Cardinality and Utilitarianism through social interactions*

Shiri Alon† and Ehud Lehrer‡

November 21, 2021

Abstract

We offer an axiomatic model of a utilitarian social planner, where cardinality of the aggregated individual preferences is not assumed but rather derived from social interactions. Furthermore, social interactions determine some (though not all) of the interpersonal comparisons between members of society.

Keywords:
JEL classification:

*The authors are thankful to Peter Wakker and Gabi Gayer for invaluable suggestions and comments. This research was done while the first author was visiting the Department of Economics at the University of Washington. The author is grateful to the department for its hospitality. The second author acknowledges the support of the grants ISF 591/21 and DFG KA 5609/1-1.

†Bar-Ilan University, Ramat-Gan 5290002, Israel. e-mail: `Shiri.Alon-Eron@biu.ac.il`

‡The School of Mathematical Sciences, Tel Aviv University, Tel Aviv 69978, Israel. e-mail: `lehrer@post.tau.ac.il`

1 Introduction

When making decisions for a society, preferences of members of society should be taken into account. If a social planner manages to elicit preferences of individuals, s/he then needs to aggregate these preferences in order to reach a social preference. The question of aggregating possibly conflicting preferences of individuals, and deciding how to balance them, is a fundamental question in social choice.

Preferences of members of society can take many forms. In many domains, for instance when considering preferences over bundles of goods, preferences are typically ordinal. Namely, the only information regarding preferences is how alternatives are ranked, without any way to measure the extent to which one alternative is better than another, or to compare tradeoffs between alternatives.

When preferences are ordinal, and there is no notion of strength of preference, aggregation is difficult. As established by Arrow in his famous impossibility theorem [2], it is impossible to aggregate preferences while abiding by a set of seemingly reasonable aggregation principles. Conceptually, part of the problem is that without knowing anything about how strongly individuals prefer one alternative to another, it is difficult to weigh the advantages of one social alternative over another in a consistent manner.[1]

As oppose to ordinal preferences, cardinal preferences convey strength of preference over alternatives. They thus facilitate interpersonal comparisons, allowing to balance conflicting tastes of individuals by give and take based on how individuals quantify tradeoffs between alternatives. And indeed, contrary to Arrow's impossibility theorem for ordinal preferences, Harsanyi [8] showed how for cardinal preferences, of both the social planner and members of society, aggregation takes on a simple form - as long as the social planner adheres to the widely accepted Pareto principle, aggregation must be utilitarian. That is, social decisions are represented by a weighted sum of the individuals' cardinal utilities. The social planner still needs to choose weights, and this is a crucial step in the comparison of benefits and harms for the different members of society.

[1]We should be careful here, though. Schmeidler and Kalai [10] showed that even with cardinal preferences, a cardinal version of IIA that was suggested by Samuelson leads again to impossibility (combined with basic assumptions). In the spirit of the original IIA, this version considers in each application two relative distances between alternatives. However when preferences are cardinal a lot of data is lost by considering only some relative distances without comparing them, for example, to some benchmark distance.

Harsanyi's result is agnostic about this issue. Yet, under cardinality of preferences the formation of a social preference is considerably simplified: it is reduced to the choice of specific weights for the individual utilities that are mutually calibrated.[2]

Cardinal preferences are typically derived when observed preferences are defined on a rich domain of alternatives, by imposing appropriate requirements on them. The key to this derivation is that the domain of alternatives is composed of (at least) two separable components. For instance, tradeoffs can be measured when alternatives are composed of both prizes and probabilities, two conceptually different components. When these two components are regarded separately, an appropriate set of axioms yields cardinal preferences. With cardinal preferences the tradeoff between two prizes is represented by the difference between their utilities.

Similarly, cardinal preferences are often derived based on alternatives that are composed of prizes and states of nature, as in Savage [16], or of prizes and time, with appropriate sets of assumptions. Both states of nature and time can be employed in a similar manner to probabilities in order to measure tradeoffs, which translate to utility differences.

To obtain simple, utilitarian aggregation of preferences under the Pareto assumption, the preferences of both the social planner and members of society need to be cardinal. There are several points where that could go wrong.

A first obstacle might exist in decision problems where the alternatives under consideration do not include a distinct component such as probabilities or the like, or if they do, preferences may not satisfy separability over that component. For example, alternatives may be bundles of goods without separability between different goods. A second problem might arise even when appropriate separability is satisfied for individual preferences. In this case, as individual and social preferences are normatively and positively different, separability may not be applicable for society. Moreover, even if alternatives are lotteries, for instance, and individual preferences take an expected utility form so that tradeoffs with respect to probabilities can be derived, these tradeoffs may not be relevant for the social decision. An individual may quantify the tradeoff between two prizes in a particular way when weighing it against probabilities, but may not feel the same when this tradeoff is used to balance her/his preferences with respect to those of others.

[2] Various methods have been proposed in the literature to resolve relative scaling of individuals' cardinal utilities. See, for example, Fleurbaey and Zuber [7] and the discussion and references therein.

In this paper we consider preferences of individuals and a social planner on a general set of alternatives, without assuming that it contains a separable element such as probabilities, time, etc. The primitives in our model are an abstract set of alternatives, a society composed of a finite number of individuals, and a family of preference relations, one per each group in society (single members, the entire society, and everything in between).[3] The interpretation for each of these relations depends on the group. Whenever a group is *organic*, in that it consists of individuals who make joint decisions of their own accord (e.g., when the group is a household), then the primitive preference relation per this group is interpreted as the joint preference relation that group members form together voluntarily, without interference of the social planner. The social planner is assumed to accept the autonomous preferences of organic groups, just as s/he accepts those of individuals. For other groups (among them the entire society), the relation is interpreted as the social planner's preference for these groups. From a normative point of view, if the social planner accepts the norms suggested by our axioms, our results provide guidelines as to how s/he should construct these preferences (including the preference of the entire society).

We formulate conditions on the family of preference relations that yield the existence of a corresponding family of cardinal utilities, one per each group in society. These utilities are utilitarian, so that for every partition of a group in society into sub-groups, the utility for this group is a weighted sum of the sub-groups' utilities. In particular it is implied that for each group in society, the preference for this group is represented by a utilitarian sum of the cardinal utilities of the members of the group.

Cardinality of preferences in our model is the result of eliciting consistent *social tradeoffs*, that are extracted from compromises that individuals make with one another. That is, the component with respect to which tradeoffs of individual preferences are measured is just other individuals. These tradeoffs are therefore inherent to the social question, derived precisely from the type of problems for which they are to be used. For example, consider a household consisting of two individuals. The individuals have their own personal preferences over alternatives, and in addition there is the preference of the household, expressing the decisions made by the individuals together. Imagine that

[3]That is, we consider a single, fixed profile of preferences, as in Harsanyi [8], and not a universal rule over multiple profiles with consistency requirements across societies, as in Arrow [2].

the individuals want to go to the mall, and need to decide whether to walk or take the bus. If each of the individuals personally prefers taking the bus then they will probably prefer to take it together. However, if they disagree, one preferring the bus and the other walking, they will have to compromise. We look at such compromises, comparing those that individuals are willing and unwilling to make, and employ them for measuring strength of preference.

The motivation for our model is twofold. First, suppose that a social planner wishes to aggregate the preferences of members of society, and finds the conditions that we formulate appealing (this would require the social planner to accept the conditions when applied to the preferences of every group in society, not only the entire society). In light of our result, the social planner can deduce that all preferences are represented by cardinal utilities, where each group utility is a utilitarian sum of the utilities of its subgroups, and specifically of its members. The social planner then needs to figure out which cardinal utilities to use.

The second element of our model is a characterization of the circumstances under which a social planner may elicit cardinal preferences of individuals and organic groups, and a utilitarian representation thereof, based on consistent social tradeoffs that members of society voluntarily make. Derivation of these tradeoffs relies on an assumption that all individuals are engaged in interactions with others, where they need to reach joint decisions. For example, individuals may be organized in households, and need to make decisions as a household, so that members of a household have their own personal preferences and on top of that there is a preference of the household.[4] For instance, we may see one of the household members walking to the mall when alone, but taking the bus with the other household member.

Whenever the preferences of individuals and their organic groups conform to our conditions, then the preferences of individuals can be represented by cardinal utilities, and the preferences of any organic group is represented by a weighted sum of the cardinal utilities of its members. The social planner may observe these preferences and elicit the implied cardinal utilities.

After eliciting the cardinal utilities representing the preferences of individuals and

[4]organic groups can overlap, in which case our assumptions require consistency in the compromises that individuals are willing to make with their peers in different organic groups, in that the social tradeoffs that these compromises convey are the same across groups.

their organic groups, these can be used in forming the cardinal utilities representing the preferences of all other, non-organic groups (including the entire society). If the social planner accepts our conditions, chiefly Pareto and consistency of social tradeoffs, then the preferences of any non-organic group is represented by a weighted sum of the cardinal utilities of its organic sub-groups, thus in particular by a weighted sum of the cardinal utilities of the group's members. In this utilitarian sum, the weights assigned to any organic sub-group are subject to the social planner's discretion, and should reflect value judgments on the planner's part. However, within any organic sub-group, the weights on the utilities of individuals are determined by their compromises within the sub-group, namely by the social tradeoffs that they themselves are willing to make with their peers in the sub-group.

Altogether we model societies where members make consistent tradeoffs when compromising with others, and the social planner is a utilitarian whose utilitarian social welfare function hinges on those social tradeoffs that members willingly make. Contrary to other social choice models that employ cardinal preferences, cardinality in our model is the direct result of social interaction, and not of assumptions made separately on each preference (as is the case if, for example, preferences are assumed to be expected utility). The price to pay is a primitive which is more complicated than in most models of social choice of that kind - we suppose a family of preference relations, one per each group in society, rather than only individual preferences and a preference for the entire society. For organic groups, the preferences assumed are supposed to be preferences that individuals voluntarily form. For non-organic groups, the social planner needs to be able to apprehend and assess conditions that apply to proper sub-groups and not only to the entire society.

It should be noted that while the representation we characterize will hold whenever preferences abide by our conditions, its normative appeal depends on our interpretation of preference reversals within organic groups. We interpret such reversals as resulting from willful compromises that individuals make. We contend that tradeoffs that result from compromises with significant others express how individuals voluntarily take into account others' preferences, and as such are a relevant construct for social decisions. However, if tradeoffs as we measure them are a result of other dynamics within organic groups, for example an excess bargaining power of one of the members, or if individuals are unaware of the true preferences of their peers or misperceive them, it may not be normatively warranted to take those tradeoffs into account in a social decision. Our model is thus adequate for use when alternatives and groups are such that the social planner can safely assume that joint decisions result from willful compromises between individuals, where these individuals know each other's preferences when making joint decisions. For instance, think of spouses who need to decide whether to walk or take the bus when going together to the mall. They discuss the matter, become aware of each other's preferences, and in case there's a conflict of preferences one of them compromises.

Utilitarianism was criticized, for example by Robbins [14] and Samuelson [15], for its assumption that utilities of individuals can be compared and weighed relative to one another. According to this criticism, it is meaningless to compare satisfaction levels of different individuals hence meaningless to aggregate utilities in the form of a weighted sum of specific utility indices. Rather, only ordinal preferences should be taken into account when reaching social decisions.

In our model we consider a general set of alternatives, and each preference, when considered independently, is merely assumed ordinal. Cardinality of preferences emerges only after individuals are placed in a social context, based on assumptions that regard their joint decisions in their organic groups. The compromises that these joint decisions require express interpersonal comparisons that individuals voluntarily perform. Thus, relative weights on utilities of individuals that belong to the same organic group are endogenously determined by the individuals themselves through the compromises that they make with one another, rather than exogenously set by the social planner. Comparison of the utilities of individuals that do not belong to the same organic group is left to the discretion of the social planner and not determined within the model (see Footnote 2). In special cases, with enough intersection between groups, all weights will be determined by individuals' voluntary choices (e.g., in the simple case where there is one individual with which everybody else forms organic groups).

Our result relies heavily on the concept of tradeoffs between outcomes. Comparing tradeoffs between outcomes as a technique for extracting cardinal utility originated in Thomsen [17], further developed in Blaschke and Bol [3] and Debreu [5], and was thoroughly investigated in Krantz, Luce, Suppes and Tversky [12] (in a form termed standard sequences). Tradeoffs as we define and employ here were defined and extensively studied by Wakker, see for example [18] and [11].

Technically, in the special case where alternatives are allocations of bundles and individuals care only about their personal allocated bundles, only basic assumptions are required to obtain that there exist cardinal individual utilities, such that the social planner's preference for the entire society is the sum of those individual utilities. Still, with only those basic assumptions the individual cardinal utilities do not convey any meaningful notion of strength of preference, while in our model cardinality of the utilities stems from the social tradeoffs that individuals are willing to make. Moreover, our setup allows for a general set of alternatives, subject only to connectivity constraints, and so can accommodate a variety of social choice problems. Even when allocations are considered, our setup can describe individuals that care about more than their personally allocated bundles, accommodating, for example, public goods and other-regarding preferences.

We mention two related papers that address the above question of interpersonal comparisons by characterizing both cardinality of preferences and a utilitarian social decision.

Piacquadio [13] starts off with ordinal preferences, extracts cardinal utilities based on social considerations, and aggregates them into a utilitarian social preference. To obtain cardinal preferences Piacquadio [13] employs conditions that express fairness considerations on the part of the social planner, translated in the representation to an endogenous choice of expanding opportunity sets. These opportunity sets reflect a moral stance of the social planner, derived from the planner's preferences. Consequently, differences of individual utility values express the tradeoffs in their well-being levels as these are perceived by the social planner based on her/his concept of fairness. By setting forth these fairness considerations, Piacquadio's model delivers as part of the cardinal utilities also the relative scaling of individuals in society, resolving interpersonal comparisons according to the value judgments derived from the social planner's preferences.

The approach in our paper is different from Piacquadio's in that in our paper the cardinality of individual utilities expresses the individuals' own strength of preference, extracted from the compromises that they make with significant others. Thus differences of utility values in our model represent tradeoffs in well-being as perceived by individuals themselves and not by the social planner, where these tradeoffs are inherently related to the social context. The additional information that we employ for that derivation are more demanding primitives - a preference for every sub-group in society, not only for individuals and the entire society. We obtain a social planner who is a utilitarian for every such sub-group, having at her/his disposal cardinal utilities to be aggregated. For individuals that belong to the same organic group, the relative scaling of their utilities is not at the discretion of the social planner, but extracted from their preferences and that of their mutual organic group. However, relative scaling of utilities for individuals that are not members of the same organic group, so in fact scaling of organic groups relative to one another, is undetermined by our model.

Another method for attributing cardinal meaning to preferences and performing interpersonal comparisons, suggested by Edgeworth [6], is by use of just noticeable differences. Angenziano and Gilboa [1] recently formulated an axiomatic foundation of that idea, using just noticeable differences to calibrate the wellbeing of individuals. The representation of personal preferences and their calibration is motivated by theories of psychological perception, and by the social planner's contention that just noticeable differences perceived by different individuals should obtain equal value.

The paper is organized as follows. Section 2 explains the setup that we use, and details the assumptions that we impose on preferences. Section 3 contains our result, with a special case presented in Subsection 3.1. The proof of the special case can be found in an online appendix[5], and all other proofs appear in Section 4.

2 Setup and assumptions

Suppose a non-empty set of outcomes, X, and a society of individuals $N = \{1, \ldots, n\}$, $2 \leq n \in \mathbb{N}$. For every nonempty subset of individuals $T \subseteq N$ there is a binary relation $\succsim^T$ over X. The asymmetric and symmetric parts of each of these relations are respectively denoted by $\succ^T$ and $\sim^T$. For $T = \{i\}$ we write $\succsim^i, \succ^i$, and $\sim^i$. An outcome $x \in X$ is called *unanimously minimal* if for every $y \in X$, and every $i \in N$, $y \succsim^i x$. An outcome $x \in X$ is called *unanimously maximal* if for every $y \in X$, and every $i \in N$, $x \succsim^i y$.

The following structural assumption delineates the type of problems that we address,

[5] https://shirialon.weebly.com/uploads/2/3/9/9/23990200/lehreralon_cardinality_and_utilitarianism_online_appendix.pdf

characterized by two topological assumptions. These will be satisfied, for example, when alternatives are bundles in $\mathbb{R}^L_+$, and all indifference classes for every relation $\succsim^T$ are connected.

A0. Structural assumption.

(a) X is a connected topological space.

(b) For every nonempty $T \subseteq N$ and every $x \in X$, the indifference class $\{y \in X \mid y \sim^T x\}$ is connected.

The first assumption contains a set of basic requirements, stating that each relation $\succsim^T$ is a continuous and non-degenerate complete order.

