Value-driven Multidimensional Welfare Analysis

A Dominance Approach with Application to Comparisons of European Populations

Argyris, Nikolaos; Østerdal, Lars Peter; Hussain, M. Azhar

Copenhagen Business School
HANDELSHØJSKOLEN

Department of Economics

Copenhagen Business School

Working paper 12-2023

Value-driven Multidimensional Welfare Analysis: A Dominance Approach with Application to Comparisons of European Populations

Nikolaos Argyris
Lars Peter Østerdal
M. Azhar Hussain

Department of Economics – Porcelænshaven 16A, 1. DK-2000 Frederiksberg

Value-driven Multidimensional Welfare Analysis: A Dominance Approach with Application to Comparisons of European Populations

Nikolaos Argyris[a,*], Lars Peter Østerdal[b], M. Azhar Hussain[c,d]

[a]*Loughborough Business School, Loughborough University, Loughborough, LE11 3TU, United Kingdom*
[b]*Department of Economics, Copenhagen Business School, Porcelænshaven 16A, DK-2000, Frederiksberg, Denmark*
[c]*Department of Finance and Economics, College of Business Administration, University of Sharjah, P.O. Box 27272, Sharjah, United Arab Emirates*
[d]*Department of Social Sciences and Business, Roskilde University, Universitetsvej 1, DK-4000, Roskilde, Denmark*

Abstract

We introduce a framework for social welfare evaluation that accommodates multiple dimensions of individual welfare, permits incorporating value judgements and enables robust social welfare comparisons. Our framework follows a dominance-based paradigm and utilises non-decreasing and potentially concave multi-attribute functions to model individual welfare. We describe how this permits capturing a variety of trade-offs between welfare attributes as well as incorporating concerns about distributional inequality in social welfare evaluation. We derive theoretical results which enable the practical implementation of our approach. Our framework incorporates a welfare measurement scale. This facilitates a richer form of analysis, compared to other dominance-based methods, from which we can gauge the overall level of social welfare in different populations relative to some meaningful benchmarks, as opposed to deriving only partial rankings. We illustrate the application of our framework with a case study investigating social welfare across 31 European countries based on the EU-SILC dataset.

Keywords: Multiple criteria analysis, social welfare, inequality aversion, value judgements, multidimensional stochastic dominance.

1. Introduction

Recent decades have witnessed a steady shift in perspectives about what constitutes societal progress. Across global and national policy arenas, there is now widespread recognition that progress extends beyond the confines of economic objectives. This is exemplified by the United Nations' adoption of the Sustainable Development Goals (UN, 2015), to serve as a broad set of shared global objectives to foster lasting prosperity for people and the planet. These goals encompass not only material conditions, such as 'no poverty' and 'decent work and economic growth', but also health, education, and equality, as represented by 'good health and well-being', 'quality education', 'gender equality' and 'reduced inequalities'. This multifaceted approach has spurred a burgeoning interest in developing multidimensional indices of societal progress, some of which garner considerable interest from policy makers and the media, such as the UN's Human Development Index and the OECD's Better Life Index (see, e.g., United Nations Development Programme, 2022; OECD, 2020). Human welfare is a key component of appraising social advancement and identifying inequalities. Both the research community and policy circles now widely acknowledge that individual well-being is inherently multidimensional, spanning not only material living standards such as individual income and wealth, but also dimensions such as health, education, social relationships, work and security (see, e.g., Stiglitz et al., 2009; Fleurbaey and Blanchet, 2013; Stiglitz et al., 2018). This conceptual framework is known as *multidimensional welfare* measurement.

[*]Corresponding author. Email: n.argyris@lboro.ac.uk

Multidimensional welfare evaluation entails comparing multidimensional distributions of various welfare measures (referred to as attributes) across several populations and/or over time. For example, comparing joint distributions of income and health, using metrics such as annual income and life-expectancy, across different countries. Such comparisons involve a double aggregation: over different attributes *and* over different individuals comprising a population. This aggregation task is fraught with ethical dilemmas but is also of foremost practical relevance: an acceptable apparatus for evaluating social welfare would be a vital input to public policy formulation and resource prioritisation (prospectively) and evaluation (retrospectively).

Broadly, approaches to multidimensional welfare measurement fall into two families: *aggregative* and *dominance-based*. Aggregative approaches employ parametric formulas to create a *composite index* from different welfare attributes. Their key advantage is that they are *decisive*: they allow for ranking all populations considered as well as assessing the level of social welfare using a meaningful measurement scale. However, their results can be highly sensitive to the choice of aggregation formula and parameter values. Additionally, many aggregative approaches work with aggregate (country-level) data, e.g. GDP per capita, which overlooks the issue of inequality within attribute distributions. These issues have spurred the development of dominance-based approaches. These seek to identify which populations are deemed superior in terms of welfare (i.e. "dominate" other populations) across a set of *admissible* aggregation possibilities. This set is modelled via imposing structural assumptions on the functions used to measure individual or social welfare, sidestepping the use of parametric formulas. Further, dominance methods are not restricted to using country-level data, but can consider the full joint distribution of attributes across a population. Dominance approaches are therefore a conceptual generalisation of aggregative approaches. A key advantage of this is that they are able to identify robust conclusions about social welfare across the populations considered, i.e. not reliant on specific formulas, parameter values, or summary statistics. However this can come with a significant loss in decisiveness (comparative power): often only a small number of robust comparisons can be made. Further, no measurement scale is used and so there is no way of identifying to what extent social welfare in some populations is higher than in some others, nor how they all fare relative to some meaningful benchmark. Finally, not all dominance approaches can utilise the cardinal nature of attributes such as income, life-expectancy, quality adjusted health (relying instead on ordinal attributes such as income quintile, access to healthcare, health quintile), and those that do require assumptions about global substitutability or complementarity in the attributes considered, e.g. assuming that the marginal effect on individual welfare arising from increases in income always increases as health increases.

In this paper we aim to address shortcomings of existing approaches for multidimensional welfare analysis, by introducing a new approach that bridges the gap between aggregative and dominance-based methods. This is done by using a framework that: (1) offers a flexible way of utilising cardinal data; (2) allows the use of a welfare measurement scale; (3) permits incorporating value judgements in the analysis; and (4) facilitates robust social welfare comparisons.

Our approach follows a dominance paradigm. As such it is designed to identify robust conclusions about social welfare in different populations. However, we use a theoretical framework that allows for utilising cardinal data without imposing global assumptions about the nature of substitution or complementarity relationships between individual welfare dimensions. Following Argyris et al. (2014), we accommodate this requirement by considering individual welfare functions that are non-decreasing and, potentially, concave. As we discuss, this allows for a variety of trade-offs between attributes but also permits introducing inequality aversion in the social welfare evaluation. A key contribution of this paper is that we introduce theoretical results to characterise admissible sets of social aggregation possibilities that utilise such functions. Based on these we derive a computational mechanism that enables the practical application of our approach, based on the solution of linear optimisation problems. Our framework can readily incorporate a welfare measurement scale and permits a far richer analysis of social welfare compared to existing dominance-based approaches. Overall, it enables reaching robust conclusions about the level of social welfare in different populations, rather than merely comparing ordinally. Our approach offers a way to extend the decisiveness of the analysis by following the paradigm of Multi-Criteria Decision Analysis (MCDA; see, e.g., Figueira et al., 2005). This involves restricting the set of admissible aggregations not only via structural assumptions, but also through incorporating value judgements (in a similar way to how MCDA approaches use 'preferences' to guide decision support). In this way we can incorporate in the evaluation a wealth of additional information about individual welfare, e.g. about attribute trade-offs or marginal effects, or social welfare, e.g. about the degree of inequality aversion. As we will see this can serve as an alternative, more flexible way of imposing additional structural assumptions. More importantly, it permits for accommodating societal views (and considering the implications of such) in a social welfare evaluation exercise. The versatility of our approach, both in terms of its theoretical features as well as the incorporation of value judgements, is illustrated with a case study that investigates welfare across 31 European countries based on the large and widely used EU-SILC dataset.

The rest of the paper is organised as follows. Section 2 provides a focused review of the related literature. Section 3 introduces notation and basic concepts. Sections 4 and 5, respectively motivate and describe the use of structural properties and value judgements in our approach. Sections 6 and 7 respectively introduce and extend our theoretical results. Section 8 describes our case-study. Section 9 provides a concluding discussion.

2. Related literature

The approach we introduce in this paper intersects a few strands of literature, namely: multidimensional stochastic orders, multidimensional welfare measurement and MCDA. We provide a focused review of these in this section.

As mentioned previously, our paper follows a dominance-based paradigm for social welfare evaluation. As such it adds to the extensive literature on (multidimensional) stochastic dominance concepts, also known as (multivariate) stochastic orders. For a general mathematical treatment we refer to the books by Müller and Stoyan (2002) and Shaked and Shanthikumar (2007). More specifically, our framework and methods are based on the concepts of multidimensional *first order dominance* and multidimensional *second order dominance*.

The multivariate first order dominance concept utilized in this paper is also referred to in the literature as the usual multivariate stochastic order (e.g. Lehmann, 1955; Levhari et al., 1975; Østerdal, 2010). A method for verifying first order dominance in the multivariate finite scenario has been outlined by Mosler and Scarsini (1991) as well as Dyckerhoff and Mosler (1997) via linear programming.[1] Arndt et al. (2012) and other subsequent studies provide computational implementations and empirical applications of such methods for welfare comparisons. However, these studies do not make use of value judgements for enhancing the analysis. Moreover, they do not consider the stronger multivariate second order dominance concept as we do in this paper.

The multivariate second order dominance concept utilized in this paper generalizes the well-known univariate second order dominance concept widely used in economics to the multivariate setting. A multivariate distribution second order dominates another if for any non-decreasing and concave utility function, the expected utility of the former is at least as high as of the other. In the stochastic dominance literature, it is known as the multivariate increasing concave order

[1] See Range and Østerdal (2019) for a discussion and a faster algorithm in the two-dimensional case.

(Shaked and Shanthikumar, 2007, Section 7.A). While the concept has been defined and theoretically studied, little is known about how to efficiently check second order dominance in two or more dimensions. To our knowledge, our study is the first to implement and apply a method to check for multivariate second order dominance. Moreover, as with the multivariate first order dominance, we make use of value judgements (preference-statements) for enhancing the analysis, which is a novel feature of our paper.

The use of value judgements considered in our paper strengthens the dominance concept by allowing more decisive comparisons. In this sense, our paper also relates to the emerging literature on multivariate almost stochastic dominance (e.g. Tsetlin and Winkler, 2018; Müller et al., 2023), in which conditions in terms of bounds on marginal utilities are used to strengthen the dominance concept. However, while multivariate almost stochastic dominance introduces restrictions on the allowed (relative) differences between marginal utilities for each dimension separately, the conditions we consider in our paper relate directly to comparisons of multivariate outcomes or distributions. Moreover, the conditions under multivariate almost stochastic dominance are required to hold generally (i.e. for any pair of outcomes) while we only impose value judgements for specific pairs of outcomes or distributions.

As mentioned in Section 1, multidimensional welfare measurement approaches fall into two families: aggregative and dominance-based. Aggregative approaches rely on specific formulas to combine data on different well-being dimensions into a composite index value (e.g. Nardo et al., 2008; Alkire and Foster, 2011). Typically, this involves taking a weighted sum, using pre-determined weights. Their main advantage is the ability to derive complete rankings of comparator populations. However, there is rarely, if ever, consensus on the "right" parameter values for the aggregation (e.g. weights) and different values may lead to different conclusions (rankings). This has spurred the development of a growing literature on dominance-based approaches to multidimensional welfare evaluation (for a number of references see below). Instead of considering a specific formula, these approaches identify which populations are deemed superior (i.e. "dominate" the others) across a range of admissible aggregation possibilities. Typically, admissible aggregation rules are defined implicitly, by recourse to a set of principles pertaining to how individual well-being is accepted to vary within and across dimensions – e.g. stipulating that individual well-being increases with income and that marginal well-being diminishes with income and/or joint increases in income and health. The advantage of dominance-based methods is that their comparative conclusions are "robust", i.e. invariant across a wide set of acceptable aggregation possibilities. However, this comes with a loss in comparative power: often robust comparisons can be derived for only a few pairs of population distribution. Some multidimensional dominance-based approaches can (and must) be used with cardinal data (e.g. Atkinson and Bourguignon, 1982; Bourguignon, 1989; Duclos et al., 2006, 2007; Gravel et al., 2009; Gravel and Mukhopadhyay, 2010; Muller and Trannoy, 2011; Duclos and Échevin, 2011; Marling et al., 2018). However, these approaches impose assumptions about the complementarity/substitutability of different well-being dimensions. For example, assuming that the marginal effect on well-being from increases in income always increases as health increases. This may only be true up to some level of health and has troubling implications about societal attitudes to inequality: it implicitly dictates a societal preference for income redistribution from unhealthier to healthier individuals. Other (first-order) dominance-based approaches can be used also with ordinal data (e.g. Arndt et al., 2012, 2016; Siersbæk et al., 2016; Hussain et al., 2016, 2020). These impose minimal aggregative assumptions, so their conclusions are most robust. However, when cardinal data is available, the richer information that this can bring into the analysis cannot be utilised. In sum, previously applied dominance-based methods can typically derive only a limited number of robust comparative conclusions and cannot utilise cardinal data without imposing restrictive assumptions on the analysis.

As detailed later in the paper, our framework enables calculation, for each population, of minimum and maximum possible values for social welfare, consistent with structural assumptions (and any value-judgements). This bears resemblance to work by Athanassoglou (2015), considering weighed aggregation as well as with Anderson et al. (2011), using a data envelopment approach to also establish lower and upper welfare bounds using weakly increasing and quasi-concave functions. These however utilise aggregate (country-level) data on several indicators, whereas our research takes a dominance-based approach so as to enable capturing population distributions across multiple indicators rather than summary measures.

Finally, as noted our method follows the MCDA paradigm of combining structural assumptions about utility functions with elicited preference statements in the context of decision support. We refer the reader to the comprehensive survey volumes by Figueira et al. (2005) and Ehrgott et al. (2010) and the systematic taxonomy by Cinelli et al. (2020). Our approach builds on the CUT method by Argyris et al. (2014) based on the use of concave and non-decreasing utility functions in the context of choice from a set of multi-attribute alternatives. We extend this method to allow for comparing probability (population) distributions over such alternatives. As previously noted the structural assumptions of our framework are particularly suited to social welfare evaluation, allowing for a variety of trade-offs between attributes but also for modelling inequality aversion – we discuss this in more detail in Section 4. Other MCDA methods utilise different structural assumptions. Of particular note are the methods on Robust Ordinal Regression (see Kadziński, 2022 for a recent survey). Like our approach, these are based on characterising a set of admissible utility functions compatible with specified structural properties and given preference statements. Original methods in this family utilised additive multi-attribute utility functions (e.g. Greco et al., 2008), but later ones (e.g. Greco et al., 2014; Angilella et al., 2016) introduced non-additive frameworks to capture interaction between attributes. The MCDA literature also intersects the broader literature on the development of composite indicators, considering both theoretical and methodological issues (e.g. Greco et al., 2019; Rowley et al., 2012; Lindén et al., 2021; Cherchye et al., 2007). MCDA methods have been extensively used in applications related to social welfare evaluation, particularly in environmental and sustainability assessments but also in health assessments. General issues and guidance on the use of MCDA in these settings are discussed in Merad et al. (2013), Cinelli et al. (2014) and Greco and Munda (2017). Most closely related to our empirical setting are the applications to compare different regions/countries in terms of sustainable development reported in Pérez-Ortiz et al. (2014), Pinar et al. (2014), Angilella et al. (2018), Greco et al. (2018), Resce and Schiltz (2021) and Cinelli et al. (2022). Other related applications include the development of robust composite measures of healthcare quality (Schang et al., 2016), as well as a composite index for corporate entities of encompassing financial, social and environmental performance (Gaganis et al., 2021). Unlike our approach, all these use aggregate data for the compared entities, not population distributions.

