Theory Dec.
DOI 10.1007/s11238-016-9537-0

On the existence of altruistic value and utility functions

Jay Simon[1]

Abstract Altruism is a popular economic explanation for a wide range of pro-social decisions and actions. It has been applied frequently in several different streams of literature, and is a descriptively compelling model of behavior. This paper provides a theoretical framework for the existence of ordinal and cardinal altruistic value functions, as well as altruistic utility functions, based on an altruistic preference relation. Representation theorems are developed to specify relatively weak conditions under which altruistic value and utility functions can be shown to exist. In addition, conditions that lead to additive forms of these functions are provided.

Keywords Preferences · Altruism · Value · Utility · Representation theorems

1 Introduction

It is widely recognized that humans often display altruistic behavior. That is, we make decisions that are inconsistent with some narrowly defined sense of self-interest. These decisions may involve charitable giving, volunteering for a non-profit organization, helping a friend move, fixing a flat tire for a stranger, or any of a number of other generous actions. All of these activities reflect an aspect of preferences that cannot be captured adequately by the traditional economic idea of value or utility achieved through consumption of goods and services.

However, the prevalence of altruistic behavior has been acknowledged widely in both economics and psychology literature, and a great deal of discussion exists exploring the reasons behind it. Arrow (1975) presents various types of preferences under

✉ Jay Simon
jaysimon@american.edu

1 Kogod School of Business, American University, Washington, DC 20016, USA

Published online: 05 February 2016

which it may be desirable for an individual to donate blood. Andreoni (1990) uses the term "impurely altruistic" to express the idea that individuals receive some benefit not only from improvement in a public good, but also from the act of giving itself. McCardle et al. (2009) use a model of preference that incorporates the recognition or acclaim received by individuals for charitable acts. There are also many evolutionary models that justify widespread altruistic behavior in humans (Hamilton 1963; Trivers 1971; Alexander 1987; Simon 1990; Bergstrom and Stark 1993; Nowak and Sigmund 2005; Montero 2008). Intergenerational altruistic models are also common; see, for instance, Phelps and Pollak (1968), Ray (1987), Hori and Kanaya (1989), and Saez-Marti and Weibull (2005). Alger and Weibull (2010) examine altruistic models between siblings rather than across generations.

Given that decision makers are likely to be altruistic, developing value and utility functions that capture altruistic preferences is desirable to help these decisions makers achieve more desirable outcomes. However, it is not clear that these value and utility functions are guaranteed to exist, as the conditions typically used to ensure existence of value and utility functions are substantially stronger when multiple individuals' preferences are being aggregated. In addition, even if such functions exist, there are differing views in the previously cited literature on the form that they should take. For example, paternalistic altruistic value or utility functions take another individual's outcome as an argument, while nonpaternalistic altruistic value or utility functions take the value or utility achieved by another individual as an argument. In this paper, we provide representation theorems for paternalistic altruistic functions, and then expand them to additive nonpaternalistic functions by introducing one further condition.

There is some disagreement in the literature regarding the names used to describe preference functions, which we will not attempt to resolve here. In this paper, a "value function" represents preferences over a set of outcomes in an environment of certainty, and may be either ordinal or cardinal, while a "utility function" represents preferences for risky decisions, i.e., preferences over gambles defined on a set of outcomes.

The goal of this paper is to define altruism precisely using preference relations, and to establish the existence of an altruistic value or utility function representing these preference relations based on relatively simple conditions. In particular, we do not assume that the altruistic preference relations are complete, transitive, or continuous; the existence of altruistic value and utility functions can be established using weaker preference conditions. The value functions may be either ordinal or cardinal, i.e., they may only preserve the ordering of elements, or they may preserve relative differences between pairs of elements. This paper also explores some of the relevant properties of such preference relations and value functions; of particular interest are the conditions required for the functions to have additive forms.

