with $k$ (the number of dimensions), r (the number of individuals) in R with $n$ (the number of individuals), u(x,i,j) (the opinion of j about the welfare of i under alternative x) in R with $W_{ij}(x)$ (the welfare of individual $i$ in dimension $j$ under alternative $x$), and f in R with $F$. It is easily seen that f satisfies R's conditions U, I, P and FC, whence, by R, there exists a strictly increasing function W with r (= $n$ in our framework) arguments such that, for all $x, y \in X$,

$$W(u(x,\cdot,i)) > W(u(y,\cdot,i)) \Rightarrow xPy \text{ (in R)},$$

i.e. for all $x, y \in X$,

$$f(\{W_{ij}(x)\}_{i\in N}) > f(\{W_{ij}(y)\}_{i\in N}) \Rightarrow xPy,$$

where the function W in R is identified with the function $f$. This contradicts condition (DOM) in our framework.

To deduce theorem 3, simply identify n in R with $n$, r in R with $k$, u(x,j,i) in R with $W_{ij}(x)$, and f in R with $F$. **Q.E.D.**

## References

Arrow, K. (1951 / 2nd ed. 1963). *Social Choice and Individual Values*. New York: Wiley.

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

Batra, R. and Pattanaik, P. K. (1972). Transitive Multi-Stage Majority Decisions with Quasi-Transitive Individual Preferences. *Econometrica* 40(6): 1121-1135.

Blair, D. H. (1988). The Primary-Goods Indexation Problem in Rawls's 'Theory of Justice'. *Theory and Decision* 24: 239-252.

Bossert, W. and Weymark, J. A. (1996). Utility in social choice. Forthcoming in S. Barberà, P. J. Hammond and C. Seidel (Eds.), *Handbook of Utility Theory, Volume 2*, Boston: Kluwer.

Fishburn, P. C. (1971). The Theory of Representative Majority Decision. *Econometrica* 39(2): 273-284.

Gibbard, A. (1979). Disparate Goods and Rawls' Difference Principle: A Social Choice Theoretic Treatment. *Theory and Decision* 11: 267-288.

Kelsey, D.: (1987). The role of information in social welfare judgements. *Oxford Economic Papers* 39: 301-317.

Khmelnitskaya, A. B. (1999). Social Welfare Orderings for Different Subgroup Utility Scales. Discussion Paper #198, Center for Rationality and Interactive Decision Theory, Hebrew University of Jerusalem.

Khmelnitskaya, A. B. and Weymark, J. A. (2000). Social choice with independent subgroup utility scales. *Social Choice and Welfare* 17: 739-748.

List, C. (2001). A Note on Introducing a 'Zero-Line' of Welfare as an Escape-Route from Arrow's Theorem. *Pacific Economic Review* 6 (special section in honour of Amartya Sen): 223-238.

List, C. (2002). Intradimensional Single-Peakedness and the Multidimensional Arrow Problem. *Theory and Decision* 52: 287-301.

Nussbaum, M. (1992). Human Functionings and Social Justice: In Defense of Aristotelian Essentialism. *Political Theory* 20: 202-246.

Plott, C. R. (1978). Rawls's theory of justice: An Impossibility result. In H. W. Gottinger and W. Leinfellner (Eds.), *Decision Theory and Social Ethics*, 201-214. Dordrecht: D. Reidel.

Plott, C. R., Little, J. T. and Parks, R. P. (1975). Individual choice when objects have 'ordinal' values. *Review of Economic Studies* 42: 403-413.

Roberts, K. (1995). Valued Opinions or Opinionated Values: The Double Aggregation Problem. In K. Basu, P. K. Pattanaik and K. Suzumura (Eds.), *Choice, Welfare and Development: A Festschrift in Honour of Amartya Sen*, 141-165. Oxford: Oxford University Press.

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

Sen, A. K. (1980/81). Plural utility. *Proceedings of the Aristotelian Society* 81: 193-215.

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

Sen, A. K. (1985). *Commodities and Capabilities*. Amsterdam: North-Holland.

Sen, A. K. (1987). *The Standard of Living*. Cambridge: Cambridge University Press.

Sen, A. K. (1991). On Indexing Primary Goods and Capabilities. Mimeo. Harvard University.

Sen, A. K. (1997). *On Economic Inequality*. Enlarged ed. Oxford: Oxford University Press.

United Nations Development Programme (2002). *Human Development Report 2002*. New York and Oxford: Oxford University Press.

Walzer, M. (1983). *Spheres of Justice*. New York: Basic Books.

## Notes

[1] *Interpersonal* and *interdimensional* comparability will be treated in formally similar ways. According to some philosophical conventions, what I call *comparability across dimensions* is called *commensurability across dimensions*. For simplicity, I will refer to *comparability* in the case of both persons and dimensions.

[2] The question of how the relevant dimensions are to be identified (i.e. the question of how the set $K$ is to be interpreted) is an important philosophical matter, but it will here be taken to lie outside the scope of social choice theory. Nussbaum's argument for a specific list of basic functionings (Nussbaum, 1992) and Walzer's account of how social meanings determine the relevant distributive spheres (Walzer, 1983) are two different approaches to the identification and demarcation of relevant dimensions, corresponding to the two different accounts of constitutive plurality briefly introduced above (essentialist and contextualist, respectively).

[3] A transformation $\phi : \mathbf{R} \rightarrow \mathbf{R}$ is *positive monotonic* if, for all $s$, $t$ in $\mathbf{R}$, $s<t$ implies $\phi(s) < \phi(t)$. $\phi : \mathbf{R} \rightarrow \mathbf{R}$ is *positive affine* if there exist $a$, $b$ in $\mathbf{R}$, with $b>0$, such that, for all $t$ in $\mathbf{R}$, $\phi(t) = a + bt$.

[4] For a more detailed discussion of the logical relation between meaningful statements and classes of admissible transformations in the one-dimensional context, see section 5 of Bossert and Weymark (1996).

[5] Note that all two-step aggregation mechanisms can also be interpreted as one-step aggregation mechanisms, but not all one-step aggregation mechanisms are expressible as two-step aggregation mechanisms.

[6] For simplicity, I will here focus only on leximin rules. The results can, however, be generalized to other positional rules. Further, the leximin rules discussed here all satisfy some form of an anonymity requirement, i.e. invariance under permutation of individuals and/or invariance under permutation of dimensions. For the discussion of non-anonymous positional rules in the one-dimensional case, see Bossert and Weymark (1996).

[7] To define a leximax rule, we simply need to replace all the "≤"-symbols with "≥"-symbols in the definition of the permutation $i \mapsto [i]$. Analogous remarks apply to the definitions below.

[8] Roberts proves this result for a domain more restricted than a universal domain (U), namely for an ordinal agreement unrestricted domain (OAU) (hence U is stated here in square brackets). He explains, however, that "OAU can be changed to NPU or U to give the same result as stated in Theorem 6" (p. 164).

[9] Roberts uses the strong Pareto principle, but points out that the weak Pareto condition is sufficient for the results of his paper (p. 144).

[10] Under the specified identification of Roberts's framework with our framework, FC is equivalent to ($CN_{dim}L_{per}U_{per}$).