A1. Non-degenerate Continuous weak orders. For every nonempty $T \subseteq N$ the relation $\succsim^T$ is a non-degenerate, continuous weak order. That is,

(a) For any $x, y \in X$, either $x \succsim^T y$ or $y \succsim^T x$ (completeness)

(b) For any $x, y, z \in X$, if $x \succsim^T y$ and $y \succsim^T z$ then $x \succsim^T z$ (transitivity)

(c) For every $x \in X$ the sets $\{y \in X \mid y \succsim^T x\}$ and $\{y \in X \mid x \succsim^T y\}$ are closed (continuity)

(d) There are $x_T, y_T \in X$ such that $x_T \succ^T y_T$ (non-degeneracy)

The following is the first Pareto condition out of the two that we employ in our characterization. This is a standard, strong Pareto condition, applied to the family of relations $\succsim^T$.

A2. Extended Pareto. For any two outcomes $x, y \in X$ and for any two nonempty, disjoint subsets $T, G \subset N$, if $x \succsim^T y$ and $x \succsim^G y$ then $x \succsim^{T \cup G} y$.
Furthermore, if any of the two antecedent conditions holds strictly, then the conclusion is strict as well.

Repeated application of Extended Pareto yields that whenever for a partition of a group T the preferences per each partition element support a ranking, then so does the preference for T. In particular, under Extended Pareto, if x_* is unanimously minimal, then for every $y \in X$ and every nonempty subset of individuals $T \subseteq N$, $y \succsim^T x_*$. Similarly, if x^* is unanimously maximal, then for every $y \in X$ and every nonempty subset of individuals $T \subseteq N$, $x^* \succsim^T y$.

Agreed Improvement, our next axiom, is a general version of monotonicity. For preferences over bundles in $\mathbb{R}^L_+$ it is enough to require that all preferences be monotone.

A3. Agreed Improvement For any two nonempty, disjoint $T, G \subset N$,

(a) For every $x, y \in X$ that are not unanimously maximal, there is an outcome $z^* \in X$ such that $z^* \succ^G x$ and $z^* \succ^T y$.

(b) For every $x, y \in X$ that are not unanimously minimal, there is an outcome $z_* \in X$ such that $x \succ^G z_*$ and $y \succ^T z_*$.

Diversity, our fourth assumption, is a richness condition. If alternatives are bundles in $\mathbb{R}^L_+$, and under the basic assumptions A1 and A3, it will fail if there are two disjoint groups that have an identical indifference class. That is, it will fail if there are two disjoint groups T and G and a bundle x such that the indifference classes of x under $\succsim^T$ and under $\succsim^G$ merge completely (if they partially merge the axiom will still be satisfied).

A4. Diversity of Tastes. For any two nonempty, disjoint subsets $T, G \subset N$, and for every $x \in X$ which is neither unanimously maximal nor unanimously minimal, there exists $y \in X$ such that either $x \succ^G y$ and $y \succ^T x$, or, $x \succ^T y$ and $y \succ^G x$.

Next we define social tradeoffs, the central construct in our paper that is used to extract cardinality of preferences. The basic idea of social tradeoffs is simple. Consider two individuals, 1 and 2, each having a personal preference over the general set of alternatives, and both having a joint preference, representing their decisions together. Suppose that 2 is personally indifferent between two alternatives x and z, as well as between alternatives y and w. This would mean that as far as 2 is concerned, the comparison between x and y is the same as the comparison between z and w.

Now inspect the joint preference of 1 and 2. Say that, together, 1 and 2 prefer x over y, and prefer w over z. We interpret such a situation as a sign that for 1, the tradeoff between x and y is larger than the tradeoff between z and w.

To see the intuition in this interpretation assume for the sake of the explanation that 2 prefers y to x, hence also w to z. As both comparisons are identical for 2, we conclude that the difference in the joint ranking is the result of 1's preference. The joint preference complies with 2 in ranking w above z, but overturns 2's preference with regard to x and y. It must therefore be that 1 prefers x over y to a greater extent than it prefers (if at all) z over w. Under these circumstances we thus conclude that for 1, the tradeoff between x and y is larger than the tradeoff between z and w, denoting it by $x \ominus y \succeq^1 z \ominus w$. An analogous measurement can be made when individuals 1 and 2 are replaced by groups.

Social tradeoffs are defined according to this rationale for every nonempty strict group in society. The definition hinges on preferences, stating that indifferent outcomes yield the same tradeoffs. That is, strength of preference, that is manifested in comparisons of tradeoffs, is defined as an attribute of indifference classes. The definition is followed by an assumption that guarantees its consistency.

Definition 1. For outcomes $x, y, z, w \in X$ and a nonempty subset $T \subset N$, write

$$x \ominus y \succeq^T z \ominus w$$

whenever there exist outcomes $x', y', z', w' \in X$ such that

$$x \sim^T x', y \sim^T y', z \sim^T z',\ w \sim^T w',$$

and a nonempty subset $G \subseteq N$, $G \cap T = \emptyset$, such that,

$$\begin{array}{ccc} x' \sim^G z' & , & y' \sim^G w' \ , \\ x' \succsim^{T \cup G} y' & \text{and} & w' \succsim^{T \cup G} z' \ . \end{array} \tag{1}$$

If it is furthermore satisfied that $x' \sim^{T \cup G} y'$ and $w' \sim^{T \cup G} z'$, write $x \ominus y \simeq^T z \ominus w$, otherwise, $x \ominus y \succ^T z \ominus w$.

Figure 1 illustrates a measurement of tradeoffs for Individual 1 using the personal preferences of Individual 2 and the joint preferences of the two individuals.

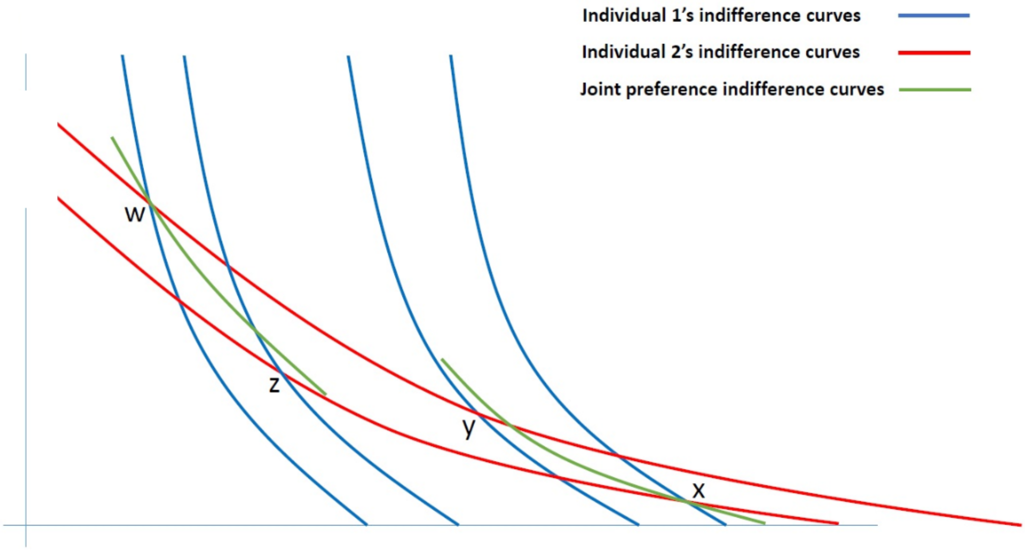


Figure 1: Definition of social tradeoffs

The tradeoffs between the blue indifference curves of Individuals 1 are measured with the aid of Individual 2's red indifference curves, and the individuals' joint preference that is sketched in green. The two individuals' personal ranking of x vs. y and of z vs. w is opposite. The outcomes satisfy, $x \sim^2 z$, $y \sim^2 w$, $x \succsim^{\{1,2\}} y$, and $w \succsim^{\{1,2\}} z$. Hence we conclude that for Individual 1 the tradeoff between x and y is higher than the tradeoff between z and w.

For individuals, as well as for organic sub-groups of organic groups (namely, sub-groups reaching joint decisions voluntarily, and further reaching joint decisions with a larger group voluntarily) it is understood that these are the tradeoffs that the individuals and the organic sub-groups willingly make when compromising with their peers in order to reach joint decisions. For non-organic groups these tradeoffs are interpreted as exhibited by the social planner.

The assumption that follows guarantees that the definition above is consistent in two respects. First, tradeoffs should be the same no matter where on the indifference classes of a group they are measured. This is illustrated in Figure 2, which shows that the ranking of tradeoffs between Individual 1's indifference curves (in blue) should be the same whether they are measured with the aid of Individual 2's continuous or broken red indifference curves.

The second type of consistency required by the assumption is across groups. If a group exhibits some ranking of tradeoffs when measured through the joint preference with another group, then the same ranking of tradeoffs will hold if measured using a different group (this will be expressed similarly to Figure 2, with the continuous and broken red lines representing preferences of different groups). This results that tradeoffs derived from voluntary compromises of individuals and sub-groups are consistent across organic groups to which they belong, and that the social planner respects those tradeoffs when constructing preferences of non-organic groups. Note that for the entire society there is no notion of tradeoffs, as the entire society has no individuals outside of it to compromise with.

A5. Consistency of Social Tradeoffs. Let $T, G, H \subseteq N$ be nonempty subsets of individuals such that $T \cap G = T \cap H = \emptyset$. Suppose that for outcomes $x, y, z, w, x', y', z', w' \in X$ it holds that:

(a) $x \sim^T x'$, $y \sim^T y'$, $z \sim^T z'$, and $w \sim^T w'$

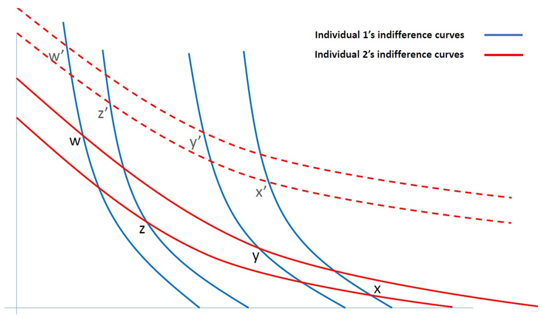
(b) $x \sim^G z$ and $y \sim^G w$

(c) $x' \sim^H z'$, and $y' \sim^H w'$

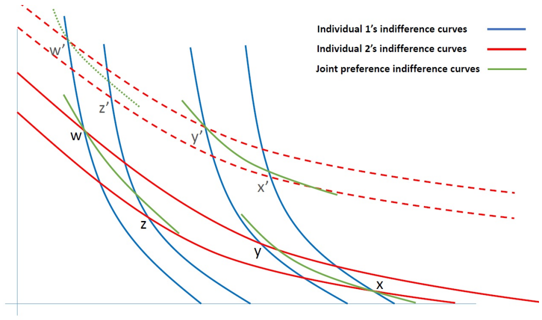
Then $x \succsim^{T \cup G} y$, $w \succsim^{T \cup G} z$, and $y' \succsim^{T \cup H} x'$ imply $w' \succsim^{T \cup H} z'$.

Our last assumption is an indifference Pareto assumption concerning tradeoffs. Within the axiom, the conditions that apply to tradeoff comparisons are satisfied only when the ranking of tradeoffs can be established as per Definition 1. When the conditions are satisfied, namely when two tradeoffs are indifferent according to two disjoint groups, then it *cannot* be that their union strictly ranks these tradeoffs. That is, either the tradeoffs are indifferent according to the union as well, or they simply cannot be compared.

A6. Tradeoff Indifference-Pareto. Let $T, G \subset N$ be nonempty, disjoint subsets



(a) The tradeoffs between the blue indifference curves of Individuals 1 are the same whether measured through Individual 2's continuous or broken red indifference curves.



(b) If the continuous green preferences hold then the dotted green preference should hold as well.

Figure 2: Consistency of tradeoff measurement

such that $T \cup G \neq N$, and let $x, y, z, w \in X$ be outcomes. Then,

$$x \ominus y \simeq^T z \ominus w \quad \text{and} \quad x \ominus y \simeq^G z \ominus w \quad \Longrightarrow \quad \neg(x \ominus y \succ^{T \cup G} z \ominus w) \ .$$

3 Result

Definition 2. Two functions, $U_G, U_T : X \longrightarrow \mathbb{R}$, are said to be *jointly improvable*, when

(a) If neither of $x, y \in X$ obtains the minimum of both U_G and U_T over X, then there exists $z_* \in X$ such that,

$$\min\left(U_G(x), U_G(y)\right) > U_G(z_*) \quad , \quad \min\left(U_T(x), U_T(y)\right) > U_T(z_*) \ .$$

(b) If neither of $x, y \in X$ obtains the maximum of both U_G and U_T over X, then there exists $z^* \in X$ such that,

$$\max\left(U_G(x), U_G(y)\right) < U_G(z^*) \quad , \quad \max\left(U_T(x), U_T(y)\right) < U_T(z^*) \ .$$

Definition 3. Two functions, $U_G, U_T : X \longrightarrow \mathbb{R}$, are said to be *diversified*, if whenever

$x \in X$ neither obtains $\min_{y\in X} U_G(y)$, $\min_{y\in X} U_T(y)$, $\max_{y\in X} U_G(y)$, nor $\max_{y\in X} U_T(y)$, then there exists $y \in X$ such that either $U_G(y) > U_G(x)$ and $U_T(x) > U_T(y)$, or $U_G(x) > U_G(y)$ and $U_T(y) > U_T(x)$.

The following is our main theorem.

Theorem 1. *Let $\{\succsim^T\}_{\emptyset\neq T\subseteq N}$ be binary relations on X. Assume that A0 holds. Then the following two statements are equivalent:*

(i) $\{\succsim^T\}_{\emptyset\neq T\subseteq N}$ satisfy A1 through A6.

(ii) There exist continuous, non-constant utility functions $\{U_T\}_{\emptyset\neq T\subseteq N}$ such that U_T represents $\succsim^T$, and for every pair of nonempty, disjoint subsets of individuals $G, T \subset N$,

$$U_{G\cup T} = \lambda_G^{G\cup T} U_G + \lambda_T^{G\cup T} U_T, \text{ for } \lambda_G^{G\cup T}, \lambda_T^{G\cup T} > 0 \ .$$

Furthermore, $\{U_T\}_{\emptyset\neq T\subseteq N}$ are jointly cardinal[6], and for every nonempty and disjoint $G, T \subset N$, U_G and U_T are diversified and jointly improvable.

The proof appears in Section 4.

Corollary 1. *Let $\{\succsim^T\}_{\emptyset\neq T\subseteq N}$ be binary relations on X. Assume that A0 holds, and that these relations satisfy A1-A6. Then there are jointly cardinal utilities $\{U_i\}_{i\in N}$, such that for every nonempty $T \subseteq N$, $\succsim^T$ is represented by,*

$$U_T = \sum_{i\in T} \lambda_i^T U_i \ , \quad \lambda_i^T > 0 \ .$$

The corollary is easily obtained by applying the theorem over and over again to T, partitioning T into $\{i\}$ and $T\setminus\{i\}$ for $i \in T$, then partitioning $T\setminus\{i\}$ into $\{j\}$ and $T\setminus\{i,j\}$ for $j \in T\setminus\{i\}$, and so on. In the same manner, for every partition $\{T_1, \ldots, T_m\}$ of T, $U_T = \sum_{\ell=1}^{m} \lambda_{T_\ell}^T U_{T_\ell}$, for $\lambda_{T_\ell}^T > 0$.

[6]Namely, if there are functions $\{\hat{U}_T\}_{\emptyset\neq T\subseteq N}$ representing $\{\succsim^T\}_{\emptyset\neq T\subseteq N}$ in the same additive manner, then there are $\theta > 0$ and ξ_T for every $\emptyset \neq T \subseteq N$, such that $\hat{U}_T = \theta U_T + \xi_T$.

3.1 A special case

It may be that G is more influential relative to H when G and H are together with members of T than when G and H are with T'. We cannot even say that there is monotonicity of the influence with respect to set inclusion, namely that if we consider nonempty, pairwise disjoint subsets $G, H, T \subset N$, then the relative influence of T is smaller when T is with $G \cup H$ than when T is with G alone, in that $\lambda_T^{G\cup H\cup T}/\lambda_{G\cup H}^{G\cup H\cup T} < \lambda_T^{G\cup T}/\lambda_G^{G\cup T}$. Decreasing influence of T with respect to set inclusion may be violated when utilities are linearly dependent. For a threesome of disjoint subsets G, H and T, the utilities U_G, U_H and U_T may be linearly dependent, in which case the coefficient $\lambda_T^{G\cup H\cup T}$ may be relatively larger than $\lambda_T^{G\cup T}$, violating the above inequality, as U_T encapsulates some of U_G and U_H.

To impose more structure on the coefficients, and obtain a representation where the relative influence of any two subsets G and H is the same, independent of the set to which G and H join, Agreed Improvement (**A3**) is replaced by the following Unanimous Improvement assumption, and two additional assumptions, stated below, are added.