The benefit of weakening the preference conditions is greater in a context of multiple individuals, because completeness, transitivity, and continuity become substantially stronger conditions when the preference relation aggregates the preferences of different people. A weaker set of preference conditions will make the existence of an altruistic value or utility function more plausible. The results need not apply only in the specific altruistic framework described in the paper. If the relevant properties and conditions are satisfied in other contexts involving multiple individuals (e.g., a parent considering outcomes affecting his/her multiple children), then the analogous representation theorems will apply, and a value or utility function will exist. The structure of the altruistic value and utility functions is similar to that of multiattribute value and utility functions for individual decision makers. Indeed, if the analogous properties of the preference relations are satisfied in that context, weaker conditions than those typically used can guarantee the existence of a multiattribute value or utility function. However, it is generally considered a reasonable assumption that rational individuals will have complete, transitive, and continuous preferences.

There is some debate in the psychology literature regarding the underlying factors driving altruistic behavior. Batson et al. (1989), for instance, claim that anticipated improvement in the decision maker's mood is not necessary for helping others under conditions of empathy. Cialdini et al. (1987), on the other hand, argue that an egotistical explanation is more appropriate. This issue is discussed in greater depth by Reyniers and Bhalla (2013). This debate serves as an interesting background for the material presented in this paper, but the concepts described here are applicable regardless of whether the factors underlying altruistic behavior are selfish or selfless.

The remainder of the paper is structured as follows. Section 2 introduces the concept of altruistic preference relations. Section 3 explores the representation of altruistic preferences with ordinal value functions. Section 4 adds preference difference relations, which can allow for the use of cardinal altruistic value functions. Section 5 considers preference relations over gambles, and presents conditions leading to the existence of altruistic utility functions. Section 6 concludes the paper.

2 Altruistic preference relations

Consider a set of outcomes X consisting of pairs of individual outcomes affecting two people designated as A and B. These two individuals have preference relations $\succsim_A$ and $\succsim_B$, respectively, over X, with corresponding indifference relations $\sim_A$ and $\sim_B$, and corresponding strict relations $\succ_A$ and $\succ_B$, defined in the standard manner. These are preference relations in the traditional sense; they do not include any interaction between the preferences of individuals A and B. They consider only the outcome experienced by that individual. They are in fact somewhat more narrowly defined than traditional preference relations, in that they are meaningful and valid in isolation of the other individual. Traditional preference relations typically do not include restrictions on what considerations can be incorporated.

However, it is not unreasonable to imagine that one individual might care about the other's preferences. Consider another possible preference relation $\succsim_{A'}$ of individual A over the set of outcomes (with corresponding $\succ_{A'}$ and $\sim_{A'}$). This new preference relation is intended to capture not only individual A's intrinsic satisfaction with the given outcomes, as reflected by $\succsim_A$, but also A's understanding of $\succsim_B$. That is, $\succsim_{A'}$ incorporates both $\succsim_A$ and $\succsim_B$.

This additional preference relation can be informative in situations in which each individual is directly affected only by a portion of each overall outcome. For example, consider $X = \{(x^A, x^B) : x^A \in [x_l^A, x_h^A], x^B \in [x_l^B, x_h^B]\}$, where x^A and x^B reflect amounts of money received by each individual. In this case, we can think of $\succsim_A$ as A's preferences over the possible levels of (x^A, x^B) when ignoring x^B, and $\succsim_{A'}$ as

A's preferences over the possible levels of $\left(x^A, x^B\right)$ when considering both x^A and x^B, where $\succsim_{A'}$ might be influenced by both $\succsim_A$ and $\succsim_B$. Of course, there is no a priori requirement that $\succsim_{A'}$ must be different from $\succsim_A$; it is possible that A does not care about B's preferences.

Let $X = X^A \times X^B$, where $X^A = I^A \subset \mathbb{R}$ and $X^B = I^B \subset \mathbb{R}$ are closed intervals. Let x^A and x^B represent arbitrary numbers in X^A and X^B, respectively. We will write x as shorthand for (x^A, x^B) to denote a generic outcome in X, and use a subscript, e.g., x_1 for (x_1^A, x_1^B), to denote an arbitrary, but specific outcome in X. Since $\succsim_A$ is unaffected by x^B, for convenience, we will sometimes use $\succsim_A$ to compare elements of X^A rather than elements of X. Similarly, we will sometimes use $\succsim_B$ to compare elements of X^B.