3. Preliminaries

In this section we introduce the notation used and the basic concepts on which our methodology is built.

3.1. Basic notation

Let I be a set of *attributes* indexed by $i = \{1, ..., m\}$. These consist of indicators that measure individual welfare across a number of dimensions.[2] Let $\mathcal{X}_i$ be the domain of each attribute $i \in I$

[2]For example, the material wealth, health and education dimensions can be respectively measured by use of the indicators 'annual equivalised income', 'self-assessed health' and 'years in education'.

and $\mathcal{X} = \mathcal{X}_1 \times ... \times \mathcal{X}_m$ be the multidimensional joint domain of the attributes. We assume that the attributes considered are cardinally measureable on interval scales.[3] On this basis we may assume with no loss of generality that all attribute scales are intervals (bounded or unbounded) on the positive part of the real line, i.e. $\mathcal{X}_i \subseteq \mathbb{R}_+ \ \forall i \in I$. For added convenience, we will assume that the attributes have a common domain, specifically $\mathcal{X}_i = \mathbb{R}_+ \ \forall i \in I$, so that $\mathcal{X} = \mathbb{R}^m_+$. This does not affect the validity of our results. We will refer to any element $x = (x^1, ..., x^m) \in \mathcal{X}$ as a multi-dimensional *outcome*, and any individual component x^i of this vector as the outcome level for attribute $i \in I$.

Let $T = \{1, ..., q\}$ be a set of *populations* (of individuals). Each population $t \in T$, will be characterised by a probability mass function $p_t(x)$ over a finite subset of $\mathcal{X}$. For convenience but without loss of generality, we will make two assumptions: (1) that the populations (distributions) considered have the same finite support, namely a set of distinct *reference outcomes* $X = \{x_1, ..., x_n\} \subset \mathcal{X}$; (2) that x_1 and x_n are, respectively, the *nadir* and *ideal* outcomes of X.[4] We index the set of reference outcomes X by $J = \{1, ..., n\}$. With these in place, each population $t \in T$ will be defined as a finite probability distribution $p_t = (p_{1t}, ..., p_{nt})$. We will use $P = \{p_1, ..., p_q\}$ to denote the set of all such probability distributions considered. Finally, we use $\mathbf{0} = (0, ..., 0)$ with the dimension of this vector determined by the context.

3.2. Social welfare evaluation

We consider a setting where a *social planner* seeks to compare a number of populations in terms of social welfare, taking into account a number of attributes. The standard concept for this purpose in welfare economics is the *social welfare function*, which aggregates individual welfare levels across a population into an overall measure of social welfare. The most commonly used aggregation formula employed in the literature (see, e.g., Atkinson and Bourguignon, 1982; Siersbæk et al., 2016) is *average utilitarianism*, which simply equates social welfare to average individual welfare. Specifically, if p is a discrete probability distribution over X, then the social welfare of the population described by p is given by the formula $\sum_{x \in X} p(x)u(x)$ (for a discussion on the axiomatic foundation of average utilitarianism see Blackorby and Donaldson, 1984; Blackorby et al., 2002). Using our set-up of common support set $X = \{x_1, ..., x_n\}$ and collection of populations described by p_t, average welfare evaluation takes the more specific form:

$$W_t = \sum_{j \in J} p_{jt} u(x_j) \ \forall t \in T. \tag{1}$$

Here, the function u is the *individual* welfare function, (commonly also referred to as an individual utility function) so that $u(x_j)$ denotes the welfare of an individual endowed with the multi-dimensional bundle of outcomes x_j. In the ensuing we will use the terms utility and individual welfare interchangeably. In this framework individual welfare is inter-personally comparable, both in terms of levels, but also in terms of increments (i.e. welfare differences) and the function u is unique up to increasing affine transformations[5]. Note, crucially, that the same function u is used for all individuals. Thus $u(x_j)$ is *not* the self-reported welfare of an individual endowed with x_j. Instead it reflects the social planner's view of the welfare level of such an individual. It may also reflect additional considerations, so that $u(x_j)$ may be more accurately described as the social planner's assessment of the contribution towards social welfare arising from the welfare of an individual endowed with x_j (see, e.g., Sen and Foster, 1997, p. 39). As we discuss subsequently,

[3]Such cardinal measurement is required only when utilising second-order but not first-order dominance (formally defined below).

[4]Specifically we have $\forall i \in I : x^i_1 = \min\{x^i_1, ..., x^i_n\}, \ x^i_n = \max\{x^i_1, ..., x^i_n\}$.

[5]v is a positive affine transformation of u if $v = au + b$ for some $a > 0, \ b \in \mathbb{R}$.

this permits that we encode u not only with properties that reflect individual welfare, but also properties that reflect desired principles of welfare aggregation (e.g. inequality aversion).

To use the formula in (1) we need to specify, to some degree, the individual welfare function u. Instead of specifying u fully, or parametrically, we follow a *dominance* paradigm: we consider a set of *admissible* individual welfare functions and seek to determine whether social welfare for some populations may be deemed superior to that of others for any admissible function u. Formally:

Definition 1. *Let $\mathcal{U}$ be a set of utility functions and let p and p' be two discrete population distributions over X. We will say that p $\mathcal{U}$-dominates p', denoted $p \trianglerighteq_{\mathcal{U}} p'$, if for any $u \in \mathcal{U}$ it holds that $\sum_{j\in J} p_j u(x_j) \geq \sum_{j\in J} p'_j u(x_j)$.*

The *dominance* relation $\trianglerighteq_{\mathcal{U}}$ is a *quasi-ordering*, i.e. a reflexive and transitive binary relation. Even though its definition uses average utilitarian social welfare comparisons, the quasi-ordering $\trianglerighteq_{\mathcal{U}}$ is *not* contingent on this. As shown in Gravel and Moyes (2013), the ordering $\trianglerighteq_{\mathcal{U}}$ generalises to the case where a broader class of social welfare functions is used (of which average utilitarianism is a special case). We discuss this in more detail in Section 7.2.

To use this dominance framework we need to specify the set of admissible functions $\mathcal{U}$. This will be done implicitly, by specifying: (1) structural properties that any admissible function must satisfy, combined with (2) value judgements that any such function must reflect. The structural properties determine what principles should underpin social welfare evaluation, both in terms of reflecting individual welfare as well as aggregating across individuals. The value judgements enable incorporating into the evaluation specific views about the welfare of individuals endowed with different multidimensional outcomes, or even populations of such individuals. We describe in the ensuing how such judgements permit incorporating a wealth of additional information, including about trade-offs between different attributes or marginal effects of different attributes on individual welfare. As already noted this approach is in the spirit of a large family of methods within the field of Multi-Criteria Decision Analysis (MCDA), where structural information about utility functions is combined with *preference statements* elicited from stakeholders in order to compare multi-dimensional alternatives and provide decision support. For this reason we will also refer to the value judgements as preference statements, but it should be borne in mind that these are meant to reflect *social* rather than individual preferences. In the following two sections we discuss how we specify $\mathcal{U}$ along these lines.

4. Structural properties

We consider two cases of imposing structural properties on u. In the first case, u is only assumed non-decreasing. In the second case, u is additionally assumed concave.[6] The non-decreasingness assumption is not controversial, especially within the context of social welfare evaluation, and is standard in the literature; it stands to reason that social welfare evaluation can be predicated by the assumption that increases in, e.g., income, health, education, ought not to make that individual worse off for the purposes of the social welfare evaluation exercise.[7] The concavity assumption is perhaps more involved. In the setting of uni-dimensional welfare comparisons, usually related to income or health, the use of univariate non-decreasing or non-decreasing and concave utility

[6]Non-decreasing means that $u(x) \geq u(x')$ for any two $x, x' \in \mathcal{X}:\ x \geq x'$. Concavity means $u(\lambda x + (1-\lambda)x') \geq \lambda u(x) + (1-\lambda)u(x')$, for any $\lambda \in [0,1]$ and any two $x, x' \in \mathcal{X}$.

[7]One may contend that empirical research does not always link increases in wealth/income or education to greater subjective well-being. Notwithstanding, social welfare evaluation requires a normative foundation. It is difficult to imagine a social welfare evaluation context where it is acceptable, *ceteris paribus*, to assume that individual welfare decreases after some income/wealth or education level.

functions, in conjunction with a utilitarian aggregation formula, has a long-standing tradition (e.g. Atkinson, 1970). This renders the use of their multivariate counterparts an obvious choice in the multi-dimensional setting. Beyond that, our choice of using concavity is additionally motivated by two fundamental reasons.

Firstly, we want to allow for interaction effects among different attributes without having to specify the form of this interaction *a priori*. As noted by Argyris et al. (2014), concave utility functions can capture constant, increasing or decreasing marginal rates of substitution between different attributes; respectively the utility function can be both concave and additive, or super- or sub-modular. Specifying only that it must be concave, permits all these other structural properties, and leaves open whether these may be imposed indirectly (through preference statements as discussed below).

Secondly, we want to allow for introducing *inequality aversion* in the social welfare evaluation, as is conventional in the literature on social evaluation. If marginal individual welfare is assumed to be diminishing in each attribute, then the formula in (1) implies that a progressive re-distribution of, say, income, from richer to poorer individuals improves social welfare. A similar situation arises for joint increases in multiple attributes (cf the property of Non-Increasing Intensities, which characterises the concavity of u as shown in Argyris et al., 2014).

We will, therefore, consider the following two sets of utility functions:

$$\mathcal{U}^1 = \{u(\cdot) : \mathcal{X} \to \mathbb{R} |\ u(x_1) = 0,\ u(x_n) = 100,\ u(\cdot) \text{ is non-decreasing}\}, \tag{2}$$

$$\mathcal{U}^2 = \{u(\cdot) \in \mathcal{U}^1 |\ u(\cdot) \text{ is concave}\}. \tag{3}$$

Note that $\mathcal{U}^2 \subseteq \mathcal{U}^1$, so $u \in \mathcal{U}^2$ is both non-decreasing and concave. The normalisation $u(x_1) = 0$ and $u(x_n) = 100$ may be assumed without loss of generality, as some positive affine transformation of a concave and non-decreasing utility function can always be constructed to meet these conditions, and this operation would preserve both properties considered. Given that $x_1,\ x_n$ were defined as the nadir/ideal outcomes respectively, this bounds $W_t \in [0, 100]\ \forall t \in T$, no matter the individual utility function u.

Using $\mathcal{U}^1$ in Definition 1 we derive the dominance relation $\trianglerighteq_{\mathcal{U}^1}$, which we refer to as (multidimensional) first order dominance. Similarly we refer to $\trianglerighteq_{\mathcal{U}^2}$ as (multidimensional) second order dominance.

5. Incorporating value judgements

We now turn to the issue of utilising value judgements (preference statements) in social welfare evaluation. This involves restricting the set of admissible individual welfare functions by requiring that any such function u is compatible with some specified preferences over multidimensional outcomes in $\mathcal{X}$. Below we first discuss the additional power that this may bring to the social welfare evaluation exercise. We then detail how specific forms of value judgements can be incorporated.

5.1. Motivation for utilising preference statements

Utilising preference statements can drastically improve the comparative power of the analysis. As already noted, dominance-based social welfare evaluation approaches may exhibit weak comparative power, i.e. many distributions may be incomparable. By introducing value judgements to a dominance-based analysis, complementary to minimal structural assumptions, we further restrict the set of compatible functions $\mathcal{U}$. This can only ever 'sharpen' the quasi-ordering $\trianglerighteq_{\mathcal{U}}$. This combination of preferences with structural properties has a long tradition in MCDA, and is a key aspect of many MCDA approaches for decision support. The results of our case study, reported later, clearly illustrate the value of this approach.

Introducing preference statements can be seen as an alternative to imposing stronger structural properties. Specifically, it permits for enforcing such properties only at specified points/regions of the multidimensional outcome space. For example, we can use value judgements (we illustrate how below) to indicate a diminishing marginal effect of income as health improves *given* that income and health are low to begin with. This would indicate substitutability of these two attributes for low attribute levels, but it would leave open the possibility that the relationship may be different for higher levels of either attribute.

Utilising preference statements can be compelling from an ethical perspective. On their own, minimal structural properties are insufficient to compare even strikingly different outcomes. Take for example the comparison of two individuals, where one is very healthy but earns marginally less income than the second individual who is in much poorer health. No structural property of the individual welfare function (whether concavity, additivity, sub- or super-modularity) can, on its own, imply any preference for either of these states. From the perspective of individual preferences, this may be appropriate. But if we also accept that the two states are incomparable in the eyes of the social planner, then so would two populations comprising, respectively, only of individuals in one of the two states. Viewed from this angle, the weak comparative power of dominance-based social welfare evaluation approaches is simply an artifact of the flexibility they are designed to accommodate from the perspective of individual preferences. It is hardly surprising that social welfare evaluation would become meaningless unless we can accept that we can make *some* comparisons of individuals in different states. A framework that incorporates preferences in social welfare evaluation would not dictate to the planner a preference between such states. Instead it allows them to make a value judgement. Nor would it compel them to compare these two specific states. Instead it allows them to state value judgements with reference to arbitrary states they wish to compare. Utilising such judgements can lead to comparative conclusions about the distributions evaluated that would otherwise be indiscernible. In sum, preferences information sharpens the ability to make social welfare comparisons in a way that is compatible with the value judgements that a social planner deems should influence social welfare evaluation.

Finally, preference statements can be used to influence the degree of inequality-aversion of the social welfare evaluation exercise. As noted previously, this is linked to concavity of the individual welfare function. In particular, the degree of concavity of individual welfare (i.e. the curvature of the individual welfare function) dictates the degree of aversion towards distributional inequalities. While all functions in set $\mathcal{U}^2$ are concave, some are more concave than others. Indeed even a linear function is, by definition, concave and is included in this set; and so are functions that are ever so slightly curved. In general, $\mathcal{U}^2$ may contain utility functions that are not sufficiently curved to induce the desired degree of inequality aversion in social welfare evaluation, a point that has been noted by Sen and Foster (1997, p. 21). As we detail in the following sections, it is possible to influence the degree of inequality aversion in the social welfare evaluation exercise through specific types of preference statements that involve comparisons of marginal welfare changes or comparisons of population distributions.

In the following sections we detail the specifics of how different types of value judgements define constraints on admissible utility functions in $\mathcal{U}$. We will assume that value judgements involve outcomes from set X or distributions from set P. This is with no loss of generality as these sets can always be enlarged to include any outcomes or distributions involved in the value judgements utilised. In practice, the use of hypothetical outcomes or distributions can help making such comparisons easier. For example, it may be easier to contemplate trade-offs between two attributes by comparing outcomes that only differ in these attributes. In comparing marginal utility of income, it may be easier to compare outcomes that only differ in income. Similarly, in comparing changes of marginal utility of income for different health levels it makes sense to compare outcomes that only differ in these two attributes. Finally, in making direct distributional comparisons, it may be easier to compare hypothetical distributions that only differ in one of the attributes and, additionally, where the difference of the marginal distributions on that same attribute can be easily understood (for example, one has higher average income but is spread over a larger range and with more probability mass at the lower end).