A3'. Unanimous Improvement

(a) For every $x, y \in X$ that are not unanimously maximal, there is an outcome $z^* \in X$ such that $z^* \succ^i x$ and $z^* \succ^i y$ for every $i \in N$.

(b) For every $x, y \in X$ that are not unanimously minimal, there is an outcome $z_* \in X$ such that $x \succ^i z_*$ and $y \succ^i z_*$ for every $i \in N$.

(c) If $z^* \succ^i z_*$ for every $i \in N$, then there is $z \in X$ such that $z^* \succ^i z \succ^i z_*$ for every $i \in N$.

The following axiom implies that the relative weights of utilities U_H and U_G are the same within the utility of any group that contains them. For that, the axiom compares tradeoffs across different groups. Essentially it states that when adding the tradeoff between z and y according to H when measured relative to G, to the tradeoff between y and x according to H when measured relative to T, we get the tradeoff between z and x according to H, measured relative to $G \cup T$. The last tradeoff is a sum of the two former ones, as the relative weights of U_H and U_G are the same within the utilities $U_{H\cup G}$ and $U_{H\cup G\cup T}$ (and the same for U_H and U_T).

The way this is expressed in the axiom is through comparison to tradeoffs between outcomes x^* and x_*. The first tradeoff, relative to G, exactly balances that between x^* and x_* according to G, and the second, relative to T, exactly balances the tradeoff between x^* and x_* according to T. To state additivity of those tradeoffs, the axiom employs outcomes x', z' that exhibit equivalent tradeoff to that of x and z according to H, and equivalent tradeoff to that of x^* and x_* according to $G \cup T$.

A7. Constant Relative Social Influence. Let $G, T, H \subsetneq N$ be three non-empty, pairwise disjoint groups of individuals, and $x, y, z, x', z', x^*, x_* \in X$ outcomes such that,

$$x \ominus y \simeq^G x^* \ominus x_* \quad \text{and} \quad x \sim^{G\cup H} y \quad ,$$
$$y \ominus z \simeq^T x^* \ominus x_* \quad \text{and} \quad y \sim^{T\cup H} z \; .$$

Then if $z', x' \in X$ satisfy,

$$x' \ominus z' \simeq^{G\cup T} x^* \ominus x_* \;\; \text{and} \;\; x' \ominus z' \simeq^H x \ominus z \; ,$$

then $x' \sim^{G\cup T\cup H} z'$.

Lastly, assumption A8 below characterizes a social preference that itself has cardinal content. Within the assumption, y is identified as the midpoint between x and z for two groups that partition the entire society (T and $N \setminus T$). This midpoint is interpreted as a midpoint for society, in that if another threesome of outcomes x', y', z' that is preference-equivalent according to the entire society satisfies that y' is a midpoint for a group in society, then it also must be perceived as the midpoint for the rest of society.

A8. Consistency of Societal Middle Ground. Let $T, G \subsetneq N$, and $x, y, z \in X$ outcomes such that

$$x \ominus y \simeq^T y \ominus z \;\; \text{and} \;\; x \ominus y \simeq^{N\setminus T} y \ominus z \; .$$

If $x', y', z' \in X$ satisfy $x' \sim^N x, y' \sim^N y$, and $z' \sim^N z$, then $x' \ominus y' \simeq^G y' \ominus z'$ implies $x' \ominus y' \simeq^{N \setminus G} y' \ominus z'$.

Before stating the result for the special case, another definition is required.

Definition 4. Functions, $u_1, \ldots, u_m : X \longrightarrow \mathbb{R}$, $m \in \mathbb{N}$, are said to be *unanimously improvable*, when

(a) If neither of $x, y \in X$ obtains the minimum of all u_i over X, $i = 1, \ldots, m$, then there exists z_* such that $\min(u_i(x), u_i(y)) > u_i(z_*)$ for every $i = 1, \ldots, m$.

(b) If neither of $x, y \in X$ obtains the maximum of all u_i over X, $i = 1, \ldots, m$, then there exists z^* such that $\max(u_i(x), u_i(y)) < u_i(z^*)$ for every $i = 1, \ldots, m$.

(c) If $u_i(z^*) > u_i(z_*)$ for every $i = 1, \ldots, m$, then there is $z \in X$ such that $u_i(z^*) > u_i(z) > u_i(z_*)$ for every $i = 1, \ldots, m$.

Under the extended set of axioms we can prove the following representation theorem. The proof is available as an online appendix.[7]

Theorem 2. *Let* $\left(\succsim^T\right)_{T \subseteq N}$ *be binary relations over* X*, and assume that A0 holds. Then the following two statements are equivalent:*

(i) The relations satisfy A1,A2,A3', and A4-A8.

(ii) There exist continuous utility functions $u_1, \ldots, u_n$ *over* X*, such that for each nonempty* $T \subseteq N$*,* $\succsim^T$ *is represented by,*

$$U_T = \sum_{i \in T} u_i \ .$$

Furthermore, the utilities u_i *are jointly cardinal and unanimously improvable, and for every nonempty and disjoint* $G, T \subset N$*,* U_G *and* U_T *are diversified.*

[7] https://shirialon.weebly.com/uploads/2/3/9/9/23990200/lehreralon_cardinality_and_utilitarianism_online_appendix.pdf

The more specific representation is simpler in its form. However when this specialized representation is in place, links between utilities of individuals in a group are restricted in the following manner: the relative influence of agents j and k is the same in the utility of every group to which they both belong, and moreover, within their joint utility $U_{\{j,k\}}$, their relative influence is determined by the relative influence of i and j within $U_{\{i,j\}}$, and the relative influence of i and k in $U_{\{i,k\}}$. These connections place restrictions on the type of voluntary preferences of organic groups that the model is able to accommodate, as well as on the social planner's flexibility to set weights on individuals' utilities that depend on the group under consideration.

4 Proof of Theorem 1

We start with the simpler direction.

4.1 Necessity: the axioms hold

Suppose there exist continuous utility functions $\{U_T\}_{\emptyset \neq T \subseteq N}$ over X as in (ii) of the theorem. Assumptions A1 and A2 immediately follow.

Agreed Improvement (**A3**) and Diversity of Tastes (**A4**) follow from the supposition that the corresponding U_G and U_T are jointly improvable and diversified.

To prove that Consistency of Social Tradeoffs (**A5**) holds let $T, G, H \subseteq N$ be nonempty subsets of individuals such that $T \cap G = T \cap H = \emptyset$, and $x, y, z, w, x', y', z', w' \in X$ outcomes, such that,

(a) $x \sim^T x'$, $y \sim^T y'$, $z \sim^T z'$, and $w \sim^T w'$,

(b) $x \sim^G z$ and $y \sim^G w$,

(c) $x' \sim^H z'$, and $y' \sim^H w'$.

Suppose that $x \succsim^{T \cup G} y$, $w \succsim^{T \cup G} z$, and $y' \succsim^{T \cup H} x'$. Then, expressing those preference relationships through the utilities representations,

(a') $U_T(x) = U_T(x')$, $U_T(y) = U_T(y')$, $U_T(z) = U_T(z')$, and $U_T(w) = U_T(w')$,

(b') $U_G(x) = U_G(z)$, $U_G(y) = U_G(w)$,

(c') $U_H(x') = U_H(z')$, $U_H(y') = U_H(w')$, and

(d') $\lambda_T^{G\cup T}U_T(x)+\lambda_G^{G\cup T}U_G(x)\geq\lambda_T^{G\cup T}U_T(y)+\lambda_G^{G\cup T}U_G(y)$,
$\lambda_T^{G\cup T}U_T(w)+\lambda_G^{G\cup T}U_G(w)\geq\lambda_T^{G\cup T}U_T(z)+\lambda_G^{G\cup T}U_G(z)$, and
$\lambda_T^{H\cup T}U_T(y')+\lambda_H^{H\cup T}U_H(y')\geq\lambda_T^{H\cup T}U_T(x')+\lambda_H^{H\cup T}U_H(x')$.

Substituting (b') in the first two inequalities in (d') yield $U_T(x)-U_T(y)\geq U_T(z)-U_T(w)$, and together with (a'), $U_T(x')-U_T(y')\geq U_T(z')-U_T(w')$. Employing (c'), the third inequality in (d') becomes, $\lambda_H^{H\cup T}\left(U_H(w')-U_H(z')\right)\geq\lambda_T^{H\cup T}\left(U_T(x')-U_T(y')\right)$, hence $\lambda_H^{H\cup T}\left(U_H(w')-U_H(z')\right)\geq\lambda_T^{H\cup T}\left(U_T(z')-U_T(w')\right)$, implying the required $w'\succsim^{T\cup H}z'$.

Tradeoff Pareto (**A6**) easily follows by noting that whenever tradeoffs are comparable, a tradeoff preference relationship holds if and only if differences of the corresponding utility are equal. Since all tradeoffs are assumed in the axiom to be comparable, the axiom simply states that inequality of utility differences according to U_T and U_G implies inequality of those differences according to the utility $U_{G\cup T}$, a trivial implication of the additive representation supposed in (ii) of the theorem.

4.2 Sufficiency: the representation holds

For clarity, we include in this subsection only the outline of the proof, composed of three logical steps, each formulated as a proposition. The detailed proofs of the propositions can be found in the appendix.

The first proposition states that for any fixed nonempty, disjoint subsets $G,T\subset N$, there exists an additive representation of $\succsim^{G\cup T}$.

Proposition 1. *Let $G,T\subset N$ be two disjoint subsets of individuals. Then there are utility functions $U_G^{G\cup T},U_T^{G\cup T}:X\longrightarrow\mathbb{R}$ such that $U_G^{G\cup T}$ represents $\succsim^G$, $U_T^{G\cup T}$ represents $\succsim^T$, and $U_G^{G\cup T}+U_T^{G\cup T}$ represents $\succsim^{G\cup T}$.*

Furthermore, the utilities that deliver this additive representation are jointly cardinal, namely if there are $\hat{U}_G,\hat{U}_T$ that obtain the same form of additive representation, then $\hat{U}_G=\tau U_G^{G\cup T}+\rho_G,\hat{U}_T=\tau U_T^{G\cup T}+\rho_T$, for $\tau>0$.

The proof can be found in the appendix, in Subsection 5.1.

According to Proposition 1, for any two non-empty, disjoint sets $T,G\subset N$, there exists a representation $U_{T\cup G}=U_T^{T\cup G}+U_G^{T\cup G}$ of $\succsim^{T\cup G}$ with continuous, jointly cardinal utilities $U_T^{T\cup G}$ and $U_G^{T\cup G}$ that represent $\succsim^T$ and $\succsim^G$, respectively. Likewise, there exists a representation $U_N=U_T^N+U_{N\setminus T}^N$ of $\succsim^N$ with continuous, jointly cardinal utilities U_T^N and $U_{N\setminus T}^N$ that represent $\succsim^T$ and $\succsim^{N\setminus T}$, respectively. The next proposition states that $U_T^{T\cup G}=\beta U_T^N+\tau$ for some $\tau,\beta\in\mathbb{R},\beta>0$. The proof of the proposition appears in the appendix in Subsection 5.2.

Proposition 2. *Let $G, T \subset N$ be two disjoint subsets of individuals. Let $U_G^{G\cup T} + U_T^{G\cup T}$ be the additive representation of $\succsim^{G\cup T}$ and $U_{N\setminus T}^N + U_T^N$ the additive representation of $\succsim^N$, according to Proposition 1. Then there are $\beta, \tau \in \mathbb{R}, \beta > 0$, such that $U_T^{T\cup G} = \beta U_T^N + \tau$.*

According to Proposition 2, all the utilities $U_T^{G\cup T}$, derived within additive representations $U_G^{G\cup T} + U_T^{G\cup T}$ of $\succsim^{G\cup T}$ for nonempty subsets of individuals $G \subset N$ disjoint from T, are cardinally the same.

Let $T \subsetneq N$ be a subset containing at least two individuals, and let U_T^N denote the cardinal utility representing $\succsim^T$ that results from applying Proposition 1 to $\succsim^T, \succsim^{N\setminus T}$ and $\succsim^N$. For a nonempty $H \subsetneq T$ Let U_T denote the cardinal utility representing $\succsim^T$ within the additive representation of $\succsim^T$ as obtained from partitioning T into $H, T \setminus H$. Both utilities represent $\succsim^T$. According to the next proposition, U_T^N and U_T are cardinally the same.

Proposition 3. *Let $T, H \subsetneq N$ be nonempty sets such that $H \subsetneq T$. Let $U_{N\setminus T}^N + U_T^N$ be the additive representation of $\succsim^N$ and $U_H^T + U_{T\setminus H}^T$ the additive representation of $\succsim^T$, as obtained according to Proposition 1. Then there are $\gamma, \xi \in \mathbb{R}$, $\gamma > 0$, such that $U_T^N = \gamma U_T + \xi$.*

The proof of the proposition can be found in Subsection 5.3 in the appendix.

We thus established that for every nonempty group of individuals $T \subsetneq N$, the utilities U_T that are obtained as the additive sum $U_T = U_H^T + U_{T\setminus H}^T$ for nonempty groups $H \subsetneq T$ (if those exist), and the utilities $U_T^{T\cup G}$ that are obtained within the additive sum $U_{T\cup G} = U_T^{T\cup G} + U_G^{T\cup G}$ for nonempty groups $G \subseteq N \setminus T$, are all cardinally related, namely they are positive affine transformations of one another.

For every $T \subseteq N$ choose a calibration and denote the utility thus calibrated by U_T. Hence there are utilities $\{U_T\}_{\emptyset \neq T \subseteq N}$ such that for every nonempty, disjoint $G, T \subset N$,

$$U_{G\cup T} = \lambda_G^{G\cup T} U_G + \lambda_T^{G\cup T} U_T, \quad \text{for } \lambda_G^{G\cup T}, \lambda_T^{G\cup T} > 0 \ .$$

The fact that the coefficients are strictly positive follows from the strong form of Pareto (**A2**) that we use.

The utilities $\{U_T\}_{\emptyset \neq T \subseteq N}$ are jointly cardinal, owing to the joint cardinality for every additive representation that is obtained at the beginning of the proof. Finally, for nonempty, disjoint $G, T \subset N$, U_G and U_T are jointly improvable as a result of Agreed Improvement (**A3**), and are diversified because of Diversity of Tastes (**A4**).

References

[1] Argenziano, R., and Gilboa, I. (2019), "Perception-Theoretic Foundations of Weighted Utilitarianism", *The Economic Journal*, 129(620), 1511-1528.

[2] Arrow, K. J. (1950), "A Difficulty in the Concept of Social Welfare", *Journal of Political Economy*, 58(4), 328-346.

[3] Blaschke, W. and G. Boll (1938), Geometric der Gewebe. Springer-Verlag, Berlin.

[4] Chateauneuf, A. and Wakker, P. (1993), "From local to global additive representation", *Journal of mathematical economics*, 22, 523-545.

[5] Debreu, Gérard (1960), "Topological Methods in Cardinal Utility Theory." In Kenneth J. Arrow, Samuel Karlin, & Patrick Suppes (1960, Eds), Mathematical Methods in the Social Sciences, 16Uf02d26, Stanford University Press, Stanford, CA.

[6] Edgeworth, F. Y. (1881), Mathematical Physics, London: Paul Kegan.

[7] Fleurbaey, M. and Zuber, S. (2017), "Fair utilitarianism", *American Economic Journal: Microeconomics*, 13(2), 370-401.

[8] Harsanyi, J. C. (1955), "Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility", *Journal of Political Economy*, 63(4), 309-321.

[9] Hiriart-Urruty, J.B. (1985), "Images of connected sets by semicontinuous multifunctions", *Journal of mathematical analysis and applications*, 111, 407-422.

[10] Kalai, E. and Schmeidler, D. (1977), "Aggregation Procedure for Cardinal Preferences: A Formulation and Proof of Samuelson's Impossibility Conjecture", *Econometrica*, 45(6), 1431-1438.

[11] Kobberling, V. and Wakker P. (2003), "Preference foundations for nonexpected utility: A generalized and simplified technique", *Mathematics of Operations Research*, 28, 395-423

[12] Krantz, D.H., R.D. Luce, P. Suppes, and A. Tversky (1971), Foundations of Measurement. New York: Academic Press.

[13] Piacquadio, P. G. (2017), "A fairness justification of utilitarianism", *Econometrica*, 85(4), 1261-1276.

[14] Robbins, L. (1935), An Essay on the Nature and Significance of Economic Science, London: Macmillan.

[15] Samuelson, P. (1947), Foundations of Economic Analysis, Cambridge: Harvard University Press.

[16] Savage, L. J. (1954), The Foundations of Statistics, New York: Wiley.

[17] Thomsen, G. (1927), "Un teorema topologico sulle schiere di curve e una caratterizzazione geometrica delle superficie isotermo-asintotiche", *Bollettino della Unione Matematica Italiana*, 6, 80-85.