Definition 1 $\succsim_{A'}$ satisfies the Pareto property if for any $(x_1^A, x_1^B), (x_2^A, x_2^B) \in X$, $x_1^A \succsim_A x_2^A$ and $x_1^B \succsim_B x_2^B$ implies $x_1 \succsim_{A'} x_2$, and if the strict relation holds for at least one of the first two comparisons, then $(x_1^A, x_1^B) \succ_{A'} (x_2^A, x_2^B)$.

The Pareto property ensures consistency between an altruistic preference relation and the underlying intrinsic preference relations. To avoid repetition, we will assume for the remainder of the paper that all preference relations and value functions presented reflect the preferences of individual A, individual B, or both (but not the preferences of anyone else).

The Pareto property is an efficiency property, meaning it ensures that if one outcome dominates another, the dominant outcome will always be preferred. However, it has a somewhat different interpretation in this context, because $\succsim_{A'}$ will generally reflect the same preferences over X^A as those captured by $\succsim_A$. It is the treatment of $\succsim_B$ by $\succsim_{A'}$ that is typically of greater interest. Specifically, the Pareto property ensures that, all else being equal, A prefers that B experience more desirable outcomes.

Definition 2 $\succsim_{A'}$ is an altruistic preference relation if $\succsim_{A'}$ satisfies the Pareto property.

It is straightforward to define an altruistic preference relation $\succsim_{B'}$ for individual B in the same manner. Note that the definition of an altruistic preference relation does not attempt to capture a degree of altruism; it simply provides a means to specify that a particular set of preferences is either altruistic or non-altruistic. Note also that this type of preference relation differs substantially from one that incorporates fairness or equity. Altruistic preferences do not necessarily favor more equitable outcomes, and preferences incorporating fairness or equity do not necessarily satisfy the Pareto property.

It is also possible to define a spiteful preference relation similarly, but with the preference comparisons of the other individual reversed. In fact, all of the results for altruistic value and utility functions in this paper could be obtained for spiteful value and utility functions as well; they depend only on the combined preference relation being monotonic in the intrinsic preference relations. A similar idea is captured by Vostroknutov (2013) in the context of social preferences that incorporate status within a group; the resulting utility functions are valid regardless of whether the decision maker would prefer others to be better or worse off. However, we will make the bold assumption that the reader is more interested in improving the quality of altruistic decisions than that of spiteful decisions, and will present the results in the altruistic context only.

3 Ordinal altruistic value functions

If $\succsim_A$ and $\succsim_B$ satisfy certain properties, then $\succsim_A$ and $\succsim_B$ can be represented by ordinal value functions $v^A\left(x^A\right)$ and $v^B\left(x^B\right)$ such that $\forall x_1^A, x_2^A \in X^A$, $x_1^A \succsim_A x_2^A$ if and only if $v^A\left(x_1^A\right) \geq v^A\left(x_2^A\right)$, and similarly for individual B. These value functions preserve only the ordering of outcomes imposed by $\succsim_A$ and $\succsim_B$; they are unique up to positive monotonic transformations. Various sets of properties which imply the existence of such ordinal value functions are provided by Debreu (1954, 1964), Rader (1963), Fishburn (1970), and Krantz et al. (1971). The properties used in this paper are defined here:

Definition 3 $\succsim_A$ is complete if for any $x_1^A, x_2^A \in X^A$, either $x_1^A \succsim_A x_2^A$, $x_2^A \succsim_A x_1^A$, or both.

Definition 4 $\succsim_A$ is transitive if for any $x_1^A, x_2^A, x_3^A \in X^A$, $x_1^A \succsim_A x_2^A$ and $x_2^A \succsim_A x_3^A$ implies $x_1^A \succsim_A x_3^A$.

Definition 5 $\succsim_A$ is continuous if for any $x_1^A, x_2^A \in X^A$ for which $x_1^A \succ_A x_2^A$, $\exists \Delta > 0$ such that for any $x_3^A \in X^A$, $|x_2^A - x_3^A| < \Delta$ implies $x_1^A \succ_A x_3^A$, and $|x_1^A - x_3^A| < \Delta$ implies $x_3^A \succ_A x_2^A$.