5.2. Ordinal preferences

The simplest case of preference statements involves comparisons of different multidimensional outcomes. To give a simple (hypothetical) example in a two-dimensional setting, the planner may consider that the well-being of an individual with income of 50000 EUR and a self-assessed health status of 5 on a 1-5 scale is no lower than that of an individual with income of 55000 EUR and health status of 2 on the same scale. We refer to such an ordinal comparison of multidimensional outcomes as a *weak* ordinal preference statements, as opposed to a *strict* one wherein the well-being of the former bundle would be deemed strictly higher (the same concept is used to distinguish between weak and strict statements in the all forms of preference statements considered). Formally, we will use the binary relation $\succeq$ (resp. $\succ$) to capture weak (resp. strict) ordinal preferences to be incorporated in the social welfare evaluation exercise[8]. Using this we may denote the comparison of the preceding example denoted by $(50000, 5) \succeq (55000, 2)$. Imposing this requirement on a set of utility functions $\mathcal{U}$ means that we require that any $u \in \mathcal{U}$ be compatible with this statement, i.e. that we have: $u(50000, 5) \geq u(55000, 2)$. In general, we encode a weak preference statement involving two outcome bundles $x_j, x_{j'} \in X$ as a constraint on utility functions $u \in \mathcal{U}$ as follows:

$$x_j \succeq x_{j'} \Leftrightarrow u(x_j) \geq u(x_{j'}). \tag{4}$$

5.3. Cardinal preferences

A second type of preference statements involves cardinal preferences over multidimensional outcomes (e.g. involving comparisons of marginal utility), also known as *preference intensities.*[9] In the same two-dimensional setting considered previously, for example, the social planner may feel that the increase in individual welfare of an extra 1000 EUR for an individual endowed with bundle $(50000, 5)$ is no less than that for an individual endowed with $(55000, 2)$. Formally, weak (resp. strict) preference intensities are captured by the quaternary relation $\succeq^*$ $(\succ^*)$. Using this we denote the comparison of the preceding example by $(51000, 5) \leftarrow (50000, 5) \succeq^* (56000, 2) \leftarrow (55000, 2)$. Imposing this requirement on a set of utility functions $\mathcal{U}$ requires that for any $u \in \mathcal{U}$ we have: $u(51000, 5) - u(50000, 5) \geq u(56000, 2) - u(55000, 2)$. In general, we encode a preference intensity statement involving four outcome bundles $x_j, x_{j'}, x_k, x_{k'} \in X$, as a constraint on utility functions $u \in \mathcal{U}$ as follows:

$$x_j \leftarrow x_{j'} \succeq^* x_k \leftarrow x_{k'} \Leftrightarrow u(x_j) - u(x_{j'}) \geq u(x_k) - u(x_{k'}). \tag{5}$$

5.4. Distributional preferences

A third possibility is to incorporate direct social welfare comparisons as a form of preference statements. This involves expressing ordinal preferences over population distributions of multidimensional outcomes, which we refer to as distributional preferences. In the same two-dimensional setting considered previously, for example, the social planner may feel that the social welfare in a population where 50% of individuals are endowed with bundle $(50000, 2)$ and 50% with bundle

[8]to be precise, $\succeq$ and $\succ$ are a specified quasi-orderings of outcomes in X.

[9]Arguments for meaningfulness of such preference statements are presented in e.g. Edwards and Von Winterfeldt (1986) and Harvey and Østerdal (2010).

$(60000, 5)$ is no higher than social welfare in a population where everyone 100% of individuals are endowed with $(55000, 3)$. Weak (resp. strict) distributional preference statements are captured by a binary preference relation $\succeq'$ ($\succ'$) over probability distributions over X. Using this we denote the preceding example by $((1, (55000, 3)) \succeq' 0.5, (50000, 2); 0.5, (60000, 5))$. Imposing this requirement on a set of utility functions $\mathcal{U}$ would then require that for any $u \in \mathcal{U}$ we have: $u(55000, 3) \geq 0.5u(50000, 2) + 0.5u(60000, 5)$. In general, we encode a preference involving two distributions $p_t, p_{t'} \in P$ as a constraint on functions $u \in \mathcal{U}$ as follows:

$$p_t \succeq' p_{t'} \Leftrightarrow W_t = \sum_{j \in J} p_{jt'} u(x_j) \geq W_{t'} = \sum_{j \in J} p_{jt'} u(x_j) \tag{6}$$

While the three types of preference statements considered are conceptually distinct, it is possible to express them in a unified form. Ordinal preferences can be re-cast as distributional preferences simply by using degenerate probability distributions with a single possible outcome[10] Cardinal preferences can be re-cast as distributional preferences using two-outcome 50%/50% distributions.[11] Therefore, with no loss of generality we will consider distributional preferences in our mathematical formulations.

5.5. *Strict preferences*

Though not explicitly stated above, strict forms of preference (i.e. given by $\succ$, $\succ^*$ or $\succ'$) are accommodated by replacing weak inequalities ($\geq$) with strict inequalities ($>$) in (4)-(6), as required. For example, the strict ordinal judgement $(50000, 5) \succ (55000, 2)$ defines the constraint: $u(50000, 5) > u(55000, 2)$. In practical computations, strict inequalities can only be enforced to some finite degree of precision, and so we may also model them as weak inequalities using a small positive threshold $\varepsilon > 0$ as a lower bound for utility differences. For example, the aforementioned strict preference statement would be modelled by the weak inequality $u(50000, 5) - u(55000, 2) \geq \varepsilon$. We can set ε to be positive but as small as the numerical precision permits. In a cardinal framework, however, the magnitude of the ε threshold is a meaningful concept: it defines a minimal discernible difference in individual welfare (given the scale). To define this independently of the scale used, we can specify a minimal proportion of the scale's range[12]. This offers an easy way of incorporating richer value judgements. In the context of preference-guided decision support, Argyris et al. (2014) additionally cluster ordinal statements based on preference intensities, by asking decision makers to additionally consider if their stated preference is 'weak', 'moderate' or 'strong', and report notable gains in the comparative power of the analysis. Similar strategies can be implemented in our setting. To give a simple example, we could implement a threshold of, say, 10% of the maximal difference (i.e. set $\varepsilon = 10$) and enforce this for any stronger preferences. In general, using higher threshold values impose stronger restrictions on the sets of admissible individual welfare functions and are expected to increase the comparative power of the analysis. Nevertheless, in our case study we use a smaller value of $\varepsilon = 1$, which corresponds to 1% of the 0-100 scale range.

5.6. *Formal set-up*

We present here how the aforementioned types of preference statements are used to restrict the set of admissible individual welfare functions. With no loss of generality, we make three

[10] $x_j \succeq x_k \Leftrightarrow (0, x_1; ...; 1, x_j; ...; 0, x_)) \succeq' (0, x_1; ...; 1, x_k; ...; 0, x_n)$.

[11] $x_j \leftarrow x_{j'} \succeq^* x_k \leftarrow x_{k'} \Leftrightarrow u(x_j) - u(x_{j'}) \geq u(x_k) - u(x_{k'}) \Leftrightarrow 0.5u(x_j) + 0.5u(x_{k'}) \geq 0.5u(x_{j'}) + 0.5u(x_k) \Leftrightarrow (0.5, x_j; 0.5, x_{k'}) \succeq' (0.5, x_{j'}; 0.5, x_k)$.

[12] Ratios of welfare differences are constant across positive affine transformations of the individual welfare function. Therefore, for any two outcomes $x, x' : x \succeq x'$ the ratio $\frac{u(x)-u(x')}{u(x_1)-u(x_n)}$ is constant across all positive affine transformations of u. Then we can define a meaningful minimum discernible proportion of the maximal welfare difference $u(x_1) - u(x_n)$, which in our case has been normalised to 100. Multiplying this value with 100 gives the minimum discernible welfare difference ε, on a 0-100 scale.

simplifications for the sake of presenting concise formulations. Firstly, we (equivalently) encapsulate all preference types of value judgements as distributional preferences, given by the relation $\succeq'$. Secondly, we assume that the preferences encapsulated in $\succeq'$ involve only outcomes from set X and distributions from set P (we can always enlarge these two sets to satisfy this assumption). Thirdly, we do not explicitly consider strict preference statements $\succeq'$ (adding them using the threshold approach does not affect our results). With these in place we may define $\mathcal{U}^1(\succeq')$, $\mathcal{U}^2(\succeq')$, the subsets of individual welfare functions compatible with $\succeq'$, as follows:

$$\mathcal{U}^d(\succeq') = \{u \in \mathcal{U}^d | \sum_{j\in J} p_{jt'}u(x_j) \geq \sum_{j\in J} p_{jt''}u(x_j) \; \forall t', t'' : p_{t'} \succeq' p_{t''}\}, \; d = 1, 2. \tag{7}$$

We are interested in the quasi-orderings obtained by using $\mathcal{U}^d(\succeq')$, $d = 1, 2$, in Definition 1, for which we reserve the following special notation:

$$\trianglerighteq^d = \trianglerighteq_{\mathcal{U}^d(\succeq')}, \; d = 1, 2, \tag{8}$$

6. Main theoretical results

We now deal with the problem of how to verify whether $p \trianglerighteq^1 p'$ or $p \trianglerighteq^2 p'$ for two population distributions $p, p' \in P$.

For any set of functions $\mathcal{U}$, we define $\mathcal{U}_{[X]}$ as the set of images of functions in $\mathcal{U}$ restricted to the set of outcomes X, i.e.:

$$\mathcal{U}_{[X]} = \{(u(x_1), ..., u(x_n)) \mid u \in \mathcal{U}\}. \tag{9}$$

We are specifically interested in the sets $\mathcal{U}^d_{[X]}$ and $\mathcal{U}^d_{[X]}(\succeq')$, $d = 1, 2$. We will derive polyhedral descriptions for these sets.

Let $\upsilon = (\upsilon_1, ..., \upsilon_n) \in \mathbb{R}^n$ denote a utility assignment (i.e. an assignment of utility values) to outcomes in X. We will consider the following two sets of utility assignments:

$$U^1 = \left\{ \upsilon \in \mathbb{R}^n \;\middle|\; \begin{array}{ll} \upsilon_j = w_{jk}x_j + \beta_{jk} & \forall j, k \in J \\ \upsilon_k \leq w_{jk}x_k + \beta_{jk} & \forall j, k \in J : j \neq k \\ w_{jk} \in \mathbb{R}^m_+, \; \beta_{jk} \in \mathbb{R} & \forall j, k \in J \\ \upsilon_1 = 0, \; \upsilon_n = 100 & \end{array} \right\}, \tag{10}$$

$$U^2 = \left\{ \upsilon \in \mathbb{R}^n \;\middle|\; \begin{array}{ll} \upsilon_j = w_j x_j + \beta_j & \forall j \in J \\ \upsilon_k \leq w_j x_k + \beta_j & \forall j, k \in J : j \neq k \\ w_j \in \mathbb{R}^m_+, \; \beta_j \in \mathbb{R} & \forall j \in J \\ \upsilon_1 = 0, \; \upsilon_n = 100 & \end{array} \right\}. \tag{11}$$

We may now state our main result which shows[13] that $U^d = \mathcal{U}^d_{[X]}$, $d = 1, 2$ (all proofs are in Appendix A).

Theorem 1. *For $d = 1, 2$, we have $\upsilon \in U^d \Leftrightarrow \upsilon \in \mathcal{U}^d_{[X]}$.*

Now consider the following sets of utility assignments:

$$U(\succeq') = \{\upsilon \in \mathbb{R}^n | \sum_{j\in J} p_{jt'}\upsilon_j \geqslant \sum_{j\in J} p_{j,t''}\upsilon_j \; \forall t', t'' : p_{t'} \succeq' p_{t''}\} \tag{12}$$

[13]For the case $d = 1$ the result is new. For the case $d = 2$ the result extends the corresponding Theorem in Argyris et al. (2014) which is restricted to the case where $X \subset \mathbb{R}^m_{++}$ (i.e. excluding the possibility for 0 coordinates in outcomes of X) and also simplifies some parts of its proof.

$$U^d(\succeq') = U^d \cap U(\succeq') \tag{13}$$

We refer to the above as sets of compatible utility assignments (i.e. compatible with $\succeq'$). We may now state the following generalisation of Theorem 1.

Corollary 1. *For $d = 1, 2$, we have: $\upsilon \in U^d(\succeq') \Leftrightarrow \upsilon \in \mathcal{U}^d_{[X]}(\succeq')$.*

With this in place it is easy to verify dominance. Specifically, let p' and p'' be two population distributions. We can verify whether $p' \trianglerighteq^d p''$ by solving a linear-programming problem, according to the following:

Theorem 2. *For $d = 1, 2$: $p' \trianglerighteq^d p'' \Leftrightarrow \min\{\sum_{j \in J} p'_j \upsilon_j - \sum_{j \in J} p''_j \upsilon_j | \ \upsilon \in U^d(\succeq')\} \geq 0$.*

Because our framework incorporates a scale for social welfare and explicitly models individual welfare (utility) values, it enables a richer form of analysis complementary to dominance-based comparisons. Based on Theorem 1, we can calculate, for each population, minimum (worst case) and maximum (best case) values for social welfare. This is formally stated in Theorem 3 below. As we illustrate in our case study, this result enables assessing the level of social welfare in each population, as well as the degree to which some populations may have higher welfare than others.

Theorem 3. *For $d = 1, 2$: $\max\{\sum_j p'_j u(x_j)| \ u \in \mathcal{U}^d(\succeq')\} = \max\{\sum_j p'_j \upsilon_j| \ \upsilon \in U^d(\succeq)\}$ and $\min\{\sum_j p'_j u(x_j)| \ u \in \mathcal{U}^d(\succeq')\} = \min\{\sum_j p'_j \upsilon_j| \ \upsilon \in U^d(\succeq)\}$.*

7. Extensions to theoretical results

In this section we outline how our main results can be extended to capture: strictly increasing individual welfare functions; and a broader set of social welfare functions.

7.1. Strictly increasing functions

Although we have considered non-decreasing functions in the definitions of $\mathcal{U}^1$ and $\mathcal{U}^2$ above, our results can be readily modified to the case where only strictly increasing functions are permitted. In principle, some non-decreasing functions may be deemed unrealisitic for social welfare evaluation, in the sense that they do not reflect an adequately positive influence on individual welfare from increases in some of the attributes. For example, even a function that is flat everywhere except for the nadir of X is included in $\mathcal{U}^1$ and $\mathcal{U}^2$. Similarly, both sets include functions that imply almost zero increase in individual welfare even after substantial increases in any or all attributes. We can enforce strictly increasing individual welfare functions via a minor modification in (10) and (11): using strict inequalities $w_{jk} > 0$ and $w_j > 0$ in these two formulations respectively. In parallel to Theorem 1, we can establish a connection between these modified formulations and the sets of strictly-increasing (and additionally concave) functions. This is formally stated in the following.