[18] Wakker, P. (1989), Additive Representations of Preferences, A New Foundation of Decision Analysis, Kluwer Academic Publishers, Dordrecht, The Netherlands.

5 Appendix

For outcomes $x, y, z, w \in X$ and a nonempty subset $T \subset N$ we say that *the tradeoffs* $x \ominus y$ *and* $z \ominus w$ *are comparable according to* $\succsim^T$ or that they are $\succsim^T$*-comparable* whenever either $x \ominus y \succeq^T z \ominus w$ or $z \ominus w \succeq^T x \ominus y$. Note that $x \ominus y$ and $z \ominus w$ for $x \sim^T y \sim^T z \sim^T w$ are always comparable, satisfying $x \ominus y \simeq^T z \ominus w$, by considering the $\succsim^T$-indifferent x, x, x, x, which trivially satisfy the condition in Definition 1. For a set E we denote the interior of E by $int(E)$, and the closure of E by $cl(E)$.

5.1 Proof of Proposition 1

By the basic assumptions there exist continuous utility functions representing $\succsim^G$ and $\succsim^T$. Denote them u_G and u_T, respectively. The proof of the proposition starts by mapping each point $x \in X$ to its utilities image, $(u_G(x), u_T(x))$, and working with the standard topology on $\mathbb{R}^2$. An additive representation $U_{G\cup T} = U_G + U_T$ of $\succsim^{G\cup T}$ is derived in this utilities space, by considering the binary relations induced there by the relations $\succsim^G, \succsim^T$, and $\succsim^{G\cup T}$. For the relations $\succsim^G$ and $\succsim^T$, the induced relations in utilities space simply translate to the real order on each of the two axes. The induced joint relation, also denoted $\succsim^{G\cup T}$, is defined by $(u_G(x), u_T(x)) \succsim^{G\cup T} (u_G(y), u_T(y)) \iff x \succsim^{G\cup T} y$. It is well defined owing to Extended Pareto (**A2**), which asserts that for two outcomes x and y, $u_G(x) = u_G(y)$ and $u_T(x) = u_T(y)$ implies $x \sim^{G\cup T} y$.

To obtain an additive representation in utilities space, Corollary 2.3 (and the related Theorem 3.3) of Chateauneuf and Wakker [4] is applied on the utilities image of X under $(u_G(\cdot), u_T(\cdot))$, namely on

$$E = \{ \ (u_G(x), u_T(x)) \mid x \in X \ \} \ .$$

Corollary 2.3 relies on an assumption, marked 2.1 in that paper, stating the connectedness of various sets, and requiring the relation $\succsim^{G\cup T}$ to satisfy monotonicity and continuity. Monotonicity of $\succsim^{G\cup T}$ is with respect to the coordinate relations, in our case monotonicity with respect to $\succsim^{G}$ and $\succsim^{T}$, which follows from Extended Pareto (**A2**). Lemmas 1-7 below establish all the other components of Assumption 2.1, as well as the additional assumption in Theorem 3.3, in Chateauneuf and Wakker [4], are satisfied.

The first lemma proves that the image of any connected set $F \subset E$ under the inverse correspondence $(u_G, u_T)^{-1}$ is connected as well. This follows from the correspondence $(u_G, u_T)^{-1}$, from E to X, being upper semicontinuous, together with a result stating that the image of a connected set under an upper semicontinuous correspondence is also connected.

Lemma 1. *Let $F \subset E$ be a connected set in the relative standard topology on E. Then $(u_G, u_T)^{-1}(a, t)$ is connected with respect to the topology over X induced by the orders $\succsim^{G}$ and $\succsim^{T}$.*

Proof. Upper semicontinuity of $(u_G, u_T)^{-1}$ requires that for every $(a, t) \in E$ and every open subset $O \subset X$ containing $(u_G, u_T)^{-1}(a, t)$, there exist a neighborhood of (a, t), $N(a, t)$, such that for every $(b, s) \in N(a, t)$, $(u_G, u_T)^{-1}(b, s) \subseteq O$. It suffices to show the desired inclusion for each of the sets of the forms: $\{z \in X \mid z \succ^G x\}, \{z \in X \mid z \succ^T x\}, \{z \in X \mid x \succ^G z\}$, and $\{z \in X \mid x \succ^T z\}$.

Fix $w \in X$, consider the open set $O = \{z \in X \mid z \succ^G w\}$, and assume $(u_G, u_T)^{-1}(a, t) \subset O$. It implies that $a > u_G(w)$. Set $N(a, t) = \{(b, s) \mid b > u_G(w)\}$, so $N(a, t)$ is a neighborhood of (a, t). For every $(b, s) \in N(a, t)$, each $y \in (u_G, u_T)^{-1}(b, s)$ satisfies $y \succ^G w$. Therefore, $(u_G, u_T)^{-1}(b, s) \subset O$. The same arguments prove the inclusion for $O = \{z \in X \mid w \succ^G z\}$ and for the same open sets with $\succ^T$ instead of $\succ^G$. It follows that the same holds for any open set in the order topology of X, generated by those sets. Consequently $(u_G, u_T)^{-1}$ is upper semicontinuous, and by a result from Hiriart-Urruty [9], the image of any connected set $F \subset E$ under it is connected as well. ■

Connectedness of inverse images under $(u_G, u_T)^{-1}$ lets us prove the following useful lemma.

Lemma 2. *For every $x \in X$ which is neither unanimously maximal nor unanimously minimal, the sets $\{(u_G(y), u_T(y)) \mid y \in X, u_G(y) = u_G(x)\}$ and $\{(u_G(y), u_T(y)) \mid y \in X, u_T(y) = u_T(x)\}$ are non-degenerate intervals.*

Proof. Our structural assumption states that for every $x \in X$, the sets $\{y \in X \mid y \sim^G x\}$ and $\{y \in X \mid y \sim^T x\}$ are connected. The utilities u_G and u_T are continuous, therefore the above sets of utility values, contained in $\mathbb{R}$, are connected. Namely, these sets are intervals. In order to prove that these intervals are non-degenerate, it is shown that Diversity (**A4**) implies that for every $x \in X$ that is neither unanimously maximal not unanimously minimal, there exist z', z'' such that $z' \sim^G x$ and $z'' \sim^T x$, while $z'' \nsim^G x$ and $z' \nsim^T x$.

Let $x \in X$. Diversity states that there exists $y \in X$ such that $y \succ^G x$ and $x \succ^T y$, or $y \succ^T x$ and $x \succ^G y$. Suppose, w.l.o.g., that the former holds. The set $\{z \in X \mid z \succ^G x\}$ is the image under $(u_G, u_T)^{-1}$ of the set $F = \{(a,t) \in E \mid a > u_G(x)\}$. According to Lemma 1, if F is connected according to the relative standard topology of E, then $\{z \in X \mid z \succ^G x\}$ is connected as well. It is now shown by negation that F is indeed connected.

Suppose on the contrary that there are two sets, A and B, closed in the relative standard topology of F, which union is F, and intersection is empty. Recall that E is connected, hence $cl(A \cup (E \setminus F))$ and $cl(B)$ must have a nonempty intersection, as these are two closed sets in the relative topology of E, which union is the entire E. The same is true for $cl(B \cup (E \setminus F))$ and $cl(A)$. Points in both these intersections must have their first coordinate equal $u_G(x)$. It follows that there is a sequence of points in B, $((b_n^B, s_n^B))_n$, which converges to a point $(u_G(x), s_B)$, and a sequence of points in A, $((b_n^A, s_n^B))_n$, converging to a point $(u_G(x), s_A)$. We generate a contradiction by proving that A and B must contain points with the same first coordinate, which by connectedness of the images of $\succsim^G$-indifference classes imply that the entire interval between those points must be contained in E, contradicting the fact that A and B are closed in the relative topology of F, and disjoint.

Suppose that A contains two points, (b_m^A, s_m^A) and (b_n^A, s_n^A), for $u_G(x) < b_n^A < b_m^A$. If A does *not* include points (b_k^A, s_k^A) for $b_n^A < b_k^A < b_m^A$, then E can be partitioned into the sets $A \cap \{(b,s) \mid b \geq \frac{2b_m^A + b_n^A}{3}\}$ and $(E \setminus F) \cup B \cup (A \cap \{(b,s) \mid b \leq \frac{b_m^A + 2b_n^A}{3}\})$, which are two disjoint, relatively closed sets which union is E. This is a contradiction to the connectedness of E. Hence whenever A contains points (b_m^A, s_m^A) and (b_n^A, s_n^A) then it also contains points (b_k^A, s_k^A), for every $b_n^A < b_k^A < b_m^A$. The same holds for B.

Let $\varepsilon > 0$, then there exists a point $(b_m^A, s_m^A) \in A$ such that $0 < b_m^A - u_G(x) < \varepsilon$, and in the same manner there are points $(b_n^A, s_n^A) \in A$ and $(b_r^B, s_r^B), (b_\ell^B, s_\ell^B) \in B$ that satisfy $0 < b_r^B - u_G(x) < b_n^A - u_G(x) < b_\ell^B - u_G(x) < b_m^A - u_G(x)$. According to the previous paragraph, it follows that both A and B contain points (b,s) for every $b_n^A \leq b \leq b_\ell^B$. Set one such value b, let $(b,s') \in A$ and $(b,s'') \in B$, and suppose w.l.o.g. that $s' \leq s''$. As was proved above, $\{(u_G(y), u_T(y)) \mid y \in X, u_G(y) = b\}$ is an interval (b is obtained as a first coordinate in E, therefore it is $u_G(z)$ for some z). As a result, F must include the entire interval $\{(b,s) \mid s' \leq s \leq s''\}$ (where $b > u_G(x)$), contradicting the choice of A and B as disjoint and closed in the relative topology of F. It is concluded that F is connected, hence $\{z \in X \mid z \succ^G x\}$, its inverse image under $(u_G, u_T)^{-1}$, is connected as well, following Lemma 1.

Consider the division of the connected $\{z \in X \mid z \succ^G x\}$ into sets, $\{z \in X \mid x \succsim^T z\} \cap \{z \in X \mid z \succ^G x\}$ and $\{z \in X \mid z \succsim^T x\} \cap \{z \in X \mid z \succ^G x\}$. These two sets are closed in the relative topology of $\{z \in X \mid z \succ^G x\}$ and their union is the entire $\{z \in X \mid z \succ^G x\}$. The first set is nonempty, as we assumed that the former option of Diversity is satisfied, that is, that there exists $y \in X$ satisfying $y \succ^G x$ and $x \succ^T y$. The second set is nonempty following Agreed Improvement (**A3**). It follows that the intersection of these two sets is nonempty, namely, that there exists $z' \in X$ such that $z' \succ^G x$ and $z' \sim^T x$. In the same manner, the set $\{z \in X \mid x \succ^T z\}$ can be shown to be connected, and by analogously considering the sets $\{z \in X \mid x \succsim^G z\} \cap \{z \in X \mid x \succ^T z\}$ and $\{z \in X \mid z \succsim^G x\} \cap \{z \in X \mid x \succ^T z\}$ it is established that there exists $z'' \in X$ such that $z'' \sim^G x$ and $x \succ^T z''$.

If the second option of diversity holds, meaning that there is $y \in X$ such that $x \succ^G y$ and $y \succ^T x$, then the proof may be repeated with connected sets $\{z \in X \mid x \succ^G z\}$ and $\{z \in X \mid z \succ^T x\}$. In any case, it results that for every $x \in X$ there exist $z', z'' \in X$ such that $z' \nsim^G x$, $z' \sim^T x$, $z'' \sim^G x$, $z'' \nsim^T x$.

Now return to the set $\{(u_G(y), u_T(y)) \mid y \in X, u_G(y) = u_G(x)\}$ for $x \in X$. It was already proved that this set is an interval. As there is $z'' \in X$ such that $z'' \sim^G x$ but $z'' \nsim^T x$, this interval is non-degenerate. The same argument leads to that conclusion for the other interval.

■

Lemma 3. $E \subseteq cl(int(E))$.

Proof. We show that for each point (a, t) in E, which is neither the utilities image of a unanimously maximal outcome, nor the utilities image of a unanimously minimal outcome every neighborhood of (a, t) intersects $int(E)$. According to Lemma 2, $\{s \mid (a, s) \in E\}$ is a non-degenerate interval which contains t, and $\{b \mid (b, t) \in E\}$ is a non-degenerate interval containing a. Suppose that on these intervals, there are $s > t$ such that $(a, s) \in E$, and $b > a$ such that $(b, t) \in E$, hence $(a, r) \in E$ for every $t \leq r \leq s$, and $(c, t) \in E$ for every $a \leq c \leq b$.

By Lemma 2, if there is a point (c, r) in E such that $a < c \leq b$, then the interval $\{(c, r') \mid t \leq r' \leq r\}$ is in E, and by applying Lemma 2 to points on this interval with $r' \leq \min(r, s)$, and to points on the interval $\{(a, r') \mid t \leq r' \leq \min(r, s)\}$, all the points in the rectangle which vertices are $(a, t), (a, \min(r, s)), (c, \min(r, s))$ and (c, t) are included in E. Every neighborhood of (a, t) intersects this rectangle, hence every neighborhood of (a, t) contains interior points of E. Symmetrically, this neighborhood exists if there is a point (c, r) in E such that $t < r \leq s$.

Otherwise, there does not exist a point (c, r) with $a < c \leq b$ nor with $t < r \leq s$. However, by **A3**, since (a, s) and (b, t) are neither unanimously minimal nor unanimously maximal, there is (d, ℓ) such that $d > b$ and $\ell > s$, generating a contradiction to connectivity of E. Therefore there must exist a point (c, r) with either $a < c \leq b$ or $t < r \leq s$, and the intersecting neighborhoods as above follow. Similar intersection of neighborhoods with points in $int(E)$ follows if on the intervals going through (a, t) there are points $(a, s), (b, t) \in E$ with $s < t$ and $b < a$ (by using (d, ℓ) such that $d < b$ and $\ell < s$).

Otherwise suppose that on the intervals going through (a, t) there are points $(a, s), (b, t) \in E$ with $s > t$ and $b < a$. Lemma 2 implies that $(a, r) \in E$ for every $t \leq r \leq s$, and $(c, t) \in E$ for every $b \leq c \leq a$. If there exists a point $(c, r) \in E$ with $c < a$ and $r > t$, then by connectivity there must be a point (c, r) with either $b \leq c < a$ or $t < r \leq s$. In that case a rectangle of interior points as above follows, intersecting each neighborhood of (a, t).

Otherwise there is no point $(c, r) \in E$ with $c < a$ and $r > t$. Suppose that there is a point $(c, r) \in E$ with $c > a$ and $r < t$. Then, again by connectivity of E, it cannot be that both rays $\{(d, t) \mid d > a\}$ and $\{(a, \ell) \mid \ell < t\}$ have an empty intersection with E. Hence, applying Lemma 2 once more, there is either an interval $\{(d, t) \mid a \leq d \leq c\}$, for $c > a$, or an interval

$\{(a,\ell) \mid r \leq \ell \leq t\}$, for $r < t$, which is contained in E, and the desired neighborhoods follow as in the cases above.

Lastly, the case where there is neither a point (c,r) with $c < a$ and $r > t$, nor a point (c,r) with $c > a$ and $r < t$, is excluded owing to part (b) of Diversity (**A4(b)**).

It is concluded that in any case, every neighborhood of (a,t) intersects $int(E)$. Denote by E^* the set E excluding the utilities image of any unanimously minimal and any unanimously maximal outcomes, if such outcomes exist. Therefore $E^* \subseteq cl(int(E))$. Following non-degeneracy (**A1(d)**) and connectedness of E, E^* is non-empty, and every neighborhood of the utilities image of a unanimously maximal or unanimously minimal outcome contains the utilities image of an outcomes which is neither unanimously maximal nor unanimously minimal. Therefore every neighborhood of the utilities image of a unanimously maximal or unanimously minimal outcome also intersects $int(E)$, and $E \subseteq cl(int(E))$. ■

Lemma 4. *$int(E)$ is connected.*

Proof. First note that since X is connected, and the mapping (u_G, u_T) is continuous, then E is connected as well. Suppose on the contrary that there are two non-empty sets, A and B, that are open in the relative (standard) topology of $int(E)$ (hence open), and that partition $int(E)$. By Lemma 3, $E \subseteq cl(int(E)) = cl(A) \cup cl(B)$. Since E is connected and non-empty, the intersection $cl(A) \cap cl(B) \cap E$ is nonempty.