Completeness and transitivity ensure that the preference relation induces an ordering on the set of outcomes. Continuity ensures that there are no "jumps" in preferences caused by infinitesimally small changes. This definition is based on the one used by Simon et al. (2014), and is chosen here for simplicity; other definitions are more appropriate for more general topological spaces. We will also assume throughout the paper that both $\succsim_A$ and $\succsim_B$ are non-trivial; that is, there exist $x_1^A, x_2^A \in X^A$ such that $x_1^A \succ_A x_2^A$, and there exist $x_1^B, x_2^B \in X^B$ such that $x_1^B \succ_B x_2^B$.

Let $\succsim_A$ and $\succsim_B$ be complete, transitive, and continuous preference relations. Then, based on Debreu (1954, 1964), corresponding continuous ordinal value functions $v^A\left(x^A\right)$ and $v^B\left(x^B\right)$ exist. We assume for simplicity that the two individuals' preferences are monotonic over X^A and X^B, respectively ($x_1^A \succsim_A x_2^A$ iff $x_1^A \geq x_2^A$, and similarly for $\succsim_B$). This assumption is not necessary to establish any of the value functions presented in the paper, as the underlying variables can always be transformed such that monotonicity holds, but it will greatly reduce the amount of notation required.

Analogously, it is also useful to be able to represent the preference relations $\succsim_{A'}$ and $\succsim_{B'}$ by ordinal value functions $v^{A'}(x)$ and $v^{B'}(x)$ such that $x_1 \succsim_{A'} x_2$ if and only if $v^{A'}(x_1) \geq v^{A'}(x_2)$, and similarly for individual B. (Bell and Keeney 2009 make the same distinction between "egotistical" value and altruistic value as is made in this paper, though we use the term "intrinsic" rather than egotistical). Unfortunately, the existence of $v^A\left(x^A\right)$ and $v^B\left(x^B\right)$ does not guarantee the existence of such $v^{A'}(x)$ and $v^{B'}(x)$. In particular, it is not assured that $\succsim_{A'}$ and $\succsim_{B'}$ are complete and transitive, thus we cannot assume they induce orderings on X. Given the challenges associated with aggregating ordinal rankings across multiple individuals (Condorcet 1785; Arrow 1951; Sen 1970), an assumption that $\succsim_{A'}$ and $\succsim_{B'}$ induce orderings on X would be rather strong. Typically, when a group ordering of outcomes is desired, other conditions are added. For example, Shubik (1982) shows that complete continuous group preferences satisfying the Pareto property must be transitive, establishing the existence of a group value function. In this paper, we assume neither completeness nor transitivity of the altruistic preference relations. We instead establish the existence of altruistic value functions by introducing two weaker conditions on $\succsim_{A'}$ and $\succsim_{B'}$.

Definition 6 $\succsim_{A'}$ satisfies the substitution property if for any $(x_1^A, x_1^B), (x_2^A, x_2^B) \in X$ for which $(x_1^A, x_1^B) \sim_{A'} (x_2^A, x_2^B)$, it holds that for any $(x_3^A, x_3^B) \in X$, $(x_1^A, x_1^B) \succsim_{A'} (x_3^A, x_3^B)$ iff $(x_2^A, x_2^B) \succsim_{A'} (x_3^A, x_3^B)$, and $(x_3^A, x_3^B) \succsim_{A'} (x_1^A, x_1^B)$ iff $(x_3^A, x_3^B) \succsim_{A'} (x_2^A, x_2^B)$.

The substitution property asserts that equally preferable outcomes can be substituted for one another without affecting the truth of a comparison. Note that substitution is a weaker condition than transitivity. It gives meaning to the notion of indifference classes, but does not preclude intransitive relationships between indifference classes.

Definition 7 $\succsim_{A'}$ satisfies the indifference property if for any $x_1, x_2, x_3 \in X$ such that $x_1 \succsim_{A'} x_2$ and $x_2 \succsim_{A'} x_3$, there exists a real number $k \in [0, 1]$ such that $(x_1^A + km^A, x_1^B + km^B) \sim_{A'} x_2$, where $m^A = x_3^A - x_1^A$ and $m^B = x_3^B - x_1^B$.