Theorem 4. *Let $\overline{\mathcal{U}}^d \subset \mathcal{U}^d$, $d = 1, 2$, respectively denote the sets of strictly increasing, and strictly increasing and concave utility functions.*[14] *Let $\overline{U}^d$, $d = 1, 2$, be defined by respectively restricting w_{jk} in (10) so that $w_{jk} > 0$, and w_j in (11) so that $w_j > 0$. Then for $d = 1, 2$ we have $\upsilon \in \overline{U}^d \Leftrightarrow \upsilon \in \overline{\mathcal{U}}^d_{[X]}$.*

[14] A function u is strictly increasing if for $x, x' \in \mathbb{R}^m : x' \geq x, \ x' \neq x$ we have $u(x') > u(x)$.

Analogously, we can adapt all results in the previous section to account for the use of only strictly increasing functions, by replacing U^d with $\overline{U}^d$ and $\mathcal{U}^d_{[X]}$ with $\overline{\mathcal{U}}^d_{[X]}$. As already commented, in practice strict inequalities are implemented via a threshold approach. This means using $w_{jk} \geq \delta$ and $w_j \geq \delta$, i.e. enforcing a positive lower bound $\delta > 0$. These variables define local gradients of individual welfare functions at various points x_j. Because gradients represent marginal changes (i.e. welfare differences), the value of the lower bound δ is a meaningful concept in a cardinal framework. By using larger values (possibly enforced only at specific x_j points), we can exclude functions that are deemed insufficiently sensitive to increases in some of the attributes. Such an approach would be similar to how 'extreme' utility functions are excluded using the concept of 'almost stochastic dominance' (Leshno and Levy, 2002) in individual choice under risk (see also Section 2 for references regarding the multivariate case). In our computational trials we found that using larger values for threshold can have noticeable impact on the dominance instances identified. Therefore, we have kept the threshold used in our case study low: setting $\delta = 10^{-3}$, which is quite small (given the 0-100 welfare scale).

7.2. Welfarist social orderings

We elaborate here on a point made previously, about the validity of our results under a more general social welfare evaluation framework (i.e. not restricted to utilitarianism). Let $\mathcal{U}$ be an arbitrary set of individual welfare functions. For any $u \in \mathcal{U}$ let $\mathbf{u}_t$ be the vector of individual welfare levels according to u for the population described by distribution p_t.

Assume for now that all populations in T have the same total number of individuals, denoted q, so that for any $u \in \mathcal{U}$, all $\mathbf{u}_t$ are of length q. Then instead of comparing averages of individual welfare across populations we may equivalently compare the sum-totals. The utilitarian quasi-ordering relative to $\mathcal{U}$, denoted $\succeq_{\mathcal{U}}$ is then defined by the pairs of populations where one has no less total welfare than another across all functions in $\mathcal{U}$. A more general framework would involve modelling the social welfare in each population as $w(\mathbf{u}_t)$ $t \in T$, where $w(\cdot)$ is a social welfare function according to which the individual welfare levels are aggregated. Let $\mathcal{W}$ be a set of admissible social welfare functions. Then the *welfarist* quasi-ordering, denoted $\trianglerighteq_{\mathcal{W}(\mathcal{U})}$ consists of the pairs of populations where social welfare is no less in one than the other for all admissible individual utility functions $u \in \mathcal{U}$ and social welfare functions $w \in \mathcal{W}$. Various properties are typically considered in the literature to specify the class $\mathcal{W}$. The property of *impartiality* is typically used, to capture the principle that the identities of the individuals involved are immaterial. This requires *symmetry* of the welfare function so that $w(\mathbf{u}_t) = w(\mathbf{\Pi}\mathbf{u}_t)$ for any $q \times q$ permutation matrix $\mathbf{\Pi}$. Further, a requirement of *efficiency* is typically imposed, to capture the principle that an increase in any individual's utilities, *ceteris paribus*, cannot decrease social welfare. This requires that the social welfare function is non-decreasing. We use $\mathcal{W}^1$ to denote the set of symmetric non-decreasing welfare functions. Finally, to introduce an aversion to inequality in the distribution of individual welfare, the welfare function is usually assumed to be *Schur-concave*. This means that $w(\mathbf{B}\mathbf{u}_t) \geq w(\mathbf{u}_t)$ for any $q \times q$ bi-stochastic matrix $\mathbf{B}$.[15] This property can be shown to be equivalent to a preference for re-distribution of welfare from a better-off to a worse-off individual. We use $\mathcal{W}^2$ to denote the set of non-decreasing and Schur-concave functions. Note that Schur-concavity implies symmetry and that sum aggregation (utilitarianism) is included in both $\mathcal{W}^1$ and $\mathcal{W}^2$.

Under mild conditions, the generality of considering sets $\mathcal{W}^1$ or $\mathcal{W}^2$, has no material impact on the ordering of populations (Gravel and Moyes, 2013). Specifically, the condition that $\mathcal{U}$ is closed under composition with non-decreasing functions is sufficient to guarantee that $\trianglerighteq_{\mathcal{W}^1(\mathcal{U})} = \trianglerighteq_{\mathcal{U}}$.

[15] A square matrix is bistochastic if the all its entries are nonnegative reals and each of its rows and columns sums to 1.

Further the condition that $\mathcal{U}$ is closed under composition with non-decreasing and concave functions is sufficient to guarantee that $\trianglerighteq_{\mathcal{W}^2(\mathcal{U})} = \trianglerighteq_{\mathcal{U}}$. In our setting, the composition of a non-decreasing function with a non-decreasing function is non-decreasing and so we obtain that $\trianglerighteq_{\mathcal{W}^1(\mathcal{U}^1)} = \trianglerighteq_{\mathcal{U}^1}$.[16] Further, a concave and non-decreasing transformation of a concave and non-decreasing function is also concave and non-decreasing, so that we also obtain $\succeq_{\mathcal{W}^2(\mathcal{U}^2)} = \succeq_{\mathcal{U}^2}$.[17] Finally, since any non-decreasing transformation of a utility function represents the same preference ordering as the original function, then for the special case where $\succeq'$ consists of only ordinal preference statements, denoted $\succeq$ we have $\succeq_{\mathcal{W}^d(\mathcal{U}^d(\succeq))} = \succeq_{\mathcal{U}^d(\succeq)}$, $d = 1, 2$.

Similar equivalences can be obtained when the sizes of the compared populations differ. This is because the comparison of two populations of different size based on average individual welfare is equivalent to the comparison of two different populations of equal size, constructed from the original two populations, using (sum) total welfare. Let q_t be the total number of individuals of population $t \in T$ and q_{jt} be the number of individuals of population $t \in T$ endowed with x_j, $j \in J$. Consider two populations $t, t' \in T$ and some $u \in \mathcal{U}$. Starting from average utilitarianism we obtain: $\sum_j p_{jt} u(x_j) \geq \sum_j p_{jt'} u(x_j) \Leftrightarrow \sum_j \frac{q_{jt}}{q_t} u(x_j) \geq \sum_j \frac{q_{jt'}}{q'_t} u(x_j) \Leftrightarrow \sum_j q_{t'} q_{jt} u(x_j) \geq \sum_j q_t q_{jt'} u(x_j)$. The left hand side of this inequality can be considered as total individual welfare in a population constructed by replicating each individual in population t as many times as the number individuals in population t', and vice versa for the right hand side, so that both constructed populations are of size $q_t q_{t'}$. Comparisons of such constructed populations via a utilitarian based quasi-ordering is equivalent to comparisons based on a welfarist quasi-ordering in the sense discussed previously.

8. Case Study: Welfare comparisons between European countries

Our framework can be readily applied in practice, to make multidimensional welfare comparisons of different populations. Here we report on an application to compare welfare across 31 European countries, using three attributes: *Income*, *Health* and *Education*. The results give a clear demonstration of the potential of our approach. Comparing with a baseline scenario of first order dominance (FOD), we find that the discriminatory power of the analysis improves markedly with the use of the second order dominance (SOD) framework. We also find that utilising value judgements (by means of preference statements) can dramatically improve the discriminatory power of the analysis. As we will see, enriching the analysis in this way reveals interesting insights. In particular, clusters of countries emerge from the dominance results, and this gives indications about the spatial distribution of welfare across Europe. Finally, the analysis of welfare ranges, made possible by our framework, provides additional valuable insights. As we report below, a picture about welfare discrepancies across Europe emerges that is entirely indiscernible based on dominance analysis alone.

8.1. Data

The European Union Statistics on Income and Living Conditions (EU-SILC) is a widely used data source for empirical analyses. This is a representative survey of households in the European Union countries as well as some countries in the periphery of the EU. EU-SILC monitor changing socio-economic conditions across Europe by measuring a range of economic and social indicators, including demographics, income, taxes and benefits, and labour market status. For most countries

[16] Strictly speaking, sets $\mathcal{U}^1$ and $\mathcal{U}^2$ are normalised such that $u(x_1) = 0$ and $u(x_n) = 100$. However this does not affect the equivalences above, because if v is some positive affine transformation of a non-decreasing (resp. and concave) function u then v is also a non-decreasing (resp. and concave) function and further for any two $t, t' \in T$ $(1, \ldots, 1)\mathbf{v}_t \geq (1, \ldots, 1)\mathbf{v}_{t'} \Leftrightarrow (1, \ldots, 1)\mathbf{u}_t \geq (1, \ldots, 1)\mathbf{u}_{t'}$. Thus for $d = 1, 2$ $\succeq_{\mathcal{U}^d}$ is the same ordering, whether we normalise $\mathcal{U}^d$ or not.

[17] See previous footnote.

considered, we use the cross-sectional survey conducted during 2019. Unfortunately this round did not include Iceland and the UK. However these were included in the 2018 round, so only for these two countries we used data from this earlier round. This practice has been adopted previously (see, e.g., Hussain, 2022), and is justified by the fact that variation from year to year is not typically large (bearing in mind that the data were collected before the start of the COVID-19 pandemic). The age range of individuals included in the sample is 25-81 years. The 31 countries included are listed in Table 1, along with their abbreviated names (two-letter ISO codes).

Country Name	**ISO Code**	**Country Name**	**ISO Code**	**Country Name**	**ISO Code**	**Country Name**	**ISO Code**
Austria	AT	Finland	FI	Latvia	LV	Serbia	RS
Belgium	BE	France	FR	Lithuania	LT	Slovakia	SK
Bulgaria	BG	Germany	DE	Luxembourg	LU	Slovenia	SI
Croatia	HR	Greece	EL	Netherlands	NL	Spain	ES
Cyprus	CY	Hungary	HU	Norway	NO	Sweden	SE
Czechia	CZ	Iceland	IS	Poland	PL	Switzerland	CH
Denmark	DK	Ireland	IE	Portugal	PT	United Kingdom	UK
Estonia	EE	Italy	IT	Romania	RO		

Table 1: Countries included with ISO codes.

The indicator we use for an individual's income is equivalized household disposable income, measured in 1,000 EUR. We calculate this from the EU-SILC variable `hy020`: 'total disposable household income'. This is the sum across all household members of gross personal market and non-market income components minus personal income and wealth taxes, adjusted for household composition. The income is made comparable across countries by transforming to a common currency (EUR) and by taking international differences in purchasing power into account (PPP). To adjust for household size the new OECD equivalence scale is used, where a single adult counts 1 and subsequent individuals aged 14+ count 0.5, while children aged 0-13 years count 0.3, such that the household weight (adult equivalents) is $1 + 0.5\times$(Number of Adults - 1) $+ 0.3\times$(Number of Children).

The indicator we use for health is the individual self-reported health level as represented by the EU-SILC variable `ph010`: 'self-perceived general health'. The possible answers are: very good, good, fair, bad, and very bad. We reversed the ordering and coded the scale numerically by using 0 for 'very bad' and 4 for 'very good' with the other categories coded 1, 2, 3 respectively, in order of improving health. This, five-point 'Likert' scale is categorical in principle, but such scales are commonly treated as interval scales in the literature (e.g., Van Praag, 1991; Ferrer-i-Carbonell and Frijters, 2004; Frey et al., 2009; Powdthavee and Van Den Berg, 2011).

The indicator we use for education is the number of years of an individual's schooling (including university). This is calculated using EU-SILC variable `pe040`: 'highest ISCED level attained' (International Standard Classification of Education). The levels are: less than primary, primary, lower secondary, upper secondary, post-secondary but non-tertiary, short cycle tertiary, bachelors degree, masters degree, and doctorate degree. We convert these to the following corresponding number of years: 2, 5, 10, 13, 14, 15, 16, 18, and 21 years of education.

The initial sample size was 655,841 individuals. We wanted to focus on individuals who had completed education, so we excluded any one below the age of 25, which means 494,110 observations remained. A further 14.2% of observations were excluded due to having an invalid response in one or more of the three welfare indicators. The final number of observations is still a staggering 423,743 respondents. The reduction is nevertheless sizeable and thus risks introducing bias in the analysis. But taking the sample as a whole the bias does not seem to be outspoken, since averages of central variables are not that different when we compared the initial and final samples. The gender distribution in the used sample is still close to 50-50, while average age is slightly lower, average income is slightly higher, average health is almost the same, while average education is a little higher.[18]

The sample of individuals is categorised into a number of groups, defined by the possible combinations of five different levels of each indicator. For health, we use the five levels used in the survey. For income, we use the five quintiles of the distribution of income across the whole sample (i.e. all countries). For education we use five levels represented by 2-5, 10, 13, 14, and 15+ years of education. For each of the three attributes, we use the category average in the original sample, for any individual in that category. The five average levels used for each category are given in Table 2. For brevity, we will refer to the five levels (in each variable) as: 'Lowest', 'Lower', 'Middle', 'Higher', 'Highest'. With five levels for each of three welfare variables the number of possible outcome combinations is $125(= 5^3)$.

	Levels				
	1: Lowest	2: Lower	3: Middle	4: Higher	5: Highest
Income (1000 EUR)	7.295	13.334	18.36112	24.578	42.824
Health (category)	0	1	2	3	4
Education (years)	4.403	10	13	14	15.149

Table 2: Average variable values in each category

This grouping of individuals is not a requirement of the methodology and can in principle be skipped altogether. It does, however, help reduce the size of the optimisation problems solved. The total number of groups (125 in our categorised case) determines the cardinality of set J in our previous formulations, and so determines the size of the linear optimisation problems that need to be solved. For this proof-of-concept application, we used a grouping that could be easily handled on a laptop with a four-core 1.6Ghz processor and 8GB of RAM. All computations used no more than 500 seconds for a full run (i.e. checking for FOD/SOD dominance across all pairs of countries given a set of value judgements) using CPLEX v22.1.1.0 as the optimisation solver. This suggests that much more finely grained groupings could easily be handled even with modestly stronger computational resources.

8.2. Value judgements

We present below the value judgements considered in different scenarios of our analysis. Although we used judgements that seemed defensible in our view, we stress that our intention is not to advocate for these specific judgements. Instead we aim to illustrate the potential impact of using such judgements on the results of the dominance analysis and the derived welfare ranges for all countries. To that end we used judgements to indicate trade-offs that mostly favour improvements in health above other attributes, but we also complemented this with a reversal in preference in specific cases. The full details are given below. In the manner explained in Section 5, each judgement can be incorporated by formulating a number of linear constraints which are then included in the definition of $U(\succeq')$ as given in equation (12).

We do not mention it explicitly to avoid repetition, but all judgements are stated *ceteris paribus*: in any statement, an attribute not mentioned is assumed constant. The term 'increment' refers to

[18]This suggests that any major bias is not seen, although the sample size reduction is very uneven across countries, e.g., only 0.03-0.05% excluded observations for Serbia, Austria, and Cyprus, and as high as 49.4-57.9% excluded observations for Finland, Iceland, and Slovenia. But even for the latter three countries the averages of welfare indicators or demographics such as gender or age are not affected much.

any increase from a lower to the next higher level in some attribute. For example, the phrase 'an increment in income up to middle level' includes both of two possibilities: a) from lowest (7.295) to lower (13.334) and b) from lower to middle (18.36112).