Lemma 2 states that for each $x \in X$, the images of x's indifference classes according to $\succsim^G$ and according to $\succsim^T$ are each an interval. Therefore if $(a,t) \in A$ then neither (a,s) nor (b,t) are in B (as the entire interval between each two points $(b,t') \in E$ and $(b,s') \in E$ in the neighborhoods of (a,t) and (a,s) is also contained in E). Denote by (b,s) a point in $cl(A) \cap cl(B) \cap E$. If we consider the four orthants around (b,s), then by the previous argument A and B must be contained in opposite orthants relative to (b,s), hence their closures reside on opposite orthants, including possibly the lines $\{(a,t) \mid a = b\}$ and $\{(a,t) \mid t = s\}$. Note that as a result (b,s) is neither a unanimously maximal nor a unanimously minimal outcome, since there must be points in either A or B with one of their coordinates larger than the corresponding coordinate of (b,s), and points in either A or B with one or their coordinates smaller than the corresponding coordinate of (b,s). As $cl(A)$ and $cl(B)$ are contained in opposite orthants, there is either no point $(a,t) \in E$ with $a > b$ and $t > s$, or there is no point $(a,t) \in E$ with $a > b$ and $t < s$. The first possibility contradicts **A3** (by setting $x = y$ in that assumption, with $b = u_G(x), s = u_T(x)$), and the second possibility contradicts **A4(b)**. It is concluded that sets A, B as supposed cannot exist. Namely, $int(E)$ is connected. ■

Lemma 5. *Any utility indifference class of $\succsim^{G\cup T}$ within $int(E)$, $\{(u_G(y), u_T(y)) \mid y \sim^{G\cup T} x\} \cap int(E)$ for every $x \in X$, is connected.*

Proof. Let $x \in X$. The structural assumption (**A0(b)**) guarantees that the indifference class of $\succsim^{G\cup T}$ containing x, in outcome space, is connected. Hence its image under the continuous mapping (u_G, u_T) is connected, in E. It should be established that it is also connected in $int(E)$. For that, note first that the image of x's indifference class according to $\succsim^{G\cup T}$ is a curve. This is since for every point (a,t) in this image, Extended Pareto (**A2**) implies that any point of the form (b,s), with $b \geq a, s \geq t$ or $b \leq a, s \leq t$, is not contained in the image of the same indifference class. Moreover, it follows that this curve is strictly decreasing.

It is next proved that any point in the relative interior of the image of the indifference class of x is an interior point of E. Let $(u_G(y), u_T(y))$ be in the relative interior of this image. Namely, there are $z', z'' \in X$ such that $z', z'' \sim^{G\cup T} x$, with their utilities satisfying $u_G(z') < u_G(y), u_T(z') > u_T(y)$ and $u_G(z'') > u_G(y), u_T(z'') < u_T(y)$. Following Lemma 2, the points in E which first coordinate is $u_G(y)$ form a non-degenerate interval, and similarly the points in E which second coordinate is $u_T(y)$. The proof continues by addressing all possible locations of $(u_G(y), u_T(y))$ within these intervals.

<u>Case 1.</u>
There are $y_1, y_2, y_3, y_4 \in X$ such that,

$$\begin{aligned}
u_G(y_1) > u_G(y) \quad &, \quad u_T(y_1) = u_T(y) \\
u_G(y_2) = u_G(y) \quad &, \quad u_T(y_2) > u_T(y) \\
u_G(y_3) < u_G(y) \quad &, \quad u_T(y_3) = u_T(y) \\
u_G(y_4) = u_G(y) \quad &, \quad u_T(y_4) < u_T(y).
\end{aligned}$$

As the curve $\{(u_G(y), u_T(y)) \mid y \sim^{G\cup T} x\}$ is connected, it contains a point (a,t) with $a > u_G(y)$ that is still smaller than both $u_G(z'')$ and $u_G(y_1)$. Following the same arguments as in the proof of Lemma 3, all points (b,s) with $u_G(y) \leq b \leq a$ and $t \leq s \leq u_T(y)$ are in E. Similarly, there is a rectangle of points with first coordinate smaller than $u_G(y)$, and second coordinate larger than $u_T(y)$, which is contained in E. Furthermore, in a similar manner to the proof of Lemma 3, connectivity implies that there is a point $(a,t) \in E$ with either $u_G(y) < a \leq u_G(y_1)$ or $u_T(y) < t \leq u_T(y_2)$, implying that there is a neighborhood of $(u_G(y), u_T(y))$ which intersection with $\{(b,s) \mid b > u_G(y), s > u_T(y)\}$ is contained in E. The same arguments imply such a neighborhood which intersection with $\{(b,s) \mid b < u_G(y), s < u_T(y)\}$ is contained in E. Altogether, there is a neighborhood of $(u_G(y), u_T(y))$ that is contained in E, and $(u_G(y), u_T(y))$ is an interior point of E.

<u>Case 2.</u>

There are $y_1, y_2, y_3 \in X$ such that,

$$\begin{aligned}
&u_G(y_1) > u_G(y)\ ,\ u_T(y_1) = u_T(y),\\
&u_G(y_2) = u_G(y)\ ,\ u_T(y_2) > u_T(y),\\
&u_G(y_3) < u_G(y)\ ,\ u_T(y_3) = u_T(y),\ \text{and}\\
&u_T(y) = \min\{t\ |(u_G(y), t) \in E\}\ .
\end{aligned}$$

The arguments employed in Case 1 may be repeated to obtain that there is a neighborhood of $(u_G(y), u_T(y))$, which intersection with $\{(b, s) \mid s > u_T(y)\}$ is contained in E. Denote the supremum of the first coordinate in this neighborhood by a. Then Lemma 2 further implies that all points (b, s) with $u_G(y) \leq b \leq \min(a, u_G(z''))$, $s < u_T(y)$, with (b, s) that is above the curve $\{(u_G(z), u_T(z)) \mid z \sim^{G \cup T} x\}$, belong to E as well.

Case 2 states that $u_T(y) = \min\{t\ |(u_G(y), t) \in E\}$. In other words, there are no points $(b, s) \in E$ with $b = u_G(y)$ and $s < u_T(y)$. Hence there are no points $(b, s) \in E$ with $b < u_G(y)$ and $u_T(z'') \leq s < u_T(y)$ (again following Lemma 2, by negation – otherwise there would be an interval in E, with second coordinate s, and first coordinate between b and the curve, going through $(u_G(y), s), s < u_T(y)$). On the other hand, it follows from Agreed Improvement (**A3**) that there exists a point $(u_G(\hat{z}), u_T(\hat{z}))$ in E such that $u_G(\hat{z}) < u_G(y)$, and $u_T(\hat{z}) < u_T(y)$. This assumption, and the lack of points in E within $\{(b, s) \mid b < u_G(y), u_T(z'') \leq s < u_T(y)\}$ and $\{(u_G(y), s) \mid s < u_T(y)\}$, generate a contradiction to the connectivity of E. Therefore there must be a point y_4 with $u_G(y_4) = u_G(y)$, and $u_T(y_4) < u_T(y)$, as in the first case, and the same conclusion follows. Symmetric cases, where either $u_T(y) = \max\{t\ |(u_G(y), t) \in E\}$, $u_G(y) = \min\{a\ |(a, u_T(y)) \in E\}$, or $u_G(y) = \max\{a\ |(a, u_T(y)) \in E\}$, are proved in the same manner.

<u>Case 3.</u>
There are $y_1, y_2 \in X$ such that,

$$\begin{aligned}
&u_G(y_1) > u_G(y)\ ,\ u_T(y_1) = u_T(y),\\
&u_G(y_2) = u_G(y)\ ,\ u_T(y_2) > u_T(y),\\
&u_G(y) = \min\{a\ |(a, u_T(y)) \in E\},\ \text{and}\\
&u_T(y) = \min\{t\ |(u_G(y), t) \in E\}\ .
\end{aligned}$$

This case is immediately eliminated, seeing that it inflicts a contradiction to connectivity, through the assumption that there is a point $(u_G(\hat{z}), u_T(\hat{z}))$ in E such that $u_G(\hat{z}) < u_G(y)$, and $u_T(\hat{z}) < u_T(y)$ (a result of **A3**). This is since y cannot be a unanimously minimal outcome, as it was assumed that there are outcomes z', z'' with $u_G(z') < u_G(y)$. Similarly, the case $u_G(y) = \max\{a\ |(a, u_T(y)) \in E\}$ and $u_T(y) = \max\{t\ |(u_G(y), t) \in E\}$ is ruled out.

Case 4.
There are $y_2, y_3 \in X$ such that,

$$\begin{aligned}
&u_G(y_2) = u_G(y) \ , \ u_T(y_2) > u_T(y), \\
&u_G(y_3) < u_G(y) \ , \ u_T(y_3) = u_T(y), \\
&u_G(y) = \max\{a \mid (a, u_T(y)) \in E\}, \text{ and} \\
&u_T(y) = \min\{t \mid (u_G(y), t) \in E\} \ .
\end{aligned}$$

As in the previous cases, there is a neighborhood of $(u_G(y), u_T(y))$ which intersection with $\{(b, s) \mid b < u_G(y), s > u_T(y)\}$ is contained in E. Within this case it is supposed that points $(u_G(y), s)$, $s < u_T(y)$, and $(b, u_T(y))$, $b > u_G(y)$, do not belong to E. As $u_G(z'') > u_G(y)$ and $u_T(z'') < u_T(y)$, and the curve $\{(u_G(z), u_T(z)) \mid z \sim^{G\cup T} x\}$ is connected, then it must also hold (employing again Lemma 2) that E contains no points (b, s) with $b < u_G(y)$ and $u_T(z'') \leq s < u_T(y)$. Similarly to Case 2, excluding all those points from E, while requiring part (b) of the structural assumption (**A0**), cannot be reconciled with connectivity of E. The case $u_G(y) = \min\{a \mid (a, u_T(y)) \in E\}$ and $u_T(y) = \max\{t \mid (u_G(y), t) \in E\}$ is analogous.

It is concluded that Case 1 is the only one that could hold for a point $(u_G(y), u_T(y))$ in the relative interior of $\{(u_G(z), u_T(z)) \mid z \in X, z \sim^{G\cup T} x\}$. Therefore, as was proved in Case 1, any such relative interior point $(u_G(y), u_T(y))$ is an interior point of E, hence $\{(u_G(y), u_T(y)) \mid y \sim^{G\cup T} x\} \cap int(E)$, for every $x \in X$, contains the entire relative interior of the connected curve $\{(u_G(y), u_T(y)) \mid y \sim^{G\cup T} x\}$, and is thus connected.

■

Lemma 6. *For every $x \in X$, the sets $\{(u_G(y), u_T(y)) \mid u_G(y) = u_G(x)\} \cap int(E)$, and $\{(u_G(y), u_T(y)) \mid u_T(y) = u_T(x)\} \cap int(E)$, are connected.*

Proof. First note that without intersecting with $int(E)$, the sets $\{(u_G(y), u_T(y)) \mid u_G(y) = u_G(x)\}$ and $\{(u_G(y), u_T(y)) \mid u_T(y) = u_T(x)\}$ are connected, being the image under a continuous function of a connected set (in outcomes space; Following assumption **A0(b)**). Suppose that there are outcomes $y, y', y'' \in X$ with $u_G(y) = u_G(y') = u_G(y'') = u_G(x)$ and $u_T(y') > u_T(y) > u_T(y'')$, where $(u_G(y'), u_T(y'))$ and $(u_G(y''), u_T(y''))$ are interior points of E. It is proved that $(u_G(y), u_T(y))$ must also be an interior point of E.

If $(u_G(y'), u_T(y'))$ and $(u_G(y''), u_T(y''))$ are interior points of E, then there is (a minimal) $\varepsilon > 0$ such that the sets,

$$\begin{aligned}
&\{(u_G(y'), u_T(y')) + \tau(\cos\theta, \sin\theta) \mid 0 < \tau < \varepsilon, \theta \in [0, 2\pi]\} \\
&\{(u_G(y''), u_T(y'')) + \tau(\cos\theta, \sin\theta) \mid 0 < \tau < \varepsilon, \theta \in [0, 2\pi]\}
\end{aligned}$$

are contained in E. However, recall that according to Lemma 2, if there are points (a, t) and (a, s) both belonging to E, then any point (a, r) with r between t and s also belongs to

E. Hence for each $0 < \tau < \varepsilon$ and each $\theta \in [0, 2\pi]$, $(u_G(y') + \tau\cos\theta, u_T(y') + \tau\sin\theta)$ and $(u_G(y'') + \tau\cos\theta, u_T(y'') + \tau\sin\theta)$ both belonging to E, and the fact that $u_G(y') = u_G(y'') = u_G(y)$, implies that $(u_G(y) + \tau\cos\theta, u_T(y) + \tau\sin\theta)$ belongs to E as well. Therefore there is a neighborhood of $(u_G(y), u_T(y))$ included in E, namely, $(u_G(y), u_T(y))$ is an interior point of E.

The above implies that there cannot be a boundary point between two interior points in the set $\{(u_G(y), u_T(y)) \mid u_G(y) = u_G(x)\}$. That is to say, if there is a boundary point then all points to (at least) one side of it must also be boundary points. Hence, the intersection of this image with $int(E)$ is connected. Analogously, the same holds for $\{(u_G(y), u_T(y)) \mid u_T(y) = u_T(x)\} \cap int(E)$. ■

Lemma 7. *$\succsim^{G\cup T}$ on E is continuous.*

Proof. In order to prove the desired continuity, it should be proved that for every $(a, t) \in E$, the sets $\{(b, s) \in E \mid (b, s) \succ^{G\cup T} (a, t)\}$ and $\{(b, s) \in E \mid (a, t) \succ^{G\cup T} (b, s)\}$ are open in the relative (standard) topology of E. Let $(a, t) \in E$. It was proved in Lemma 5 that $\{(b, s) \in E \mid (b, s) \sim^{G\cup T} (a, t)\}$ is a strictly downward sloping curve, and that any point in its relative interior is an interior point of E. It is now shown that the opposite is also true. That is to say, if a point $(c, r) \in \{(b, s) \in E \mid (b, s) \sim^{G\cup T} (a, t)\}$ is a relative boundary point of this curve, then it is also a boundary point of E. Specifically, we next prove that for the upper most point on this curve there is no point in E with a smaller first coordinate and larger second coordinate, and for the lowest point on this curve there is no point in E with a larger first coordinate and a smaller second coordinate.

Suppose a point (c, r) such that $(c, r) \sim^{G\cup T} (a, t)$, and for which there is no other point $(c', r') \sim^{G\cup T} (a, t)$ with $c' > c$ and $r' < r$. We prove by negation that in that case, there can be no points $(c', r') \in E$ with $c' > c$ and $r' < r$. Suppose on the contrary that there was such a point (c', r'). Then it cannot hold that both rays, $\{(b, r) \mid b > c\}$ and $\{(c, s) \mid s < r\}$ have an empty intersection with E, because that would generate a contradiction to the connectivity of E. Therefore (also employing Lemma 2) suppose that there is an interval $\{(c, s) \mid \ell \le s \le r\}$, for $\ell < r$, which is contained in E. Still, for $\{(b, r) \mid b > c\}$ not to intersect E it should also hold that the ray $\{(c', s) \mid r < s\}$ has an empty intersection with E. However, as (c', r') is not unanimously maximal, that would again generate a contradiction to the connectivity of E, based on Agreed Improvement (**A3**). Supposing first a nonempty intersection with $\{(b, r) \mid b > c\}$ yields an analogous contradiction.

The previous paragraph implies that there are intervals, $\{(c, s) \mid \ell \le s \le r\}$ for $\ell < r$, and $\{(b, r) \mid c \le b \le d\}$ for $d > c$, contained in E. Same as in the proof of Lemma 3, there exists a rectangle with vertices $(c, r), (d', r), (b, \ell'), (c, \ell')$, for $d' > c$ and $\ell' < r$, which is contained in E. Denote $R = \{(b, s) \mid c \le b \le d', \ell' \le s \le r\}$. To generate a contradiction it is shown that $R \setminus \{(c, r)\}$ must contain a point $(c', r') \sim^{G\cup T} (c, r)$, contradicting the choice of (c, r).

To show that a contradiction ensues, it is next proved that $(u_G, u_T)^{-1}(R \setminus \{(c, r)\})$ is connected. For that, we show that the correspondence $(u_G, u_T)^{-1}$, from E to X, is upper semicontinuous, and employ a result stating that the image of a connected set under an upper semicontinuous correspondence is also connected. Upper semicontinuity of $(u_G, u_T)^{-1}$ requires that for every $(a,t) \in E$ and every open subset $O \subset X$ containing $(u_G, u_T)^{-1}(a,t)$, there exist a neighborhood of (a,t), $N(a,t)$, such that for every $(b,s) \in N(a,t)$, $(u_G, u_T)^{-1}(b,s) \subseteq O$. It suffices to show the desired inclusion for sets $\{z \in X \mid z \succ^G x\}, \{z \in X \mid z \succ^T x\}, \{z \in X \mid x \succ^G z\}$, and $\{z \in X \mid x \succ^T z\}$.

Consider an open set $O = \{z \in X \mid z \succ^G w\}$ for $w \in X$, such that $(u_G, u_T)^{-1}(a,t) \subset O$, hence $a > u_G(w)$. Set $N(a,t) = \{(b,s) \mid b > u_G(w)\}$, so $N(a,t)$ is a neighborhood of (a,t), and for every $(b,s) \in N(a,t)$, every $y \in (u_G, u_T)^{-1}(b,s)$ satisfies $y \succ^G w$, therefore $(u_G, u_T)^{-1}(b,s) \subset O$. The same arguments prove the inclusion for $O = \{z \in X \mid w \succ^G z\}$ and for the same open sets with $\succ^T$ instead of $\succ^G$. It follows that the same holds for any open set in the order topology of X, generated by those sets. Consequently $(u_G, u_T)^{-1}$ is upper semicontinuous, and by a result from Hiriart-Urruty [9], the image of the connected $R \setminus \{(c,r)\}$ under it is connected as well.

Let $x, y \in X$ be such that $u_G(x) = c, u_T(x) = r, u_G(y) = d', u_T(y) = \ell'$. Then,

$$(u_G, u_T)^{-1}(R \setminus \{(c,r)\}) = \{z \in X \mid y \succ^G z \succ^G x, x \succ^T z \succ^T y\} \setminus \{z \in X \mid z \sim^G x, z \sim^T x\} \ .$$

Denote,

$$\begin{aligned} A &= \{z \in X \mid z \succsim^{G \cup T} x\} \cap (u_G, u_T)^{-1}(R \setminus \{(c,r)\}) \\ B &= \{z \in X \mid x \succsim^{G \cup T} z\} \cap (u_G, u_T)^{-1}(R \setminus \{(c,r)\}) \ . \end{aligned}$$

A and B are closed in the relative topology of $R \setminus \{(c,r)\}$. They are nonempty since for $y' \in (u_G, u_T)^{-1}(d', r)$, $y' \succ^{G \cup T} x$, and for $y'' \in (u_G, u_T)^{-1}(c, \ell')$, $x \succ^{G \cup T} y''$, and their union is $R \setminus \{(c,r)\}$. Hence, by connectedness of $R \setminus \{(c,r)\}$, their intersection must be nonempty. Namely, there should exist $z \in X$ such that $z \in (u_G, u_T)^{-1}(R \setminus \{(c,r)\})$ and $z \sim^{G \cup T} x$, generating a contradiction to the assumption that no point $(c', r') \sim^{G \cup T} (c,r)$ satisfies $c' > c$ and $r' < r$. It is concluded that there can be no point $(b,s) \in E$ such that $b > c$ and $s < r$. Analogue arguments prove that whenever $(c,r) \sim^{G \cup T} (a,t)$ is such that no $(c', r') \sim^{G \cup T} (a,t)$ satisfies $c' < c, r' > r$, then there is no point $(b,s) \in E$ such that $b < c, s > r$.

The previous arguments imply that the indifference set $\{(b,s) \in E \mid (b,s) \sim^{G \cup T} (a,t)\}$ is a downward sloping curve, dividing E into two parts, one above and one below that curve. Therefore, $\{(b,s) \in E \mid (b,s) \succ^{G \cup T} (a,t)\}$ consists of all points in E which are strictly above that curve. More formally, to establish that this is indeed an open set, denote by (c', r') the point on this curve with maximal second coordinate (the upper-most point on the curve), and by (c'', r'') the point on this curve with maximal first coordinate (the lowest point on the curve). Consider the (continuous) curve in $\mathbb{R}^2$,

$$\begin{aligned} \mathcal{L} \;=\; & \{(b,s) \in (R)^2 \mid b < c', s > r', b+s = c'+r'\} \cup \\ & \{(b,s) \in E \mid (b,s) \sim^{G\cup T} (a,t)\} \cup \\ & \{(b,s) \in (R)^2 \mid b > c'', s < r'', b+s = c''+r''\} \ . \end{aligned}$$

Then,

$$\{(b,s) \in E \mid (b,s) \succ^{G\cup T} (a,t)\} = \left(\bigcup_{(b,s)\in\mathcal{L}} \{(d,\ell) \in \mathcal{R}^2 \mid d > b, \ell > s\} \right) \bigcap E \ ,$$

which is open in the relative standard topology of E. Similarly, the set $\{(b,s) \in E \mid (a,t) \succ^{G\cup T} (b,s)\}$ consists of all point in E which are strictly below that curve, hence it is an open set. As a result, $\succsim^{G\cup T}$ is continuous on E. ■

The above lemmas establish that all the components of Assumption 2.1, as well as the additional assumption in Theorem 3.3, in Chateauneuf and Wakker [4], are satisfied. To derive the desired representation through Theorem 3.3 and Corollary 2.3 of [4], it is required that the relation in question satisfy an additivity condition. According to Theorem III.6.6 in Wakker [18] it suffices to show that $\succsim^{G\cup T}$ over E satisfies the Reidemeister condition. This is a straightforward implication of Consistency of Social Tradeoffs (**A5**), shown in the lemma below.

Lemma 8. *Let $(a,t),(b,s),(c,t),(d,s),(a,t'),(b,s'),(c,t'),(d,s') \in E$. If $(a,t) \sim^{G\cup T} (b,s)$, $(d,s) \sim^{G\cup T} (c,t)$, and $(b,s') \sim^{G\cup T} (a,t')$, then $(d,s') \sim^{G\cup T} (c,t')$.*

Proof. Being in E, there are outcomes $x,y,z,w,x',y',z',w' \in X$ such that,

$$\begin{aligned} a = u_G(x) \quad &, \quad t = u_T(x) \\ b = u_G(y) \quad &, \quad s = u_T(y) \\ c = u_G(z) \quad &, \quad t = u_T(z) \\ d = u_G(w) \quad &, \quad s = u_T(w) \\ a = u_G(x') \quad &, \quad t' = u_T(x') \\ b = u_G(y') \quad &, \quad s' = u_T(y') \\ c = u_G(z') \quad &, \quad t' = u_T(z') \\ d = u_G(w') \quad &, \quad s' = u_T(w') \end{aligned}$$

Equivalence of utility values implies,

$$\begin{aligned} & x \sim^G x', y \sim^G y', z \sim^G z', w \sim^G w', \text{ and} \\ & x \sim^T z, y \sim^T w, x' \sim^T z', \text{and } y' \sim^T w' \ , \end{aligned}$$

and by the induced relation $\succsim^{G\cup T}$ on utility pairs, $x \sim^{G\cup T} y$, $w \sim^{G\cup T} z$, and $y' \sim^* x'$. Therefore Consistency of Social Tradeoffs (**A5**) renders $w' \sim^{G\cup T} z'$, which translates to the required $(d, s') \sim^{G\cup T} (c, t')$. ■

Corollary 2.3 of [4] can now be applied. We conclude that there are $V_G, V_T : \mathbb{R} \longrightarrow \mathbb{R}$, which are continuous and strictly increasing, representing the real order over the x-axis and y-axis, respectively. They form an additive representation $V(a,t) = V_G(a)+V_T(t)$ of $\succsim^{G\cup T}$ on $int(E)$. Furthermore, V_G and V_T are jointly cardinal. According to Lemma 3, $int(E) \subseteq E \subseteq cl(int(E))$, hence Theorem 3.3 of [4] may be applied to extend the additive representation to the entire E.

V_G and V_T are continuous, increasing functions over $\mathbb{R}$, hence $U_G = V_G{\cdot}u_G$ and $U_T = V_T{\cdot}u_T$, their compositions over the continuous utility functions u_G and u_T, are themselves continuous utility functions, representing $\succsim^G$ and $\succsim^T$, respectively. For every $x, y \in X$,

$$x \succsim^{G\cup T} y \;\Leftrightarrow\; (u_G(x), u_T(x)) \succsim^{G\cup T} (u_G(y), u_T(y)) \;\Leftrightarrow$$
$$V_G(u_G(x)) + V_T(u_T(x)) \geq V_G(u_G(y)) + V_T(u_T(y)) \;\; .$$

Namely, for every $x, y \in X$, $x \succsim^{G\cup T} y$, if and only if, $U_G(x) + U_T(x) \geq U_G(y) + U_T(y)$, where U_G and U_T are continuous utility functions, representing the relations $\succsim^G$ and $\succsim^T$, respectively, on X. These functions are jointly cardinal, that is, if $(\hat{U}_G, \hat{U}_T)$ is another additive representation as above, the two must be jointly cardinal, namely $\hat{U}_G = \tau U_G + \rho_G, \hat{U}_T = \tau U_T + \rho_T$, for $\tau > 0$. This establishes Proposition 1.

5.2 Proof of Proposition 2

Lemma 9. *Let $T, G \subset N$ be two nonempty, disjoint sets, and denote,*

$$E^{T,G} = \{(U_T^{T\cup G}(t), U_G^{T\cup G}(t)) \mid t \in X\} \; .$$

Then for any $\theta \in X$ which is neither unanimously maximal for $\succsim^T, \succsim^G$ nor unanimously minimal, the sets,

$$\{(U_T^{T\cup G}(y), U_G^{T\cup G}(y)) \mid U_T^{T\cup G}(y) = U_T^{T\cup G}(\theta)\} \cap int(E^{T,G})$$
$$\{(U_T^{T\cup G}(y), U_G^{T\cup G}(y)) \mid U_G^{T\cup G}(y) = U_G^{T\cup G}(\theta)\} \cap int(E^{T,G})$$
$$\{(U_T^{T\cup G}(y), U_G^{T\cup G}(y)) \mid U_T^{T\cup G}(y) + U_G^{T\cup G}(y) = U_T^{T\cup G}(\theta) + U_G^{T\cup G}(\theta)\} \cap int(E^{T,G})$$

are nonempty.

Proof. Let $\theta \in X$ be an outcome that is neither unanimously maximal nor unanimously minimal for $\succsim^T, \succsim^G$. According to Lemma 2, the utilities indifference curve $\{(U_T^{T\cup G}(y), U_G^{T\cup G}(y)) \mid U_T^{T\cup G}(y) = U_T^{T\cup G}(\theta)\}$ is a non-degenerate interval, and according to Lemma 6 its intersection with $int(E^{T,G})$ is connected, namely, an interval itself. Suppose on the contrary that it is an empty interval. Namely, suppose,

$$\{(U_T^{T\cup G}(y), U_G^{T\cup G}(y)) \mid U_T^{T\cup G}(y) = U_T^{T\cup G}(\theta)\} \cap int(E^{T,G}) = \emptyset \ .$$

Lemmas 3 and 4 assert that $int(E^{T,G})$ is connected, and that $E^{T,G} \subseteq cl(int(E^{T,G}))$, therefore, as indifference curves are intervals, one of two options must hold: either (a) for every $(U_T^{T\cup G}(y), U_G^{T\cup G}(y))$ on this utilities indifference curve there is no $(U_T^{T\cup G}(z), U_G^{T\cup G}(z))$ such that $U_G^{T\cup G}(z) = U_G^{T\cup G}(y)$ and $U_T^{T\cup G}(z) < U_T^{T\cup G}(y) = U_T^{T\cup G}(\theta)$, or (b) for every $(U_T^{T\cup G}(y), U_G^{T\cup G}(y))$ on this utilities indifference curve there is no $(U_T^{T\cup G}(z), U_G^{T\cup G}(z))$ such that $U_G^{T\cup G}(z) = U_G^{T\cup G}(y)$ and $U_T^{T\cup G}(z) > U_T^{T\cup G}(\theta)$. Assume that (a) holds.

Choose two points on the relative interior of this curve (which exist as the curve is a non-degenerate interval), $(U_T^{T\cup G}(y'), U_G^{T\cup G}(y'))$ and $(U_T^{T\cup G}(y''), U_G^{T\cup G}(y''))$, such that $y', y'' \in X$ and $U_T^{T\cup G}(y') = U_T^{T\cup G}(y'') = U_T^{T\cup G}(\theta)$. Being in the relative interior, y' and y'' are neither unanimously minimal nor unanimously maximal, therefore according to Agreed Improvement (**A3**) there is an outcome z_* such that $U_T^{T\cup G}(z_*) < U_T^{T\cup G}(\theta)$ and $U_G^{T\cup G}(z_*) < \min(U_G^{T\cup G}(y'), U_G^{T\cup G}(y''))$. Since for every $(U_T^{T\cup G}(\theta), U_G^{T\cup G}(y)) \in E^{G,T}$ there is no $(U_T^{T\cup G}(z), U_G^{T\cup G}(y))$ such that $U_T^{T\cup G}(z) < U_T^{T\cup G}(\theta)$, then
$U_G^{T\cup G}(z_*) \le \inf_{\{y: U_T^{T\cup G}(y) = U_T^{T\cup G}(\theta)\}} U_G^{T\cup G}(y)$, for a finite infimum. Together with the fact that there are points $(U_T^{T\cup G}(\theta), U_G^{T\cup G}(y)) \in E^{G,T}$ for $U_G^{T\cup G}(y) > \inf_{\{y: U_T^{T\cup G}(y) = U_T^{T\cup G}(\theta)\}} U_G^{T\cup G}(y)$ (again by Agreed Improvement), a contradiction is inflicted, either to the fact that $int(E^{G,T})$ is connected (Lemma 4), or to the fact that $E^{G,T} \subseteq cl(int(E))$ (Lemma 3). If option (b) above holds, the same arguments are applied, only with an outcome z^* used, which dominates the utility images of y' and y'' from above.

It is concluded that $\{(U_T^{T\cup G}(y), U_G^{T\cup G}(y)) \mid U_T^{T\cup G}(y) = U_T^{T\cup G}(\theta)\}$ contains interior points. The proof for $\{(U_T^{T\cup G}(y), U_G^{T\cup G}(y)) \mid U_G^{T\cup G}(y) = U_G^{T\cup G}(\theta)\}$ is analogous.

To show that $\{(U_T^{T\cup G}(y), U_G^{T\cup G}(y)) \mid U_T^{T\cup G}(y) + U_G^{T\cup G}(y) = U_T^{T\cup G}(\theta) + U_G^{T\cup G}(\theta)\} \cap int(E^{T,G})$ note that if θ is neither unanimously maximal nor unanimously minimal for $\succsim^T$ and $\succsim^G$, then the indifference curve of θ according to $\succsim^{T\cup G}$, before intersecting it with $int(E^{T,G})$, is a non-degenerate interval. Therefore the relative interior of this utilities image is non-empty. According to Lemma 5, each such point is an interior point of $E^{T,G}$, and the result follows. ■

According to Proposition 1, for any two non-empty, disjoint sets $T, G \subset N$, there exists a representation $U_{T\cup G} = U_T^{T\cup G} + U_G^{T\cup G}$ of $\succsim^{T\cup G}$ with continuous, jointly cardinal utilities $U_T^{T\cup G}$ and $U_G^{T\cup G}$ that represent $\succsim^T$ and $\succsim^G$, respectively. Likewise, there exists a representation $U_N = U_T^N + U_{N\setminus T}^N$ of $\succsim^N$ with continuous, jointly cardinal utilities U_T^N and $U_{N\setminus T}^N$ that represent $\succsim^T$ and $\succsim^{N\setminus T}$, respectively. It is next proved that $U_T^{T\cup G} = \beta U_T^N + \tau$ for some $\tau, \beta \in \mathbb{R}, \beta > 0$. For that, the next two lemmas show that, locally, whenever utility differences are equal according to U_T^N, then they are equal according to $U_T^{T\cup G}$.

Lemma 10. *Let T, G be two nonempty, disjoint sets, and $\theta \in X$ an outcome such that $M \succ^T \theta \succ^T m$ for some $m, M \in X$. Then there are $y^*, y_* \in X$, satisfying $y^* \succ^T \theta \succ^T y_*$, such that,* *(1) For every* $(u_1,u_2) \in (U_T^{T\cup G}(y_*), U_T^{T\cup G}(y^*)) \times (U_G^{T\cup G}(y_*), U_G^{T\cup G}(y^*))$ *there exists an outcome* $t \in X$ *such that* $U_T^{T\cup G}(t) = u_1$ *and* $U_G^{T\cup G}(t) = u_2$.