The following two (ordinal) judgements are designed to capture a preference for improvements in health vs any of the other two attributes, except in cases where an individual is already of above middle health but has relatively low income or education respectively.

Judgement 1 (Trade-offs between Health and Income). *An increment in health up to "higher" level is preferred to any increment in income. An increment in income up to "middle" level is preferred to an increment in health from "higher" to "highest" level.*

Judgement 2 (Trade-offs between Health and Education). *An increment in health up to "higher" level is preferred to any increment in education. An increment in education up to "middle" level is preferred to an increment in health from "higher" to "highest" level.*

The following (ordinal) judgement is designed to capture a preference for improvements in income vs education except in cases where income is already at higher level.

Judgement 3 (Trade-offs between Income and Education). *An increment in income up to "higher" level is preferred to any increment in education. An increment in education is preferred to an increment in income from "higher" to "highest" level.*

To illustrate the use of cardinal preference statements, we used the following judgement. It is designed to introduce sub-modularity in the underlying welfare utility function, except that this is imposed locally and only for increments in income up to higher level.

Judgement 4 (Effects of Health increases). *An increment in health leads to a higher increase in individual welfare when income is lower, provided income is no higher than at "middle" level.*

To illustrate the usefulness of SOD and the use of value judgments, we repeated the analysis for different scenarios where we successively refined the set of admissible individual welfare functions. For the baseline we used FOD without any additional value judgements, i.e. assuming only that individual welfare functions are strictly increasing. As a second scenario we considered SOD without any judgements. We also combined SOD with the Judgements 1-4 stated above, adding them one at a time in the order they were stated. This gave four additional scenarios which, for ease of reference, we label by SOD[J1], SOD[J1-2], SOD[J1-3] and SOD[J1-4] respectively.

8.3. Results: dominance identification

The results of the dominance analysis are given in Table 3 and Figure 1. Table 3 shows the number of pairs of countries where the population distribution of one is found to dominate the population distribution of the other, across the different scenarios considered. To get a sense of the strength of the identified quasi-orderings, the percentages (shown in parentheses) can be used. These are calculated based on a theoretical maximum value of 465 dominance instances that would be implied by a full ranking of the 31 countries ($465 = \binom{31}{2}$, i.e. the number of distinct country pairs). Figure 1 gives the full results of the dominance analysis. A positive number in any cell indicates that the population distribution of the country in the corresponding matrix row dominates that of the country in the corresponding column. The values shown indicate the scenario in which a specific dominance instance was first identified. For example, the dominance instances of scenario 1 (FOD) are indicated by 1; the dominances instances of scenario 2 (SOD) are indicated by 1 or 2, with 2 being the additional dominance instances attributed relative to the previous scenario, and

Scenario	Dominance Instances
1: FOD	45 (9.7%)
2: SOD	86 (18.5%)
3: SOD[J1]	132 (28.4%)
4: SOD[J1-2]	239 (51.4%)
5: SOD[J1-3]	410 (88.2%)
6: SOD[J1-4]	426 (91.6%)

Table 3: Dominance Instances (and percentages of maximum possible dominance instances: 465) across Scenarios

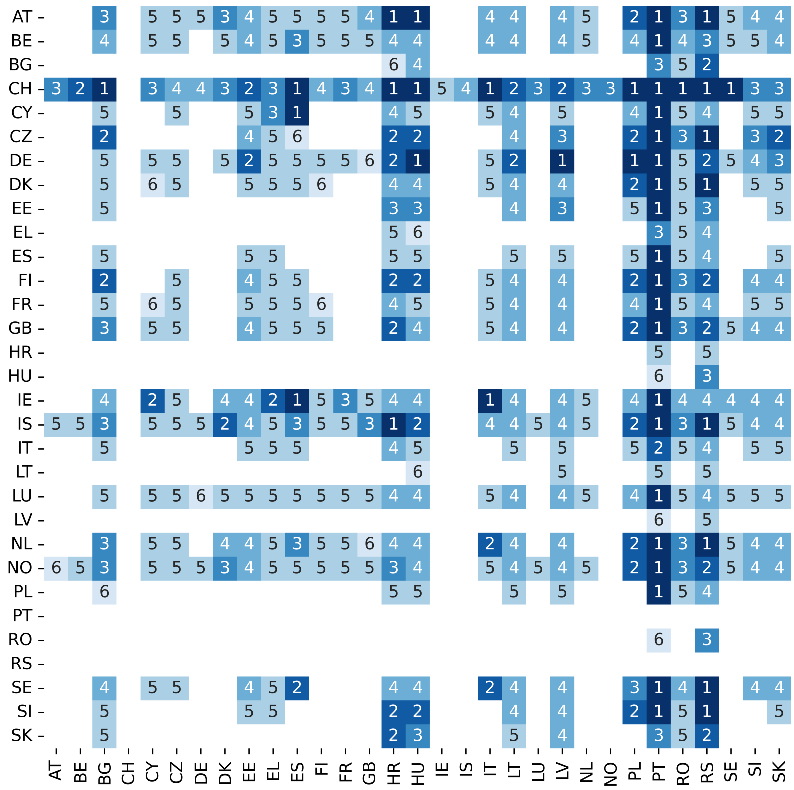

Figure 1: Matrix of dominance results

so on. An empty cell implies that the row country is not found to dominate the column country, but the reverse may be the case if the corresponding cell shows a positive value.

The FOD scenario is the baseline for assessing the value of the SOD approach (with or without judgements) introduced in this paper. There are few FOD instances identified, but of course these would be very robust findings, not contingent on any concavity assumptions or value judgements. Portugal (PT) is dominated by the majority of countries. The Republic of Serbia (RS) is also dominated by a few countries. At the other end, Switzerland (CH) is found to dominate several countries.

The use of a SOD framework and the use of value judgements can improve the comparative strength of the dominance analysis significantly. Just the use of SOD doubles the dominance instances compared to the baseline. From there, the use of judgements can have a dramatic effect, leading to identifying between 28%-92% of the maximum number of dominance instances across the scenarios. Of course the results will, in general, depend on the specific value judgements employed; nevertheless the potential of incorporating value judgments and SOD stands out clearly. Several countries are found to dominate or be dominated by many others only when value judgements are considered. The biggest impact seems to stem from incorporating trade-offs between health and education or income and education (scenarios 4 and 5). Lithuania (LT) and Greece (EL), for example, are not dominated by many countries up to scenario 3 but are dominated by most countries from scenario 4 onwards. Belgium (BE) and Luxembourg (LU) are found to dominate most other countries, but only from scenario 4 onwards. The quasi-ordering of Scenario 6 (SOD[J1-4]) is shown in the graph of Figure 2, in which the nodes indicate countries and the edges indicate dominance in the direction shown by the arrow. Such graphs can be used to define (ordered) clusters of countries based on the dominance instances identified. The top cluster includes countries that are not dominated by any other country. In this case the top cluster only includes Switzerland (CH). Each subsequent cluster includes countries that are dominated only by countries in preceding clusters (though not necessarily all such countries). We have drawn the graph such that countries in the same cluster are on the same vertical level. It is evident that countries around North/Central Europe populate most of the top level clusters, whereas countries around South/East Europe tend to populate the lower level clusters.

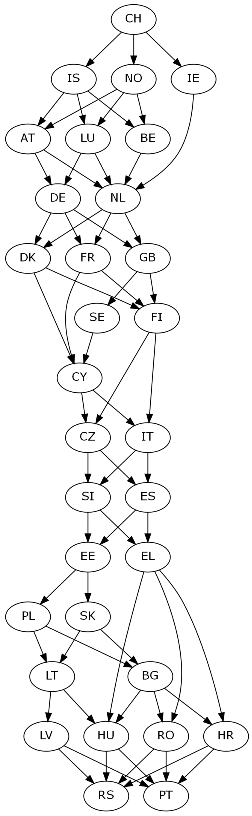

Figure 2: Induced graph of the quasi-ordering of Scenario 6 (SOD[J1-4]).

8.4. Results: welfare ranges

Although dominance analysis helps identify the relative position of countries in terms of social welfare, it does not allow us to identify the degree to which some countries may outperform others, or how all countries perform relative to some benchmark. In Figure 2, we cannot get a sense of how high welfare is in each country. The graph tells us that welfare in Switzerland is deemed higher than welfare in, e.g., Portugal; but it does not tell us if welfare in either country should be deemed high or low. It could well be that it is high for both or high for Switzerland and low for Portugal. We cannot distinguish between these two possibilities based on the dominance structure identified. The missing piece of information required for this is a *scale* on which to measure welfare.

The framework we introduce provides this missing piece of information, thus enriching social welfare evaluation based on dominance analysis. Recall that any individual welfare function can be normalised by fixing the utility levels of two arbitrary outcomes. In our case we set the utility levels of the nadir and ideal points (x_1 and x_n) to 0 and 100 respectively. This defines a 0-100 scale for social welfare (since average utilitarianism simply takes the average of utility levels across individuals). Further, as indicated in Theorem 3, we can calculate lower and upper bounds for social welfare in any country t, for each scenario of value judgements considered.

In Figure 3 we illustrate the potential ranges for social welfare across countries, for three of the scenarios define previously: FOD (scenario 1), SOD (scenario 2), SOD[J1-4] (Scenario 6). Figure B.4 in Appendix B, provides the full results across all of the six scenarios considered. Countries are plotted in increasing order of the computed minimum bounds for average welfare. We have annotated the welfare scale with four intermediate values, defining welfare intervals of length 20, which we refer to as: *low; moderately low; moderate; moderately high; high.* The results very clearly illustrate the value of the introduced approach. The baseline, FOD, scenario provides no meaningful information about social welfare levels: the welfare range spans almost the entire 0-100 scale for most countries. Moving on from that to simple SOD already provides some meaningful bounds, particularly lower bounds: we can see clearly that for about one third of countries average welfare would be at a moderate level or above in the worst case, and almost all countries attain at least a moderately low level of social welfare in the worst case. Moving on to the scenario incorporating all value judgements defined previously, we can see a dramatic reduction of potential welfare ranges for all countries. About a fifth of countries can be seen to comfortably attain a level of high average welfare. Crucially, with one exception, all countries can be seen to attain at least a moderate level of average welfare (and in fact also the Republic of Serbia almost attains this with a minimum of 48). Further most countries attain at least a moderately-high level in the worst case (lower bound). This paints an informative picture for social welfare across Europe, which is completely indiscernible based on investigating dominance instances only. Similarly, we can get a sense of the magnitude of differences in social welfare across countries by looking at the variation in social welfare ranges. For example, the top 6 countries, as ranked by their lower welfare bounds, can be seen to attain a worst case welfare level that is higher than the best case welfare level (upper bound) in the bottom 12 countries, suggesting a prominent gap between these two groups. Such findings can be very useful to policy makers, for example in appraising the outcomes of existing spending/policies designed to improve social welfare in different countries, or evaluating and prioritising future policy initiatives based on needs.

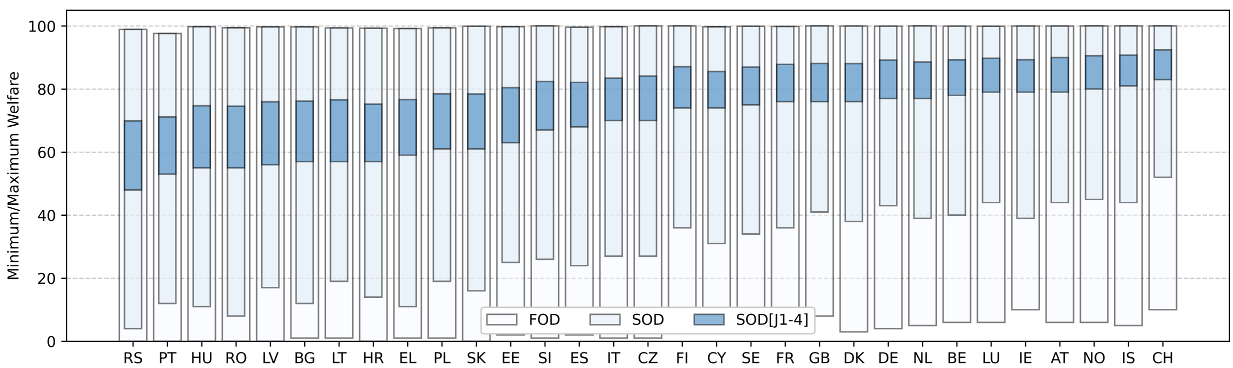

Figure 3: Social welfare ranges across countries for three different scenarios

9. Conclusion

We have introduced a new dominance-based framework to multidimensional social welfare evaluation. Our framework permits, but does not mandate, influencing social welfare evaluation through value judgements. Such judgments may encompass a variety of considerations: about attribute trade-offs, marginal effects or the degree of inequality aversion. This allows for accommodating societal views, or considering their implications on social welfare evaluation. We have derived theoretical results that enable the practical application of our approach. We have shown that the incorporation of such features has the potential to significantly enhance the comparative power of the analysis. As previously noted, the specific results of such analyses may depend on the value judgements considered. Notwithstanding, the effects of value judgements is made transparent and auditable, and the results presented demonstrate the added value of this new form of analysis of welfare ranges – a distinct contribution of our approach. Moreover, the exploration of different counterfactual scenarios of value judgements may be pursued and the implications in terms of social welfare evaluation contrasted. This can inform policy-making by highlighting the connection between specific values and the implied policy priorities. Such a value-driven approach could facilitate consensus-building in policy-making process involving multiple stakeholders, by focusing deliberations on value judgements that demonstrably impact the results of policy appraisal.

We conclude with two suggestions for further research. Firstly, it would be interesting to consider using alternative forms of preference statements in a more extensive application. This could involve direct ratings of different attribute bundles on a welfare scale (e.g. 0-100 as we used in this paper) or employing imprecise evaluations, introducing lower and upper bounds on ratings. Alternatively, stakeholders could be asked to categorise outcomes into predefined welfare scale classes (e.g. splitting the 0-100 scale in five intervals). Secondly, our approach can be used in the context of decision support under risk. Clearly there is a conceptual connection between the problem of comparing population distributions and that of comparing probability distributions over a multidimensional domain. The results we introduced can be directly used to derive a quasi-ordering of risky multidimensional outcomes, as well as refining this via the incorporation of preference statements. The latter could involve direct comparison of certain multidimensional outcomes or lotteries of such outcomes. Therefore the exploration of methodological and practical issues in utilising our results to design decision support frameworks in the risky choice context is another interesting direction for future research.

Acknowledgements

We thank Peter Bogetoft and participants at the EURO 2022 conference in July 2022 for helpful comments on an early version.

Appendix A. Proofs

We assume throughout the below that $\mathcal{X}_i = \mathbb{R}_+ \; \forall i \in I$. This does not affect the generality of our results, because any function u defined over $\mathbb{R}_+$ is also defined on any subset of $\mathbb{R}_+$.

Proof of Theorem 1. We deal separately with two cases for $d = 1$ and $d = 2$. Each will be split in separate parts, as annotated below.

Case 1 ($d = 1$).