(2) For every $\tilde{x}, \tilde{y}, \tilde{z}, \tilde{w} \in X$ *for which* $y^* \succ^T \tilde{x}, \tilde{y}, \tilde{z}, \tilde{w} \succ^T y_*$ *and* $\tilde{x} \succsim^T \tilde{y}$,

$$U_T^{T\cup G}(\tilde{x}) - U_T^{T\cup G}(\tilde{y}) \geq U_T^{T\cup G}(\tilde{z}) - U_T^{T\cup G}(\tilde{w}) \iff \tilde{x} \ominus \tilde{y} \succeq^T \tilde{z} \ominus \tilde{w}$$

(3) For every $\tilde{x}, \tilde{y}, \tilde{z}, \tilde{w} \in X$ *for which* $y^* \succ^G \tilde{x}, \tilde{y}, \tilde{z}, \tilde{w} \succ^G y_*$ *and* $\tilde{x} \succsim^G \tilde{y}$,

$$U_G^{T\cup G}(\tilde{x}) - U_G^{T\cup G}(\tilde{y}) \geq U_G^{T\cup G}(\tilde{z}) - U_G^{T\cup G}(\tilde{w}) \iff \tilde{x} \ominus \tilde{y} \succeq^G \tilde{z} \ominus \tilde{w} \ .$$

Proof. Let $\theta \in X$ be such that $M \succ^T \theta \succ^T m$ for some $m, M \in X$. Consider the utilities image under $U_T^{T\cup G}$ and $U_G^{T\cup G}$,

$$E^{T,G} = \{ \ (U_T^{T\cup G}(t), U_G^{T\cup G}(t)) \mid t \in X \ \} \ .$$

By assumption, θ is neither unanimously maximal nor unanimously minimal for $\succsim^T, \succsim^G$. Thus according to Lemma 2 the image in $E^{T,G}$ of its indifference class under $\succsim^T$, $\{(U_T^{T\cup G}(t), U_G^{T\cup G}(t)) \mid U_T^{T\cup G}(t) = U_T^{T\cup G}(\theta)\}$, is a non-degenerate interval. By Lemma 6, this image is further connected in $int(E^{T,G})$, and by Lemma 13 it contains interior points. Hence there are outcomes $a, a' \sim^T \theta$, which images, $(U_T^{T\cup G}(a), U_G^{T\cup G}(a))$ and $(U_T^{T\cup G}(a'), U_G^{T\cup G}(a'))$, are in $int(E^{T,G})$, and $U_G^{T\cup G}(a) > U_G^{T\cup G}(a')$. By the same arguments, the images of the indifference classes of a and a' under $\succsim^G$, $\{(U_T^{T\cup G}(t), U_G^{T\cup G}(t)) \mid U_G^{T\cup G}(t) = U_G^{T\cup G}(a)\}$ and $\{(U_T^{T\cup G}(t), U_G^{T\cup G}(t)) \mid U_G^{T\cup G}(t) = U_G^{T\cup G}(a')\}$, are connected in $int(E^{T,G})$, and there exist outcomes $b, b' \in X$, which images, $(U_T^{T\cup G}(b), U_G^{T\cup G}(b))$ and $(U_T^{T\cup G}(b'), U_G^{T\cup G}(b'))$, are also in $int(E^{T,G})$, such that $U_G^{T\cup G}(b) = U_G^{T\cup G}(a)$, $U_G^{T\cup G}(b') = U_G^{T\cup G}(a')$, and $U_T^{T\cup G}(b), U_T^{T\cup G}(b') > U_T^{T\cup G}(\theta)$.

Let t_1 denote an outcome for which $U_G^{T\cup G}(t_1) = U_G^{T\cup G}(a')$ and $U_T^{T\cup G}(t_1) = \min(U_T^{T\cup G}(b), U_T^{T\cup G}(b'))$. Such an outcome exists on account of connectedness of $\{(U_T^{T\cup G}(t), U_G^{T\cup G}(t)) \mid t \sim^G a'\}$. Note that $U_{T\cup G}(t_1) > U_{T\cup G}(a')$ since $U_T^{T\cup G}(t_1) > U_T^{T\cup G}(a')$. Set y_1 to be an outcome such that $U_G^{T\cup G}(y_1) = U_G^{T\cup G}(a')$ and $U_T^{T\cup G}(y_1) = \min(U_T^{T\cup G}(t_1), U_T^{T\cup G}(a) + \frac{1}{2}(U_G^{T\cup G}(a) - U_G^{T\cup G}(a')))$. Connectedness again implies that such an outcome y_1 exists. By the choice of utilities is satisfies,

$$\begin{aligned} U_{T\cup G}(a') &< U_{T\cup G}(y_1) = U_T^{T\cup G}(y_1) + U_G^{T\cup G}(y_1) \\ &< U_T^{T\cup G}(a) + U_G^{T\cup G}(a) - U_G^{T\cup G}(a') + U_G^{T\cup G}(a') = U_{T\cup G}(a) \ . \end{aligned}$$

Let r denote an outcome such that $U_T^{T\cup G}(r) = U_T^{T\cup G}(\theta)$ and $U_G^{T\cup G}(r) = U_{T\cup G}(y_1) - U_T^{T\cup G}(\theta)$, therefore $U_{T\cup G}(r) = U_{T\cup G}(y_1)$. Observe that since y_1 was chosen so that $U_{T\cup G}(a') < U_{T\cup G}(y_1) < U_{T\cup G}(a)$, then $U_G^{T\cup G}(a') < U_G^{T\cup G}(r) < U_G^{T\cup G}(a)$. Therefore such an outcome r indeed exists, following connectedness arguments as detailed above. As the utilities indifference curve of θ according to $\succsim^T$ is connected in $int(E^{T,G})$, and the images of a and a' are in that interior, the image of r must lie in $int(E^{T,G})$ as well. Therefore there exists an outcome y_2 with $U_{T\cup G}(y_1) - U_G^{T\cup G}(a) < U_T^{T\cup G}(y_2) < U_T^{T\cup G}(\theta)$ and $U_{T\cup G}(y_2) = U_{T\cup G}(y_1)$ (so that $U_G^{T\cup G}(a') < U_G^{T\cup G}(y_2) < U_G^{T\cup G}(a)$), such that there also exists y_* with $U_T^{T\cup G}(y_*) = U_T^{T\cup G}(y_2)$ and $U_G^{T\cup G}(y_*) = U_G^{T\cup G}(a')$. Set $y^* \in X$ to be an outcome such that $U_T^{T\cup G}(y^*) = U_T^{T\cup G}(y_1)$ and $U_G^{T\cup G}(y^*) = U_G^{T\cup G}(y_2)$. Such an outcome exists on account of the same connectedness arguments as above.

Altogether, following the definitions of y^* and y_*, together with connectedness of images of indifference classes within $E^{T,G}$ (Lemma 2), for every $(u_1, u_2) \in (U_T^{T\cup G}(y_*), U_T^{T\cup G}(y^*)) \times (U_G^{T\cup G}(y_*), U_G^{T\cup G}(y^*))$ there exists an outcome $t \in X$ such that $U_T^{T\cup G}(t) = u_1$ and $U_G^{T\cup G}(t) = u_2$. This proves part (1) of the lemma.

For part (2) of the lemma, let $\tilde{x}, \tilde{y}, \tilde{z}$, and $\tilde{w}$ be outcomes that satisfy $\tilde{x} \succ^T \tilde{y}$, $y^* \succ^T \tilde{x}, \tilde{y}, \tilde{z}, \tilde{w} \succ^T y_*$. Set x, y, z, w to be outcomes which obtain the following utilities (and which existence is guaranteed as explained in the previous paragraphs):

$$\begin{aligned}
U_T^{T\cup G}(x) = U_T^{T\cup G}(\tilde{x}) \quad &, \quad U_G^{T\cup G}(x) = U_G^{T\cup G}(y_*) \\
U_T^{T\cup G}(y) = U_T^{T\cup G}(\tilde{y}) \quad &, \quad U_G^{T\cup G}(y) = U_T^{T\cup G}(x) + U_G^{T\cup G}(x) - U_T^{T\cup G}(y) \\
U_T^{T\cup G}(z) = U_T^{T\cup G}(\tilde{z}) \quad &, \quad U_G^{T\cup G}(z) = U_G^{T\cup G}(x) \\
U_T^{T\cup G}(w) = U_T^{T\cup G}(\tilde{w}) \quad &, \quad U_G^{T\cup G}(w) = U_G^{T\cup G}(y)
\end{aligned}$$

The above choice of utilities, together with the assumed equality of utility differences, imply:

$$\begin{aligned}
& x \sim^T \tilde{x} \ , \ y \sim^T \tilde{y} \ , \ z \sim^T \tilde{z} \ , \ w \sim^T \tilde{w} \quad , \\
& x \sim^G z \ , \ y \sim^G w \quad , \\
& x \sim^{T\cup G} y \ .
\end{aligned}$$

Therefore, by definition, $x \ominus y \succeq^T z \ominus w$ if and only if $w \succsim^{T\cup G} z$, in other words, if and only if, $U_T^{T\cup G}(w) + U_T^{T\cup G}(x) + U_G^{T\cup G}(x) - U_T^{T\cup G}(y) \geq U_T^{T\cup G}(z) + U_G^{T\cup G}(x)$, or equivalently, $U_T^{T\cup G}(x) - U_T^{T\cup G}(y) \geq U_T^{T\cup G}(z) - U_T^{T\cup G}(w)$. Therefore, by the respective indifference of x, y, z, w to $\tilde{x}, \tilde{y}, \tilde{z}, \tilde{w}$ according to $\succsim^T$, $\tilde{x} \ominus \tilde{y} \succeq^T \tilde{z} \ominus \tilde{w}$, if and only if, $U_T^{T\cup G}(\tilde{x}) - U_T^{T\cup G}(\tilde{y}) \geq U_T^{T\cup G}(\tilde{z}) - U_T^{T\cup G}(\tilde{w})$.

Symmetric arguments prove part (3) of the lemma, by employing x which obtains $U_T^{G\cup T}(x) = U_T^{T\cup G}(y_*)$.

∎

Lemma 11. *Let $\theta \in X$ be such that $M \succ^T \theta \succ^T m$ for some $m, M \in X$. Then there are $x^*, x_* \in X$, $x^* \succ^T \theta \succ^T x_*$, such that for every $\tilde{x}, \tilde{y}, \tilde{z}, \tilde{w} \in X$, $\tilde{x} \succ^T \tilde{y}$, that satisfy*

$x^* \succ^T \tilde{x}, \tilde{y}, \tilde{z}, \tilde{w} \succ^T x_*$,

$$\begin{aligned} U_T^N(\tilde{x}) - U_T^N(\tilde{y}) &= U_T^N(\tilde{z}) - U_T^N(\tilde{w}) \quad \Longleftrightarrow \\ U_T^{T\cup G}(\tilde{x}) - U_T^{T\cup G}(\tilde{y}) &= U_T^{T\cup G}(\tilde{z}) - U_T^{T\cup G}(\tilde{w}) \ . \end{aligned}$$

Proof. According to Lemma 10 there are outcomes $y^* \succ^T y_*$, satisfying $y^* \succ^T \theta \succ^T y_*$, such that for every $(u_1, u_2) \in (U_T^N(y_*), U_T^N(y^*)) \times (U_{N\backslash T}^N(y_*), U_{N\backslash T}^N(y^*))$, there exists an outcome $t \in X$ such that $U_T^N(t) = u_1$ and $U_{N\backslash T}^N(t) = u_2$. In the same manner, there are outcomes $z^* \succ^T z_*$, $z^* \succ^T \theta \succ^T z_*$, such that for every $(u_1, u_2) \in (U_T^{T\cup G}(z_*), U_T^{T\cup G}(z^*)) \times (U_G^{T\cup G}(z_*), U_G^{T\cup G}(z^*))$, there exists an outcome $t' \in X$ such that $U_T^{T\cup G}(t') = u_1$ and $U_G^{T\cup G}(t') = u_2$. Set $x^* = y^*$ if $z^* \succsim^T y^*$ and $x^* = z^*$ otherwise, and set $x_* = y_*$ if $y_* \succsim^T z_*$ and $x_* = z_*$ otherwise. These outcomes also satisfy $x^* \succ^T \theta \succ^T x_*$.

Let $\tilde{x}, \tilde{y}, \tilde{z}$, and $\tilde{w}$ be outcomes that satisfy $\tilde{x} \succ^T \tilde{y}$, $x^* \succ^T \tilde{x}, \tilde{y}, \tilde{z}, \tilde{w} \succ^T x_*$, and $U_T^N(\tilde{x}) - U_T^N(\tilde{y}) = U_T^N(\tilde{z}) - U_T^N(\tilde{w})$. Set $x, y, z, w, x', y', z', w'$ to be outcomes which obtain the following utilities (and which existence is guaranteed as explained in the previous paragraphs):

$$\begin{aligned} U_T^N(x) = U_T^N(\tilde{x}) \quad &, \quad U_{N\backslash T}^N(x) = U_{N\backslash T}^N(y^*) \\ U_T^N(y) = U_T^N(\tilde{y}) \quad &, \quad U_{N\backslash T}^N(y) = U_T^N(\tilde{x}) + U_{N\backslash T}^N(y^*) - U_T^N(\tilde{y}) \\ U_T^N(z) = U_T^N(\tilde{z}) \quad &, \quad U_{N\backslash T}^N(z) = U_{N\backslash T}^N(y^*) \\ U_T^N(w) = U_T^N(\tilde{w}) \quad &, \quad U_{N\backslash T}^N(w) = U_T^N(\tilde{x}) + U_{N\backslash T}^N(y^*) - U_T^N(\tilde{y}) \end{aligned}$$

$$\begin{aligned} U_T^{T\cup G}(x') = U_T^{T\cup G}(\tilde{x}) \quad &, \quad U_G^{T\cup G}(x') = U_G^{T\cup G}(z^*) \\ U_T^{T\cup G}(y') = U_T^{T\cup G}(\tilde{y}) \quad &, \quad U_G^{T\cup G}(y') = U_T^{T\cup G}(\tilde{x}) + U_G^{T\cup G}(z^*) - U_T^{T\cup G}(\tilde{y}) \\ U_T^{T\cup G}(z') = U_T^{T\cup G}(\tilde{z}) \quad &, \quad U_G^{T\cup G}(z') = U_G^{T\cup G}(z^*) \\ U_T^{T\cup G}(w') = U_T^{T\cup G}(\tilde{w}) \quad &, \quad U_G^{T\cup G}(w') = U_T^{T\cup G}(\tilde{x}) + U_G^{T\cup G}(z^*) - U_T^{T\cup G}(\tilde{y}) \end{aligned}$$

The above choice of utilities, together with the assumed equality of utility differences, $U_T^N(\tilde{x}) - U_T^N(\tilde{y}) = U_T^N(\tilde{z}) - U_T^N(\tilde{w})$, imply:

$$\begin{aligned} & x \sim^T x' \ , \ y \sim^T y' \ , \ z \sim^T z' \ , \ w \sim^T w' \quad , \\ & x \sim^{N\backslash T} z \ , \ y \sim^{N\backslash T} w \quad , \\ & x' \sim^G z' \ , \ y' \sim^G w' \quad , \\ & x \sim^N y \ , \ z \sim^N w \ , \ x' \sim^{T\cup G} y' \ . \end{aligned}$$

Consistency of Social tradeoffs then entails that $z' \sim^{T\cup G} w'$, resulting in, $U_T^{T\cup G}(\tilde{z}) - U_T^{T\cup G}(\tilde{w}) = U_T^{T\cup G}(\tilde{x}) - U_T^{T\cup G}(\tilde{y})$, as required. Assuming that conclusion will in turn, using the same arguments, yield the same utility differences for U_T^N.

■

Since both functions, $U_T^{T\cup G}$ and U_T^N, represent $\succsim^T$, then local equality of utility differences implies that for every $\theta \in X$ such that $M \succ^T \theta \succ^T m$ for some $m, M \in X$, there are $x^* \succ^T \theta \succ^T x_*$ and $\beta, \tau \in \mathbb{R}, \beta > 0$, such that on $\{t \in X \mid x^* \succ^T t \succ^T x_*\}$, $U_T^{T\cup G} = \beta U_T^N + \tau$. Applying that conclusion to outcomes θ with overlapping preference-intervals yields that the same is true globally.

5.3 Proof of Proposition 3

Lemma 12. *Let $\theta \in X$ be such that $M \succ^T \theta \succ^T m$ for some $m, M \in X$. Then there are $y^*, y_* \in X$, $y^* \succ^T \theta \succ^T y_*$, such that for every $\tilde{x}, \tilde{y}, \tilde{z}, \tilde{w} \in X$ that satisfy $y^* \succ^T \tilde{x} \succ^T \tilde{y} \succ^T y_*$ and $y^* \succ^T \tilde{z} \succ^T \tilde{w} \succ^T y_*$, there are $x', y', z', w', y'', w'' \in X$, such that,*

$$\begin{aligned}
&x' \sim^T \tilde{x} \ , \ y' \sim^T \tilde{y} \ , \ z' \sim^T \tilde{z} \ , \ w' \sim^T \tilde{w} \ , \\
&x' \sim^H y' \sim^H z' \sim^H w' \ , \\
&y'' \sim^{T\setminus H} y' \ , \ w'' \sim^{T\setminus H} w' \ , \\
&y'' \sim^T \tilde{x} \ .
\end{aligned}$$

Proof. Consider the mapping $(U_{T\setminus H}^T, U_H^T) : X \to \mathbb{R}^2$, and denote its range by $E^{T\setminus H,H} = \{(U_{T\setminus H}^T(x), U_H^T(x)) \mid x \in X\}$. According to the proof of Lemma 5, every point in the relative interior of the utility indifference curve $\{(U_{T\setminus H}^T(x), U_T^T(x)) \mid x \in X,\ U_{T\setminus H}^T(x) + U_T^T(x) = U_{T\setminus H}^T(\theta) + U_T^T(\theta)\}$ is also in $int(E^{T\setminus H,H})$. Therefore, if there is more than one point in this set, there exists a point $\hat{\theta} \in int(E^{T\setminus H,H})$, with $\hat{\theta} \sim^T \theta$.

Suppose on the contrary that $\{(U_{T\setminus H}^T(x), U_H^T(x)) \mid x \in X,\ U_{T\setminus H}^T(x) + U_H^T(x) = U_{T\setminus H}^T(\theta) + U_H^T(\theta)\}$ is a singleton. Hence $(U_{T\setminus H}^T(\theta), U_H^T(\theta)) \notin int(E^{T\setminus H,H})$, and $int(E^{T\setminus H,H})$ contains no points (a,t) such that $a + t = U_{T\setminus H}^T(\theta) + U_H^T(\theta)$. By assumption, θ is not unanimously minimal nor unanimously maximal. Hence Agreed Improvement (**A3**) implies that there is $(a,t) \in E^{T\setminus H,H}$ such that $a > U_{T\setminus H}^T(\theta)$ and $t > U_H^T(\theta)$. According to the proof of Lemma 3, every neighborhood of (a,t) intersects $int(E^{T\setminus H,H})$, therefore there is $(a',t') \in int(E^{T\setminus H,H})$ such that $a' + t' > U_{T\setminus H}^T(\theta) + U_H^T(\theta)$. Analogously there is $(b',s') \in int(E^{T\setminus H,H})$ such that $b' + s' < U_{T\setminus H}^T(\theta) + U_H^T(\theta)$.

Consider the sets $\{(a,t) \in int(E^{T\setminus H,H}) \mid a+t > U_{T\setminus H}^T(\theta)+U_H^T(\theta)\}$ and $\{(b,s) \in int(E^{T\setminus H,H}) \mid b+s < U_{T\setminus H}^T(\theta) + U_H^T(\theta)\}$. According to the previous paragraph these sets are non-empty, and following the contrary assumption they form a partition of $int(E^{T\setminus H,H})$. However, this constitutes a contradiction to $int(E^{T\setminus H,H})$ being connected (Lemma 4). Therefore $\{(U_{T\setminus H}^T(x), U_H^T(x)) \mid x \in X,\ U_{T\setminus H}^T(x) + U_H^T(x) = U_{T\setminus H}^T(\theta) + U_H^T(\theta)\}$ cannot be a singleton. Since it is connected (as the image of a connected set under a continuous transformation), its relative interior is non-empty, hence there is $\hat{\theta} \sim^T \theta$ such that $\hat{\theta} \in int(E^{T\setminus H,H})$.

Since $\hat{\theta}$ is an interior point of $E^{T\setminus H,H}$, there are $x^1, y^1 \in X$ such that $x^1, y^1 \sim^H \hat{\theta}$, $x^1 \succ^T \hat{\theta} \succ^T y^1$. If $(U_{T\setminus H}^T(y^1), U_T(x^1) - U_{T\setminus H}^T(y^1)) \in E^{T\setminus H,H}$, set $y^* = x^1, y_* = y^1$. Otherwise let

x^2, y^2 be outcomes such that $x^2, y^2 \sim^H \hat{\theta}$,
$U^T_{T\setminus H}(x^2) = \frac{1}{2}\left(U^T_{T\setminus H}(\hat{\theta}) + U^T_{T\setminus H}(x^1)\right)$, $U^T_{T\setminus H}(y^2) = \frac{1}{2}\left(U^T_{T\setminus H}(\hat{\theta}) + U^T_{T\setminus H}(y^1)\right)$. Such outcomes exist on account of Lemma 2.

If $(U^T_{T\setminus H}(y^2), U_T(x^2) - U^T_{T\setminus H}(y^2)) \in E^{T\setminus H,H}$, set $y^* = x^2, y_* = y^2$, otherwise set x^3, y^3 in the same manner, and so on. As there is a neighborhood of $\hat{\theta}$ contained in $E^{T\setminus H,H}$, and $(U^T_{T\setminus H}(\hat{\theta}), U_T(\hat{\theta}) - U^T_{T\setminus H}(\hat{\theta}))$ is in $E^{T\setminus H,H}$, the process will stop after a finite number of steps. Namely, there are $x^k, y^k \sim^H \hat{\theta}$, $x^k \succ^T \hat{\theta} \succ^T y^k$, hence also $x^k \succ^{T\setminus H} \hat{\theta} \succ^{T\setminus H} y^k$, such that $(U^T_{T\setminus H}(y^k), U_T(x^k) - U^T_{T\setminus H}(y^k)) \in E^{T\setminus H,H}$. Set $y^* = x^k, y_* = y^k$.

By connectedness of all utility indifference curves, and Lemma 2, for every (a,t) such that $t \geq U^T_H(\hat{\theta})$, $U^T_{T\setminus H}(y^*) \geq a \geq U^T_{T\setminus H}(y_*)$, and $a + t \leq U^T_{T\setminus H}(y^*) + U^T_H(y^*)$ there is $z \in X$ such that $U^T_{T\setminus H}(z) = a$ and $U^T_H(z) = t$. For any outcomes $\tilde{x} \succ^T \tilde{y}$ and $\tilde{z} \succ^T \tilde{w}$, such that $y^* \succ^T \tilde{x}, \tilde{y}, \tilde{z}, \tilde{w} \succ^T y_*$, let x', y', z', w', y'', w'' be outcomes such that,

$$\begin{aligned}
U^T_H(x') = U^T_T(\hat{\theta}) \quad &, \quad U^T_{T\setminus H}(x') = U_T(\tilde{x}) - U^T_H(\hat{\theta}) \ , \\
U^T_H(y') = U^T_T(\hat{\theta}) \quad &, \quad U^T_{T\setminus H}(y') = U_T(\tilde{y}) - U^T_H(\hat{\theta}) \ , \\
U^T_H(y'') = U_T(\tilde{x}) - U^T_{T\setminus H}(y') \quad &, \quad U^T_{T\setminus H}(y'') = U^T_{T\setminus H}(y') \ , \\
\\
U^T_H(z') = U^T_H(\hat{\theta}) \quad &, \quad U^T_{T\setminus H}(z') = U_T(\tilde{z}) - U^T_H(\hat{\theta}) \ , \\
U^T_H(w') = U^T_H(\hat{\theta}) \quad &, \quad U^T_{T\setminus H}(w') = U_T(\tilde{w}) - U^T_H(\hat{\theta}) \ , \\
U^T_H(w'') = U^T_H(y'') \quad &, \quad U^T_{T\setminus H}(w'') = U^T_{T\setminus H}(w') \ .
\end{aligned}$$

According to the above, these outcomes exist. By choice of the utilities they satisfy,

$$\begin{aligned}
&x' \sim^T \tilde{x} \ , \ y' \sim^T \tilde{y} \ , \ z' \sim^T \tilde{z} \ , \ w' \sim^T \tilde{w} \ , \\
&x' \sim^H y' \sim^H z' \sim^H w' \ , \\
&y'' \sim^{T\setminus H} y' \ , \ w'' \sim^{T\setminus H} w' \ , \\
&y'' \sim^T \tilde{x} \ .
\end{aligned}$$

■

Employing the above lemma we turn to prove that U^N_T and U_T that is obtained as the additive representation of $\succsim^H$ and $\succsim^{T\setminus H}$ maintain local equality of utility differences.

Lemma 13. *Let $\theta \in X$ be such that $M \succ^T \theta \succ^T m$ for some $m, M \in X$. Then there are $x^*, x_* \in X$, $x^* \succ^T \theta \succ^T x_*$, such that for every $\tilde{x}, \tilde{y}, \tilde{z}, \tilde{w} \in X$, $\tilde{x} \succ^T \tilde{y}$, that satisfy $x^* \succ^T \tilde{x}, \tilde{y}, \tilde{z}, \tilde{w} \succ^T x_*$,*

$$\begin{aligned}
U^N_T(\tilde{x}) - U^N_T(\tilde{y}) \quad &= \quad U^N_T(\tilde{z}) - U^N_T(\tilde{w}) \quad \Longleftrightarrow \\
U_T(\tilde{x}) - U_T(\tilde{y}) \quad &= \quad U_T(\tilde{z}) - U_T(\tilde{w}) \ .
\end{aligned}$$

Proof. The previous lemma establishes that there are outcomes y^*, y_*, $y^* \succ^T \hat{\theta} \succ^T y_*$, such that for any outcomes $\tilde{x} \succ^T \tilde{y}$ and $\tilde{z} \succ^T \tilde{w}$, that satisfy $y^* \succ^T \tilde{x}, \tilde{y}, \tilde{z}, \tilde{w} \succ^T y_*$, there are outcomes x', y', z', w', y'', w'' that satisfy,

$$
\begin{aligned}
&x' \sim^T \tilde{x} \ , \ y' \sim^T \tilde{y} \ , \ z' \sim^T \tilde{z} \ , \ w' \sim^T \tilde{w} \ , \\
&x' \sim^H y' \sim^H z' \sim^H w' \ , \\
&y'' \sim^{T \setminus H} y' \ , \ w'' \sim^{T \setminus H} w' \ , \\
&y'' \sim^T \tilde{x} \ , \ w'' \sim^T \tilde{z} \ .
\end{aligned}
$$

Moreover, in the same manner as in the proof of Lemma 11, there are $z^*, z_* \in X$, $z^* \succ^T \theta \succ^T z_*$, such that whenever outcomes $\tilde{x} \succ^T \tilde{y}$ and $\tilde{z} \succ^T \tilde{w}$ satisfy $y^* \succ^T \tilde{x}, \tilde{y}, \tilde{z}, \tilde{w} \succ^T y_*$, there are outcomes x''', y''', z''', w''' for which,

$$
\begin{aligned}
&x''' \sim^T \tilde{x} \ , \ y''' \sim^T \tilde{y} \ , \ z''' \sim^T \tilde{z} \ , \ w''' \sim^T \tilde{w} \ , \\
&x''' \sim^{N \setminus T} z''' \ , \ y''' \sim^{N \setminus T} w''' \ , \\
&x''' \sim^N y''' \ .
\end{aligned}
$$

If $z^* \succsim^T y^*$ set $x^* = y^*$, otherwise set $x^* = z^*$. Similarly, if $y_* \succsim^T z_*$ set $x_* = y_*$, otherwise set $x_* = z_*$. Let $\tilde{x}, \tilde{y}, \tilde{z}, \tilde{w}$ be outcomes that satisfy $x^* \succ^T \tilde{x}, \tilde{y}, \tilde{z}, \tilde{w} \succ^T x_*$ and $\tilde{x} \succ^T \tilde{y}$. Then there are outcomes x', y', z', w', outcomes x'', y'', z'', w'', and outcomes x''', y''', z''', w''', which satisfy,

$$
\begin{aligned}
&x' \sim^T \tilde{x} \ , \ y' \sim^T \tilde{y} \ , \ z' \sim^T \tilde{z} \ , \ w' \sim^T \tilde{w} \ , \\
&x' \sim^H y' \sim^H z' \sim^H w' \ ,
\end{aligned} \tag{2}
$$

$$
\begin{aligned}
&x'' = x' \ , \ y'' \sim^{T \setminus H} y' \ , \ z'' = z' \ , \ w'' \sim^{T \setminus H} w' \ , \\
&x'' \sim^H z'' \ , \ y'' \sim^H w'' \ , \\
&x'' \sim^T y'' \ ,
\end{aligned} \tag{3}
$$

$$
\begin{aligned}
&x''' \sim^T \tilde{x} \ , \ y''' \sim^T \tilde{y} \ , \ z''' \sim^T \tilde{z} \ , \ w''' \sim^T \tilde{w} \ , \\
&x''' \sim^{N \setminus T} z''' \ , \ y''' \sim^{N \setminus T} w''' \ , \\
&x''' \sim^N y''' \ .
\end{aligned} \tag{4}
$$

First, suppose $U_T(\tilde{x}) - U_T(\tilde{y}) = U_T(\tilde{z}) - U_T(\tilde{w})$. The same equality then holds for x', y', z', w' and x''', y''', z''', w'''. Under that equality of utility differences, indifference relationships (2) and (3) imply $z'' \sim^T w''$. Hence, by definition, $x' \ominus y' \simeq^{T \setminus H} z' \ominus w'$.

On the other hand, indifference relationships (4) imply that the tradeoff between x''' and y''' is $\succeq^T$-comparable to the tradeoff between z''' and w''': $x''' \ominus y''' \succeq^T z''' \ominus w''' \iff w''' \succsim^N z'''$.

As $x' \sim^T x''', y' \sim^T y''', z' \sim^T z'''$, and $w' \sim^T w'''$, and according to the definition of tradeoffs comparison, this is equivalent to, $x' \ominus y' \succeq^T z' \ominus w' \iff w''' \succsim^N z'''$. However, since $x' \ominus y' \simeq^H z' \ominus w'$ (as they are all $\succsim^H$ indifferent, and $x' \ominus x' \simeq^H x' \ominus x'$) and $x' \ominus y' \simeq^{T \setminus H} z' \ominus w'$, Tradeoff Pareto (**A6**) implies that necessarily also $x' \ominus y' \simeq^T z' \ominus w'$, and therefore it must be that $w''' \sim^N z'''$. Translating that indifference to the additive representation of $\succsim^N$ using T and $N \setminus T$, and employing the indifference relationships $x''' \sim^{N \setminus T} z'''$, $y''' \sim^{N \setminus T} w'''$, and $x''' \sim^N y'''$,

$$\begin{aligned}
U_N(z''') &= U_N(w''') \iff \\
U_T^N(z''') + U_{N \setminus T}^N(z''') &= U_T^N(w''') + U_{N \setminus T}^N(w''') \iff \\
U_T^N(z''') - U_T^N(w''') &= U_{N \setminus T}^N(w''') - U_{N \setminus T}^N(z''') \iff \\
U_T^N(z''') - U_T^N(w''') &= U_{N \setminus T}^N(y''') - U_{N \setminus T}^N(x''') \iff \\
U_T^N(z''') - U_T^N(w''') &= U_T^N(x''') - U_T^N(y''') \ .
\end{aligned}$$

Since $x''' \sim^T \tilde{x}$, $y''' \sim^T \tilde{y}$, $z''' \sim^T \tilde{z}$, and $w''' \sim^T \tilde{w}$, it follows that, $U_T^N(\tilde{z}) - U_T^N(\tilde{w}) = U_T^N(\tilde{x}) - U_T^N(\tilde{y})$, as required.

In the other direction, suppose $U_T^N(\tilde{x}) - U_T^N(\tilde{y}) = U_T^N(\tilde{z}) - U_T^N(\tilde{w})$. Employing indifference relationships (4) it follows that $z''' \sim^N w'''$, and therefore $x''' \ominus y''' \simeq^T z''' \ominus w'''$, which is equivalent to $x' \ominus y' \simeq^T z' \ominus w'$, as $x' \sim^T x''', y' \sim^T y''', z' \sim^T z'''$, and $w' \sim^T w'''$. In addition, (3) implies that $x'' \ominus y'' \succeq^{T \setminus H} z'' \ominus w'' \iff w'' \succsim^T z''$, or, equivalently, $x' \ominus y' \succeq^{T \setminus H} z' \ominus w' \iff w'' \succsim^T z''$. Similarly to the above, Tradeoff Pareto (**A6**) implies that $x' \ominus y' \simeq^{T \setminus H} z' \ominus w'$, hence it must be that $w'' \sim^T z''$. Substituting using the additive representation $U_T = U_H^T + U_{T \setminus H}^T$ and the relationships in (2) and (3), it follows that $U_T(x') - U_T(y') = U_T(z') - U_T(w')$, and the same then holds for $\tilde{x}, \tilde{y}, \tilde{z}$ and $\tilde{w}$. ■

In the same manner as in the proof of Proposition 2, as both functions, U_T and U_T^N, represent $\succsim^T$, then local equality of utility differences implies that for every $\theta \in X$ such that $M \succ^T \theta \succ^T m$ for some $m, M \in X$, there are $x^* \succ^T \theta \succ^T x_*$ and $\gamma, \xi \in \mathbb{R}$, $\gamma > 0$, such that on $\{t \in X \mid x^* \succ^T t \succ^T x_*\}$, $U_T^N = \gamma U_T + \xi$. Applying that conclusion to outcomes θ with overlapping preference-intervals yields that the conclusion holds globally.