(1.1) Let $u \in \mathcal{U}^1$ and define $\upsilon_j = u(x_j) \; \forall j \in J$. Note that by construction we have $\upsilon_1 = 0$, $\upsilon_n = 100$. We will show the existence of $w_{jk} \in \mathbb{R}_+^m$, $\beta_{jk} \in \mathbb{R}$ (for arbitrary $j, k \in J$), such that the corresponding equality and inequality in (10) are satisfied.

(1.1.1) Suppose, first, that $\upsilon_j \geq \upsilon_k$. Set $w_{jk} = \mathbf{0}, \beta_{jk} = \upsilon_j$. Then we have:

$$w_{jk} x_j + \beta_{jk} = \mathbf{0} x_j + \upsilon_j = \upsilon_j,$$
$$w_{jk} x_k + \beta_{jk} = \mathbf{0} x_k + \upsilon_j = \upsilon_j \geq \upsilon_k,$$

i.e. the required inequalities in (10) are satisfied.

(1.1.2) Suppose that $v_j < v_k$. Suppose, further, that $x_j \geq x_k$. Then from non-decreasingness of the utility function u we obtain $v_j = u(x_j) \geq u(x_k) = v_k$ which contradicts the supposition. Therefore $\exists i' \in I : x_j^{i'} < x_k^{i'}$.

Since we have $v_k - v_j > 0$ (by supposition), there must exist a sufficiently small $\varepsilon > 0$ such that $\varepsilon(v_k - v_j) > 0$ is sufficiently small so that we have:

$$\begin{aligned} x_j^{i'} &< x_k^{i'} - \varepsilon(v_k - v_j) \\ \Rightarrow \varepsilon v_k &< x_k^{i'} - x_j^{i'} + \varepsilon v_j \\ \Rightarrow v_k &< \frac{1}{\varepsilon} x_k^{i'} + (v_j - \frac{1}{\varepsilon} x_j^{i'}), \end{aligned}$$

so that by setting $w_{jk}^{i'} = \frac{1}{\varepsilon}, w_{jk}^{i} = 0 \ \forall i \in I, \beta_{jk} = (v_j - \frac{1}{\varepsilon} x_j^{i'})$, we directly obtain that $v_k < w_{jk} x_k + \beta_{jk}$ and also $w_{jk} x_j + \beta_{jk} = \frac{1}{\varepsilon} x_k^{i'} + v_j - \frac{1}{\varepsilon} x_j^{i'} = v_j$, i.e. that the respective inequalities in (10).

(1.2) Let $v \in U^1$. We need to show the existence of a $u \in \mathcal{U}^1 : u(x_j) = v_j \ \forall j \in J$.

For any $x \in \mathbb{R}_+^m$ define $J^{\leq}(x) = \{j \in J \mid x_j \leq x\}$ and $u(x) = \max_{j \in J^{\leq}(x)}\{v_j\}$.

(1.2.1) To show that u is non-decreasing, consider $x, x' \in \mathbb{R}_+^m : x' \geq x$.

Directly from the definition of u, we obtain that $\exists j \in J : u(x) = v_j$ as well as that: $x_j \leq x$. Then we have: $x' \geq x \Rightarrow j \in J^{\leq}(x') \Rightarrow u(x') \geq v_j = u(x)$. Since also $u(x_1) = 0$ and $u(x_n) = 100$ by construction, we have $u \in \mathcal{U}^1$.

(1.2.2) To show that $u(x_j) = v_j \ \forall j \in J$, consider an arbitrary $j \in J$. By the definition of u we obtain that $\exists k \in J : v_k = u(x_j)$ and also that $x_k \leq x_j$. In the trivial case that $k = j$ we obtain the requirement directly: $u(x_j) = v_k = v_j$. So it remains to consider the case where $k \neq j$.

Since it holds, trivially, that $j \in J^{\leq}(x_j)$, we directly obtain $u(x_j) \geq v_j$.

Let w_{jk} and β_{jk} be an arbitrarily chosen set of constants, $j, k \in J$, that satisfy the conditions of (10). From the corresponding inequality in (10) for the pair (j, k) combined with $x_k \leq x_j$ (as noted above) we obtain:

$$\begin{aligned} v_k \leq w_{jk} x_k + \beta_{jk} \leq w_{jk} x_j + \beta_{jk} &= v_j \\ \Rightarrow u(x_j) &\leq v_j. \end{aligned}$$

Combining the above gives $u(x_j) = v_j$ as required.

Case 2 ($d = 2$)**.**

(2.1) Let $u \in \mathcal{U}^2$. Define $v_j = u(x_j) \forall j \in J$. Note that $v_1 = 0$ and $v_n = 100$ by construction. Now consider arbitrary $j \in J$. We will show the existence of $w_j \in \mathbb{R}_+^m, \beta_j \in \mathbb{R}$ such that the corresponding equalities and inequalities in (11) are satisfied.

(2.1.1). Suppose that $\exists a_j = (a_j^0, a_j^1, \ldots, a_j^m)$ with $a_j^0 > 0$ and $(a_j^1, \ldots, a_j^m) \leq \mathbf{0}$ such that $a_j(v_j, x_j) \geq a_j(v_k, x_k) \forall k \in J$. [19] Then we may define $\beta_j = \frac{a_j(v_j, x_j)}{a_j^0}$ and $w_j = (-\frac{a_j^1}{a_j^0}, \ldots, -\frac{a_j^m}{a_j^0})$. From these we may obtain the desired equality as follows:

$$\begin{aligned} \beta_j &= \frac{a_j(v_j, x_j)}{a_j^0} \\ &= \frac{a_j^0}{a_j^0} v_j + (\frac{a_j^1}{a_j^0}, ..., \frac{a_j^m}{a_j^0}) x_j \\ \Rightarrow v_j &= w_j x_j + \beta_j, \end{aligned}$$

[19] The notation $a_j(v_j, x_j)$ means the inner product of a and (v_j, x_j), i.e. $a(v_j, x_j) = (a_j^0, a_j^1, \ldots, a_j^m) \cdot (v_j, x_j^1, \ldots, x_j^m) = a_j^0 v_j + a_j^1 x_j^1 + \ldots + a_j^m x_j^m$.

as well as the desired inequality (for arbitrary $k \in J$) as follows:

$$
\begin{aligned}
a_j(\upsilon_j, x_j) &\geq a_j(\upsilon_k, x_k) \\
\Rightarrow \frac{a_j(\upsilon_j, x_j)}{a_j^0} &\geq \frac{a_j(\upsilon_k, x_k)}{a_j^0} \\
\Rightarrow \beta_j &\geq \frac{a_j^0}{a_j^0}\upsilon_k + (\frac{a_j^1}{a_j^0}, ..., \frac{a_j^m}{a_j^0})x_k \\
\Rightarrow \upsilon_k &\leq w_j x_k + \beta_j.
\end{aligned}
$$

Based on (2.1.1), we will complete the proof of (2.1) by showing that: for any $j \in J$, $\exists a_j = (a_j^0, a_j^1, ..., a_j^m)$ with $a_j^0 > 0$ and $(a_j^1, ..., a_j^m) \leq \mathbf{0}$ such that $a_j(\upsilon_j, x_j) \geq a_j(\upsilon_k, x_k) \forall k \in J$.

(2.1.2) Define $H = \{(\upsilon, x) \in \mathbb{R}_+^{m+1} | \upsilon \leq u(x)\}$, the hypograph of u. Consider an arbitrary $j \in J$. As an intermediary step we will first show that $\exists h_j = (h_j^0, h_j^1, ..., h_j^m) \neq \mathbf{0}$ with $h_j^0 \geq 0$ and $(h_j^1, ..., h_j^m) \leq \mathbf{0}:\ h_j(\upsilon_j, x_j) \geq h_j(\upsilon, x)\ \forall\ (\upsilon, x) \in H$.

Since u is a concave function it follows that H is a convex set. Further, point (υ_j, x_j) is a boundary point of H (if not then for some small $\varepsilon > 0$ we have $(\upsilon_j + \varepsilon, x_j) \in H$ which, by the definition of H, leads to the contradiction $\upsilon_j < \upsilon_j + \varepsilon \leq u(x_j) = \upsilon_j$).

From the above we obtain by the Supporting Hyperplane Theorem (e.g. Corollary 11.6.1 in Rockafellar, 1970) that there exists a supporting hyperplane at point (υ_j, x_j) of H, namely that $\exists h_j = (h_j^0, h_j^1, ..., h_j^m) \neq \mathbf{0}:\ h_j(\upsilon_j, x_j) \geq h_j(\upsilon, x)\ \forall\ (\upsilon, x) \in H$.

Further, we can show that $h_j^0 \geq 0$ and $(h_j^1, h_j^2, ..., h_j^m) \leq \mathbf{0}$.

For the former, let $\varepsilon > 0$. Then $\upsilon_j \geq \upsilon_j - \varepsilon \Rightarrow (\upsilon_j - \varepsilon, x_j) \in H \Rightarrow h_j^0 \upsilon_j + \sum_{i \in I} h_j^i x_j^i \geq h_j^0 \upsilon_j - h_j^0 \varepsilon + \sum_{i \in I} h_j^i x_j^i \Rightarrow h_j^0 \varepsilon \geq 0 \Rightarrow h_j^0 \geq 0$.

Similarly, for arbitrary $i' \in I$ let $\mathbf{1}_{i'} \in \mathbb{R}^m$ be a vector with 1 in the i'-th position and 0 elsewhere and define $x'_j = x_j + \mathbf{1}_{i'}\varepsilon$. Since u is non-decreasing we have $x'_j \geq x_j \Rightarrow u(x'_j) \geq u(x_j) \Rightarrow (u(x_j), x'_j) \in H \Rightarrow h_j(u(x_j), x_j) \geq h_j(u(x_j), x'_j) \Rightarrow h_j^{i'} x_j \geq h_j^{i'}(x_j + \varepsilon) \Rightarrow h_j^{i'} \varepsilon \leq 0 \Rightarrow h_j^{i'} \leq 0$.

We now return to showing the existence of a desired a_j as stated previously. We will do this separately for three cases: (1) $x_j > \mathbf{0}$, (2) $x_j \not> \mathbf{0}$ and $x_j \neq \mathbf{0}$, (3) $x_j = \mathbf{0}$. To avoid cumbersome notation, onwards from this point onwards we will drop the index j and just write a and h instead of a_j and h_j respectively, but it should be borne in mind that a and h are defined for a specific (though arbitrary) j in remainder of this proof.

(**2.1.3**) Suppose $x_j > \mathbf{0}$. Set $a = h$. This means that $(a^1, ..., a^m) \leq \mathbf{0}$ and $a^0 \geq 0$. It remains to show that $a^0 \neq 0$. By contradiction, assume that $a^0 = 0$. From the preceding, $(u(\mathbf{0}), \mathbf{0}) \in H \Rightarrow a(\upsilon_j, x_j) \geq a(u(\mathbf{0}), \mathbf{0}) \Rightarrow 0\upsilon_j + (a^1, ..., a^m)x_j \geq a^0 u(\mathbf{0}) + (a^1, ..., a^m)\mathbf{0} \Rightarrow (a^1, ..., a^m)x_j \geq 0$. As shown before $(a^1, ..., a^m) \leq \mathbf{0}$ and since we have assumed $x_j > \mathbf{0}$ we have $(a^1, ..., a^m)x_j \leq 0$. Combining gives $(a^1, ..., a^m)x_j = 0$ and since we have assumed $x_j > \mathbf{0}$ we obtain $(a^1, ..., a^m) = \mathbf{0}$. Now if $a^0 = 0$ this would mean that $a = (a^0, a^1, ..., a^m) = \mathbf{0}$ which is a contradiction (recall that the supporting hyperplane at point (υ_j, x_j) is non-trivial, i.e. $a = h \neq \mathbf{0}$). Thus $a^0 \neq 0 \Rightarrow a^0 > 0$.

(**2.1.4**) Suppose that $x_j \not> \mathbf{0}$ and $x_j \neq \mathbf{0}$, We may assume for convenience that $\exists \bar{m} < m : x_j^i > 0\ \forall i \in \{1, ..., \bar{m}\}$, $x_j^i = 0\ \forall i \in \{\bar{m}+1, ..., m\}$. Accordingly we may partition any $x \in \mathbb{R}_+^m$ as follows: $x = (\bar{x}, x^{\bar{m}+1}, ..., x^m)$. Note that $x_j = (\bar{x}_j, 0, ..., 0)$.

Define $\bar{u}(\bar{x})) = u(\bar{x}, 0, ..., 0)$. For $\bar{y} \geq \bar{x}$ we have $(\bar{y}, 0, ..., 0) \geq (\bar{x}, 0, ..., 0)$, therefore $\bar{u}(\bar{y}) = u(\bar{y}, 0, ..., 0) \geq u(\bar{x}, 0..., 0) = \bar{u}(\bar{x})$, i.e. $\bar{u}$ is non-decreasing. Further, for $\lambda \in [0, 1]$ we have $\bar{u}(\lambda\bar{x} + (1-\lambda)\bar{y}) = u(\lambda\bar{x} + (1-\lambda)\bar{y}, 0, ..., 0) = u(\lambda(\bar{x}, 0, ..., 0) + (1-\lambda)(\bar{y}, 0, ..., 0)) \geq \lambda u(\bar{x}, 0, ..., 0) + (1-\lambda)u(\bar{y}, 0, ..., 0) = \lambda\bar{u}(\bar{x}) + (1-\lambda)\bar{u}(\bar{y})$, i.e. $\bar{u}$ is concave.

Then from the preceding (cf 2.1.2 and 2.1.3) we know that for any $j \in J : \exists \bar{a} = (a^0, a^1, ..., a^{\bar{m}})$ with $a^0 > 0$ and $(a^1, ..., a^{\bar{m}}) \leq \mathbf{0}$ such that $\bar{a}(\bar{u}(\bar{x}_j), \bar{x}_j) \geq \bar{a}(\bar{u}(\bar{x}_k), \bar{x}_k) \forall k \in J$.

Since $x_j = (\bar{x}_j, 0, ..., 0)$ we have $\bar{u}(\bar{x}_j) = u(x_j) = \upsilon_j$. Further, note that for any $x \in \mathbb{R}_+^m :\ x \geq (\bar{x}, 0, ..., 0) \in \mathbb{R}_+^m \Rightarrow\ u(x) \geq u(\bar{x}, 0, .., 0))$.

Now consider an arbitrary $k \in J$, let $E > 0$, and define $\epsilon = u(x_k) - u(\bar{x}_k, 0, ..., 0)$ and $a =$

$(\bar{a}, -E, ..., -E)$. Then we have:

$$\begin{aligned}
a(u(x_k), x_k) &= (\bar{a}, -E, ..., -E)(u(x_k), x_k) \\
&= (\bar{a}, -E, ..., -E)(u(\bar{x}_k, 0, ..., 0) + \epsilon, \bar{x}_k, x_k^{\bar{m}+1}, ..., x_k^m) \\
&= \bar{a}(u(\bar{x}_k, 0, ..., 0), \bar{x}_k) + a^0\epsilon - E(\mathbf{1}(x_k^{\bar{m}+1}, ..., x_k^m)) \\
&= \bar{a}(\bar{u}(\bar{x}_k), \bar{x}_k) + a^0\epsilon - E(\mathbf{1}(x_k^{\bar{m}+1}, ..., x_k^m)) \\
&\leq \bar{a}(\bar{u}(\bar{x}_j), \bar{x}_j) + a^0\epsilon - E(\mathbf{1}(x_k^{\bar{m}+1}, ..., x_k^m)) \\
&= \bar{a}(\bar{u}(\bar{x}_j), \bar{x}_j) + (-E, ..., -E)(0, ..., 0) + a^0\epsilon - E(\mathbf{1}(x_k^{\bar{m}+1}, ..., x_k^m)) \\
&= (\bar{a}, -E, ..., -E)(\bar{u}(\bar{x}_j), \bar{x}_j, 0..., 0) + a^0\epsilon - E(\mathbf{1}(x_k^{\bar{m}+1}, ..., x_k^m)) \\
&= a(\bar{u}(\bar{x}_j), \bar{x}_j, 0..., 0) + a^0\epsilon - E(\mathbf{1}(x_k^{\bar{m}+1}, ..., x_k^m)) \\
&= a(u(x_j), x_j) + a^0\epsilon - E(\mathbf{1}(x_k^{\bar{m}+1}, ..., x_k^m)).
\end{aligned}$$

Note that $(x_k^{\bar{m}+1}, ..., x_k^m) \geq \mathbf{0}$. Now consider the case where $(x_k^{\bar{m}+1}, ..., x_k^m) \neq \mathbf{0}$. Then for $E > 0$ we have $E(\mathbf{1}(x_k^{\bar{m}+1}, ..., x_k^m)) > 0$. But since $E > 0$ is freely chosen, we can chose a value arbitrarily large so that $E(\mathbf{1}(x_k^{\bar{m}+1}, ..., x_k^m) \geq a^0\epsilon$ no matter what the values $a^0, x_k^{\bar{m}+1}, ..., x_k^m$ may be (and for any k). Then we would have $a(u(x_j), x_j) \geq a(u(x_j), x_j) + a^0\epsilon - E(\mathbf{1}(x_k^{\bar{m}+1}, ..., x_k^m)) \geq a(u(x_k), x_k)$ as required. On the other hand when $(x_k^{\bar{m}+1}, ..., x_k^m) = \mathbf{0}$ then $E(\mathbf{1}(x_k^{\bar{m}+1}, ..., x_k^m)) = 0$. Further $x_k = (\bar{x}_k, 0, ..., 0) \Rightarrow \bar{u}(\bar{x}_k) = u(x_k) \Rightarrow \epsilon = 0$ so the above directly gives $a(u(x_j), x_j) \geq a(u(x_k), x_k)$ as required.

(**2.1.5**) Suppose that $x_j = \mathbf{0}$. Then $x_k \neq \mathbf{0}$ for any $k \neq j$ since all outcomes are distinct. Recall that $u(\mathbf{0}) = 0$ by assumption and also that $h^0 \geq \mathbf{0}$ and $(h^1, ..., h^m) \leq \mathbf{0}$. The case where $h^0 > 0$ is trivial as we may then set $a = h$ and obtain the desired result directly. Consider the case where $h^0 = 0$. Since h defines a hyperplane it follows that $(h^1, ..., h^m) \neq \mathbf{0}$. Combining the above we have $(h^1, ..., h^m)(x_k^1, ..., x_k^m) < 0$ for any $k \neq j$. Then there must exist a positive but small enough $\varepsilon > 0 :\ (h^1, ..., h^m)(x_k^1, ..., x_k^m) + \varepsilon u(x_k) < 0$, for any $k \in J$ (recall $u(x_k) \geq u(\mathbf{0}) = 0$ by assumption). Set $a = (\varepsilon, h^1, ..., h^m)$. Then we may obtain (for arbitrary k) the desired inequality as follows:

$$\begin{aligned}
a(\upsilon_j, x_j) &= (\varepsilon, h^1, ..., h^m)(\upsilon_j, x_j) \\
&= (\varepsilon, h^1, ..., h^m)(0, \mathbf{0}) \\
&= 0 \\
&> \varepsilon u(x_k) + (h^1, ..., h^m)(x_k^1, ..., x_k^m) \\
&= (\varepsilon, h^1, ..., h^m)(\upsilon_k, x_k) \\
&= a(\upsilon_k, x_k).
\end{aligned}$$

(2.2) Let $\upsilon \in U^2$. We need to show the existence of $u \in \mathcal{U}^2 : u(x_j) = \upsilon_j\ \forall j \in J$. Let w_j and β_j be an arbitrarily chosen set of constants, $j \in J$, that satisfy the conditions of (11). Define a function u as $u(x) = \min_{j\in J}\{w_j x + \beta_j\}$. We will show that u is concave, non-decreasing, and that $u(x_j) = \upsilon_j \forall j \in J$.

Since u is defined as the minimum of a finite number of affine (thus concave) functions $w_j x + \beta_j$ it follows that u is also concave.

Now consider $x, y \in \mathbb{R}_+^m$ with $x \geq y$. Then we have: $u(y) = \min_{j\in J}\{w_j y + \beta_j\} \leq \min_{j\in J}\{w_j x + \beta_j\} = u(x)$. This shows that u is non-decreasing.

Finally consider arbitrary $k \in J$. Then we have $\upsilon_k = w_k x_k + \beta_k \geq \min_{j\in J}\{w_j x_k + \beta_j\} = u(x_k)$. At the same time we have $\upsilon_k \leq w_j x_k + \beta_j\ \forall j \in J \Rightarrow \upsilon_k \leq \min_{j\in J}\{w_j x_k + \beta_j\} = u(x_k)$. This gives $u(x_k) = \upsilon_k \forall k \in J$. ■

Proof of Corollary 1. It is easy to see that $\upsilon \in U(\succeq') \Leftrightarrow \upsilon \in \mathcal{U}_{[X]}(\succeq')$. Combining with Theorem 1 we have (for $d = 1, 2$): $\upsilon \in U^d(\succeq) \Leftrightarrow \left(\upsilon \in U^d\right) \wedge (\upsilon \in U(\succeq')) \Leftrightarrow \left(\upsilon \in \mathcal{U}^d_{[X]}\right) \wedge \left(\upsilon \in \mathcal{U}_{[X]}(\succeq')\right) \Leftrightarrow \upsilon \in \mathcal{U}^d_{[X]}(\succeq')$. ■

Proof of Theorem 2. $p' \trianglerighteq^d p'' \Leftrightarrow \sum_{j\in J} p'_j u(x_j) \geq \sum_{j \in J} p''_j u(x_j)\ \forall u(\cdot) \in \mathcal{U}^d(\succeq') \Leftrightarrow \min\{\sum_{j\in J} p'_j u(x_j) - \sum_{j\in J} p''_j u(x_j)|\ u(\cdot) \in \mathcal{U}^d(\succeq')\} \geq 0 \Leftrightarrow \min\{\sum_{j\in J} p'_j \upsilon_j - \sum_{j\in J} p''_j \upsilon_j|\ \upsilon \in \mathcal{U}^d_{[X]}(\succeq')\} \geq 0 \Leftrightarrow \min\{\sum_{j\in J} p'_j \upsilon_j - \sum_{j\in J} p''_j \upsilon_j|\ \upsilon \in U^d(\succeq')\} \geq 0$. ■

Proof of Theorem 3. This is obtained immediately from an application of Corollary 1, to replace $\mathcal{U}^d_{[X]}(\succeq')$ with $U^d(\succeq')$ (and vice versa). ■

Proof of Theorem 4.
Case 1 ($d = 1$).

Let $u \in \overline{\mathcal{U}}^1$. Define $\upsilon_j = u(x_j) \forall j \in J$. As in the proof of Theorem 1, note that $\upsilon_1 = 0$ and $\upsilon_n = 100$.

To show that $\upsilon \in \overline{U}^1$, it suffices to show the existence of $w_{jk} \in \mathbb{R}^m_{++}$ and $\beta_{jk} \in \mathbb{R}$ such that: (A) $\upsilon_j = w_{jk}x_j + \beta_{jk}$ and (B) $\upsilon_k \leq w_{jk}x_k + \beta_{jk}$, $\forall j, k \in J$.

Note that if u is strictly increasing then it is also non-decreasing, i.e. $u \in \mathcal{U}^1$. Then, by Theorem 1, $\upsilon \in U^1$ and so for any $j, k \in J$, there exists $\bar{w}_{jk} \in \mathbb{R}^m_+$, $\bar{\beta}_{jk} \in \mathbb{R}$: (A') $\upsilon_j = \bar{w}_{jk}x_j + \bar{\beta}_{jk}$ and (B') $\upsilon_k \leq \bar{w}_{jk}x_k + \bar{\beta}_{jk}$.

Let $\epsilon^i_{jk} > 0\ \forall i \in I$. Define $w^i_{jk} = \bar{w}^i_{jk} + \epsilon^i_{jk}, \forall i \in I$, $\beta_{jk} = \bar{\beta}_{jk} - \sum_i \epsilon^i_{jk}x^i_j$. Then, using (A'), we obtain:

$$w_{jk}x_j + \beta_{jk} = \sum_i (\bar{w}^i_{jk} + \epsilon^i_{jk})x^i_j + (\bar{\beta}_{jk} - \sum_i \epsilon^i_{jk}x^i_j) = \bar{w}_{jk}x_j + \bar{\beta}_{jk} = \upsilon_j$$

which establishes (A). For (B) to hold we require:

$$\begin{aligned} &\upsilon_k \leq w_{jk}x_k + \beta_{jk} \\ \Leftrightarrow\ &\upsilon_k \leq \bar{w}_{jk}x_k + \sum_i \epsilon^i_{jk}x^i_k + \bar{\beta}_{jk} - \sum_i \epsilon^i_{jk}x^i_j \\ \Leftrightarrow\ &\upsilon_k - (\bar{w}_{jk}x_k + \bar{\beta}_{jk}) \leq \sum_i \epsilon^i_{jk}(x^i_k - x^i_j). \end{aligned} \tag{A.1}$$

Now recall that the ϵ^i_{jk} values were restricted to be positive but otherwise freely chosen, and we can show that they can be chosen to ensure that the inequality (A.1), which is equivalent to (B), holds. We consider two cases.

Firstly, suppose that $x^i_k \leq x^i_j\ \forall i \in I$. Then because u is strictly increasing we obtain $\upsilon_k = u(x_k) < u(x_j) = \upsilon_j$. Further, it is easy to see that (A') and (B') above can be satisfied by setting $\bar{w}_{jk} = \mathbf{0}$ and $\bar{\beta}_{jk} = \upsilon_j$. Thereby the left hand side of the inequality in (A.1) becomes $\upsilon_k - (\bar{w}_{jk}x_k + \bar{\beta}_{jk}) = \upsilon_k - \upsilon_j < 0$. From our assumptions $x^i_k < x^i_j$ and $\epsilon^i_{jk} > 0\ \forall i \in I$ we obtain that the right hand side of the inequality in (A.1) is also negative, i.e. $\sum_i \epsilon^i_{jk}(x^i_k - x^i_j) < 0$. Recall however that all ϵ^i_{jk} values are positive but otherwise freely chosen. This means we can chose them to be small enough so that $\sum_i \epsilon^i_{jk}(x^i_k - x^i_j) < 0$ is arbitrarily close to 0, i.e. larger than some other given negative value, which ensures that the inequality in (A.1) is satisfied.

Similarly, for the second case, where $x^{i'}_k - x^{i'}_j > 0$ for some $i' \in I$, we can choose the value $\epsilon^{i'}_{jk}$ to be sufficiently larger than all other ϵ^i_{jk} values so as to ensure that $\sum_i \epsilon^i_{jk}(x^i_k - x^i_j) > 0$ is positive. Then, combining with (B') we obtain $\upsilon_k - (\bar{w}_{jk}x_k + \bar{\beta}_{jk}) \leq 0 < \sum_i \epsilon^i_{jk}(x^i_k - x^i_j) > 0$, i.e. that condition (A.1) is satisfied.

Now let $\upsilon \in \overline{U}^1$. We need to show that $\upsilon \in \overline{\mathcal{U}}^1_{[X]}$, i.e. that there exists a strictly increasing function u such that $u(x_j) = \upsilon_j\ \forall j \in J$.

Define the function u as follows:

$$u(x) = \max_{j \in J^{\leq}(x)} \{\upsilon_j + \epsilon(\sum_i x^i - x^i_j)\},$$

where $\epsilon > 0$ is an arbitrarily chosen constant. Note that, by definition, $j \in J^{\leq}(x) \Rightarrow \epsilon(\sum_i x^i - x^i_j) \geq 0 \Rightarrow \upsilon_j + \epsilon(\sum_i x^i - x^i_j) \geq \upsilon_j$

To show that u is strictly increasing consider $y, z \in \mathbb{R}^m_+ : y \geq z, \ y \neq z$. Then $\forall j \in J : \upsilon_j + \epsilon(\sum_i y^i - x^i_j) > \upsilon_j + \epsilon(\sum_i z^i - x^i_j)$ and $J^{\leq}(y) \supseteq J^{\leq}(z)$. Combining we obtain:

$$\begin{aligned} u(y) &= \max_{j \in J^{\leq}(y)} \{\upsilon_j + \epsilon(\sum_i y^i - x^i_j)\} \\ &\geq \max_{j \in J^{\leq}(z)} \{\upsilon_j + \epsilon(\sum_i y^i - x^i_j)\} \\ &> \max_{j \in J^{\leq}(z)} \{\upsilon_j + \epsilon(\sum_i z^i - x^i_j)\} = u(z). \end{aligned}$$

To show that u reproduces the utility assignment υ, consider $r \in J$. We need to show that $\upsilon_r = u(x_r)$. We do this by showing that $\upsilon_r \leq u(x_r)$ and $\upsilon_r \geq u(x_r)$.

Since $r \in J^{\leq}(x_r)$ we obtain:

$$\upsilon_r \leq \max_{j \in J^{\leq}(x_j)} \{\upsilon_j\} \leq \max_{j \in J^{\leq}(x_j)} \{\upsilon_j + \epsilon(\sum_i x^i_r - x^i_j)\} = u(x_r).$$

Further, consider $k' = \arg\max_{j \in J^{\leq}(x_r)} \{\upsilon_j + \epsilon(\sum_i x^i_r - x^i_j)\} \in J$, so that $u(x_r) = \upsilon_{k'}$. Note also that $k' \in J^{\leq}(x_r) \Rightarrow x_{k'} \leq x_r$. Let $w_{rk'}$ and $\beta_{rk'}$ be an arbitrary chosen set of constants, $r, k' \in J$, that satisfy the conditions of (10). Then we obtain:

$$u(x_r) = \upsilon_{k'} \leq w_{rk'} x_{k'} + \beta_{rk'} \leq w_{rk'} x_r + \beta_{rk'} = \upsilon_r.$$

Case 2 ($d = 2$)

Let $u \in \overline{\mathcal{U}}^2$ and define $\upsilon_j = u(x_j) \ \forall \ j \in J$. To show that $\upsilon \in \overline{U}^2$, we need to show the existence of $w_j \in \mathbb{R}^m_{++}, \ \beta_j \in \mathbb{R} \ \forall j \in J$ such that the conditions in (11) are satisfied.

Since $\overline{\mathcal{U}}^2 \subset \mathcal{U}^2$ we know from Theorem 1 that $\upsilon \in U^2$ and so $\exists w_j \in \mathbb{R}^m_+, \ \beta_j \in \mathbb{R} \ \forall j \in J$ such that the conditions in (11) are satisfied. It thus remains to show that $w_j > \mathbf{0}$.

Suppose that $w^{i'}_j = 0$ for some $i' \in I$. Consider an arbitrary $j \in J$. From the proof of Theorem (1), we know that the w_j defines a super-gradient of the concave function u at point x_j, so that we have $u(x) \leq w_j x + \beta_j \ \forall x \in \mathbb{R}^m_+$. Let $\epsilon > 0$ and suppose $x^{i'} = x^{i'}_j + \epsilon, \ x^i = x^i_j \ \forall i \in I : i = i'$. From the above we have: $u(x) \leq w_j x + \beta_j = w_j x_j + w^{i'}_j x^{i'}_j + \beta_j = w_j x_j + \beta_j = \upsilon_j$. However since $x > x_j$ and u is non-decreasing we also have obtain $u(x) > u(x_j) = \upsilon_j$, which is a contradiction. Therefore we must have that $w_j > \mathbf{0}$.

Now let $\upsilon \in \overline{U}^2$. We need to show that $\upsilon \in \overline{\mathcal{U}}^2_{[X]}$, i.e. the existence of a strictly increasing and concave function u that reproduces υ. We may define $u(x) = \min_j \{w_j x + \beta_j\}$ exactly as in the proof of Theorem 1. Then the parts of that proof showing that u is concave and reproduces υ apply here directly. Finally the corresponding part of that proof can be modified slightly to show that u is strictly increasing, by assuming further that $x \neq y$ so that by also taking into account that $w_j > \mathbf{0}$. ∎

Appendix B. Additional Figure

In Figure B.4 we provide the potential ranges for social welfare across countries, for all six scenarios considered.

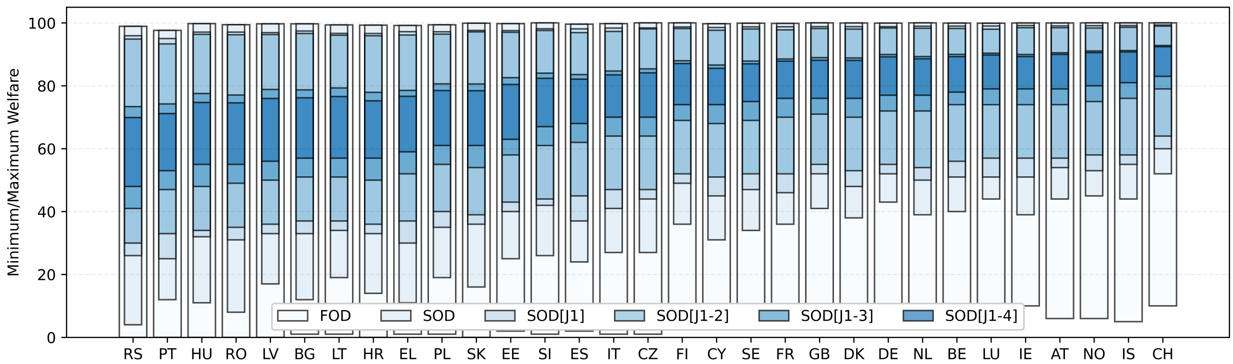

Figure B.4: Social welfare ranges across countries for all scenarios considered

References

Alkire, S., Foster, J., 2011. Counting and multidimensional poverty measurement. Journal of Public Economics 95, 476–487.

Anderson, G., Crawford, I., Leicester, A., 2011. Welfare rankings from multivariate data, a nonparametric approach. Journal of Public Economics 95, 247–252.

Angilella, S., Catalfo, P., Corrente, S., Giarlotta, A., Greco, S., Rizzo, M., 2018. Robust sustainable development assessment with composite indices aggregating interacting dimensions: The Hierarchical-SMAA-Choquet integral approach. Knowledge-Based Systems 158, 136–153.

Angilella, S., Corrente, S., Greco, S., Słowiński, R., 2016. Robust ordinal regression and stochastic multiobjective acceptability analysis in multiple criteria hierarchy process for the choquet integral preference model. Omega 63, 154–169.

Argyris, N., Morton, A., Figueira, J.R., 2014. CUT: A multicriteria approach for concavifiable preferences. Operations Research 62, 633–642.

Arndt, C., Distante, R., Hussain, M.A., Østerdal, L.P., Huong, P.L., Ibraimo, M., 2012. Ordinal welfare comparisons with multiple discrete indicators: A first order dominance approach and application to child poverty. World Development 40, 2290–2301.

Arndt, C., Hussain, A.M., Salvucci, V., Tarp, F., Østerdal, L.P., 2016. Poverty mapping based on first-order dominance with an example from Mozambique. Journal of International Development 28, 3–21.

Athanassoglou, S., 2015. Multidimensional welfare rankings under weight imprecision: a social choice perspective. Social Choice and Welfare 44, 719–744.

Atkinson, A.B., 1970. On the measurement of inequality. Journal of Economic Theory 2, 244–263.

Atkinson, A.B., Bourguignon, F., 1982. The comparison of multi-dimensioned distributions of economic status. The Review of Economic Studies 49, 183–201.

Blackorby, C., Bossert, W., Donaldson, D., 2002. Chapter 11 utilitarianism and the theory of justice, in: Handbook of Social Choice and Welfare. Elsevier. volume 1 of *Handbook of Social Choice and Welfare*, pp. 543–596.

Blackorby, C., Donaldson, D., 1984. Social criteria for evaluating population change. Journal of Public Economics 25, 13–33.

Bourguignon, F., 1989. Family size and social utility: Income distribution dominance criteria. Journal of Econometrics 42, 67–80.

Cherchye, L., Moesen, W., Rogge, N., Puyenbroeck, T.V., 2007. An introduction to 'benefit of the doubt'composite indicators. Social indicators research 82, 111–145.

Cinelli, M., Coles, S.R., Kirwan, K., 2014. Analysis of the potentials of multi criteria decision analysis methods to conduct sustainability assessment. Ecological indicators 46, 138–148.

Cinelli, M., Kadziński, M., Gonzalez, M., Słowiński, R., 2020. How to support the application of multiple criteria decision analysis? let us start with a comprehensive taxonomy. Omega 96, 102261.

Cinelli, M., Koffler, C., Askham, C., Amadei, A., Arendt, R., Bachmann, T., Barros, B., Bjørn, A., Dias, L., Laurent, A., et al., 2022. Criteria used to review weighting methods as part of the un environment life cycle initiative's global guidance on environmental life cycle impact assessment indicators (glam) project, in: SETAC Europe 32nd Annual Meeting: Towards a Reduced Pollution, Society of Environmental Toxicology and Chemistry. pp. 559–560.

Duclos, J.Y., Échevin, D., 2011. Health and income: A robust comparison of Canada and the US. Journal of Health Economics 30, 293–302.

Duclos, J.Y., Sahn, D.E., Younger, S.D., 2006. Robust multidimensional poverty comparisons. The Economic Journal 116, 943–968.

Duclos, J.Y., Sahn, D.E., Younger, S.D., 2007. Robust multidimensional poverty comparisons with discrete indicators of well-being, in: Jenkins, S.P., Mickleright, J. (Eds.), Poverty and Inequality Re-examined. Oxford University Press Oxford, UK, pp. 185–208.

Dyckerhoff, R., Mosler, K., 1997. Orthant orderings of discrete random vectors. Journal of Statistical Planning and Inference 62, 193–205.

Edwards, W., Von Winterfeldt, D., 1986. Decision analysis and behavioral research. Cambridge University Press 604, 6–8.

Ehrgott, M., Figueira, J., Greco, S., 2010. Trends in multiple criteria decision analysis. volume 6. Springer.

Ferrer-i-Carbonell, A., Frijters, P., 2004. How important is methodology for the estimates of the determinants of happiness? The economic journal 114, 641–659.

Figueira, J., Greco, S., Ehrgott, M. (Eds.), 2005. Multiple criteria decision analysis: state of the art surveys. Springer Science & Business Media.

Fleurbaey, M., Blanchet, D., 2013. Beyond GDP: Measuring welfare and assessing sustainability. Oxford University Press.

Frey, B.S., Luechinger, S., Stutzer, A., 2009. The life satisfaction approach to valuing public goods: The case of terrorism. Public Choice 138, 317–345.

Gaganis, C., Pasiouras, F., Tasiou, M., Zopounidis, C., 2021. Cisef: A composite index of social, environmental and financial performance. European Journal of Operational Research 291, 394–409.

Gravel, N., Moyes, P., 2013. Utilitarianism or welfarism: does it make a difference? Social Choice and Welfare 40, 529–551.

Gravel, N., Moyes, P., Tarroux, B., 2009. Robust international comparisons of distributions of disposable income and regional public goods. Economica 76, 432–461.

Gravel, N., Mukhopadhyay, A., 2010. Is india better off today than 15 years ago? a robust multidimensional answer. The Journal of Economic Inequality 8, 173–195.

Greco, S., Ishizaka, A., Matarazzo, B., Torrisi, G., 2018. Stochastic multi-attribute acceptability analysis (smaa): an application to the ranking of italian regions. Regional studies 52, 585–600.

Greco, S., Ishizaka, A., Tasiou, M., Torrisi, G., 2019. On the methodological framework of composite indices: A review of the issues of weighting, aggregation, and robustness. Social indicators research 141, 61–94.

Greco, S., Mousseau, V., Słowiński, R., 2014. Robust ordinal regression for value functions handling interacting criteria. European Journal of Operational Research 239, 711–730.

Greco, S., Mousseau, V., Słowiński, R., 2008. Ordinal regression revisited: multiple criteria ranking using a set of additive value functions. European Journal of Operational Research 191, 416–436.

Greco, S., Munda, G., 2017. Multi-criteria evaluation in environmental policy analysis, in: Spash, C.L. (Ed.), The Routledge Handbook of Ecological Economics: Nature and Society. Routledge, p. 10.

Harvey, C.M., Østerdal, L.P., 2010. Cardinal scales for health evaluation. Decision Analysis 7, 256–281.

Hussain, M.A., 2022. Poverty in the five welfare regimes of Europe, in: De Gruyter Handbook of Contemporary Welfare States. Walter de Gruyter GmbH & Co KG. volume 1, p. 321.

Hussain, M.A., Jørgensen, M.M., Østerdal, L.P., 2016. Refining population health comparisons: a multidimensional first order dominance approach. Social Indicators Research 129, 739–759.

Hussain, M.A., Siersbæk, N., Østerdal, L.P., 2020. Multidimensional welfare comparisons of EU member states before, during, and after the financial crisis: a dominance approach. Social Choice and Welfare 55, 645–686.

Kadziński, M., 2022. Robust ordinal regression for multiple criteria decision aiding, in: Intelligent Decision Support Systems: Combining Operations Research and Artificial Intelligence-Essays in Honor of Roman Słowiński. Springer, pp. 185–205.

Lehmann, E., 1955. Ordered families of distributions. The Annals of Mathematical Statistics , 399–419.

Leshno, M., Levy, H., 2002. Preferred by "all" and preferred by "most" decision makers: almost stochastic dominance. Management Science 48, 1074–1085.

Levhari, D., Paroush, J., Peleg, B., 1975. Efficiency analysis for multivariate distributions. The Review of Economic Studies 42, 87–91.

Lindén, D., Cinelli, M., Spada, M., Becker, W., Gasser, P., Burgherr, P., 2021. A framework based on statistical analysis and stakeholders' preferences to inform weighting in composite indicators. Environmental Modelling & Software 145, 105208.

Marling, T.G., Range, T.M., Sudhölter, P., Østerdal, L.P., 2018. Decomposing bivariate dominance for social welfare comparisons. Mathematical Social Sciences 95, 1–8.

Merad, M., Dechy, N., Serir, L., Grabisch, M., Marcel, F., 2013. Using a multi-criteria decision aid methodology to implement sustainable development principles within an organization. European Journal of Operational Research 224, 603–613.

Mosler, K., Scarsini, M., 1991. Some theory of stochastic dominance, in: Mosler, K., Scarsini, M. (Eds.), Stochastic Orders and Decision Under Risk. Institute of Mathematical Statistics, Heyward. volume 19, pp. 261–284.

Müller, A., Scarsini, M., Tsetlin, I., Winkler, R.L., 2023. Multivariate almost stochastic dominance: Transfer characterizations and sufficient conditions under dependence uncertainty. Operations Research .

Müller, A., Stoyan, D., 2002. Comparison methods for stochastic models and risks. volume 389. Wiley.

Muller, C., Trannoy, A., 2011. A dominance approach to the appraisal of the distribution of well-being across countries. Journal of Public Economics 95, 239–246.

Nardo, M., Saisana, M., Saltelli, A., Tarantola, S., A, H., E, G., 2008. Handbook on Constructing Composite Indicators: Methodology and User Guide. 2 ed., OECD publishing, Paris.

OECD, 2020. How's Life? 2020. URL: https://www.oecd-ilibrary.org/content/publication/9870c393-en.

Østerdal, L.P., 2010. The mass transfer approach to multivariate discrete first order stochastic dominance: direct proof and implications. Journal of Mathematical Economics 46, 1222–1228.

Pinar, M., Cruciani, C., Giove, S., Sostero, M., 2014. Constructing the feem sustainability index: A choquet integral application. Ecological Indicators 39, 189–202.

Powdthavee, N., Van Den Berg, B., 2011. Putting different price tags on the same health condition: Re-evaluating the well-being valuation approach. Journal of Health Economics 30, 1032–1043.

Pérez-Ortiz, M., de la Paz-Marín, M., Gutiérrez, P., Hervás-Martínez, C., 2014. Classification of eu countries' progress towards sustainable development based on ordinal regression techniques. Knowledge-Based Systems 66, 178–189.

Range, T.M., Østerdal, L.P., 2019. First-order dominance: stronger characterization and a bivariate checking algorithm. Mathematical Programming 173, 193–219.

Resce, G., Schiltz, F., 2021. Sustainable development in europe: A multicriteria decision analysis. Review of Income and Wealth 67, 509–529.

Rockafellar, R.T., 1970. Convex analysis. Princeton university press.

Rowley, H.V., Peters, G.M., Lundie, S., Moore, S.J., 2012. Aggregating sustainability indicators: Beyond the weighted sum. Journal of Environmental Management 111, 24–33.

Schang, L., Hynninen, Y., Morton, A., Salo, A., 2016. Developing robust composite measures of healthcare quality–ranking intervals and dominance relations for scottish health boards. Social Science & Medicine 162, 59–67.

Sen, A., Foster, J.E., 1997. On economic inequality. Oxford University Press.

Shaked, M., Shanthikumar, J.G., 2007. Stochastic orders. Springer Science & Business Media.

Siersbæk, N., Østerdal, L.P., Arndt, C., 2016. Multidimensional first-order dominance comparisons of population wellbeing. Measuring poverty and wellbeing in developing countries. In Arndt, C. and Tarp ,F. (eds.): Measuring poverty and wellbeing in developing countries , 24.

Stiglitz, J.E., Fitoussi, J.P., Durand, M., 2018. Beyond GDP: measuring what counts for economic and social performance. OECD Publishing Paris.

Stiglitz, J.E., Sen, A., Fitoussi, J.P., et al., 2009. Report by the commission on the measurement of economic performance and social progress.

Tsetlin, I., Winkler, R.L., 2018. Multivariate almost stochastic dominance. Journal of Risk and Insurance 85, 431–445.

UN, 2015. Transforming our world: the 2030 agenda for sustainable development. URL: https://wedocs.unep.org/20.500.11822/9814.

United Nations Development Programme, 2022. Human Development Report 2021/2022. URL: https://www.un-ilibrary.org/content/books/9789210016407.

Van Praag, B.M., 1991. Ordinal and cardinal utility: an integration of the two dimensions of the welfare concept. Journal of Econometrics 50, 69–89.