The indifference property is a specific type of solvability. It simply asserts that if outcome x_1 is preferred to x_2, and x_2 is preferred to x_3, there must exist a convex combination of x_1 and x_3 (i.e., an outcome on the line segment between them) that is indifferent to x_2. Note that this property does not follow from continuity on $\succsim_A$ and $\succsim_B$. For example, consider lexicographic preferences on (x^A, x^B) consistent with $\succsim_A$ and $\succsim_B$: $x_1 \succsim_{A'} x_2$ iff $x_1 \succ_A x_2$, or $x_1 \sim_A x_2$ and $x_1 \succsim_B x_2$. The isoquant on which x_2 falls consists of the set of points $\{x_i : x_i \sim_A x_2 \text{ and } x_i \sim_B x_2\}$ (which may be only x_2 itself). That set does not, in general, include a convex combination of x_1 and x_3.

Theorem 1 *The following two statements are equivalent for given $\succsim_A$, $\succsim_B$, X, and a preference relation $\succsim_{A'}$ defined based on $\succsim_A$, $\succsim_B$, and X:*

1. *$\succsim_{A'}$ is an altruistic preference relation satisfying the substitution and indifference properties.*
2. *There exists a continuous ordinal $v^{A'}(x^A, x^B)$ representing $\succsim_{A'}$ that is monotonically increasing in x^A and x^B.*

See the Appendix for the proof of Theorem 1. As in the case of traditional ordinal value functions, this ordinal altruistic value function is unique up to continuous positive monotonic transformations. As stated earlier, the substitution and indifference properties are weaker than the set of conditions typically used to show existence of ordinal value functions. Theorem 1 is possible because $\succsim_{A'}$ satisfies the Pareto property, which establishes a basic set of pairs of elements of X for which $\succsim_{A'}$ must hold. It is likely that a similar existence theorem could be developed for other properties of interpersonal preferences (e.g., inequity aversion); however, there is no guarantee that the conditions and proof for such a theorem would be similar to those of Theorem 1.

If the conditions for Theorem 1 hold and an altruistic value function $v^{A'}$ exists, a natural question arises regarding what form $v^{A'}$ might take. In particular, assessment can be eased if preferences can be represented by an additive value function. To show that an additive altruistic value function exists, we require the Thomsen condition, which is a particular form of double cancelation. The Thomsen condition asserts that if $(x_1^A, x_2^B) \sim_{A'} (x_2^A, x_1^B)$ and $(x_2^A, x_3^B) \sim_{A'} (x_3^A, x_2^B)$, then $(x_1^A, x_3^B) \sim_{A'} (x_3^A, x_1^B)$. It is commonly used as a condition for the existence of additive value functions of two attributes (Krantz et al. 1971; Keeney and Raiffa 1976). Alternatively, the hexagon condition presented by Karni and Safra (1998), which is the two-attribute case of the corresponding tradeoff condition used by Keeney and Raiffa (1976), may be used instead.

Theorem 2 *The following two statements are equivalent for a given* $\succsim_A$, $\succsim_B$, X, *and a preference relation* $\succsim_{A'}$ *satisfying the conditions of Theorem 1:*

1. $\succsim_{A'}$ *satisfies the Thomsen condition.*
2. *There exists a continuous ordinal* $v^{A'}(x)$ *representing* $\succsim_{A'}$ *given by:*

$$v^{A'}(x) = w^A v^A \left(x^A\right) + w^B v^B \left(x^B\right), \tag{1}$$

where v^A *and* v^B *represent* $\succsim_A$ *and* $\succsim_B$*, respectively, and* w^A *and* w^B *are positive constants.*

See the Appendix for the proof of Theorem 2.

If an altruistic value function of the form shown in Eq. 1 exists, it can be assessed from a decision maker using procedures similar to those used for additive multiattribute value functions. Midvalue splitting (Keeney and Raiffa 1976; Kirkwood 1997) can be used to assess v^A and v^B, where the midvalue of $[x_1^A, x_2^A]$ is the level x_{mid}^A such that if $(x_1^A, x_2^B) \sim_{A'} (x_{\text{mid}}^A, x_1^B)$, for some $x_1^B, x_2^B \in X^B$, then $(x_{\text{mid}}^A, x_2^B) \sim_{A'} (x_2^A, x_1^B)$. That is, expressed via tradeoffs with X^B, the increase in value achieved by moving from x_1^A to x_{mid}^A is equal to the increase in value achieved by moving from x_{mid}^A to x_2^A. Alternatively, if the appropriate conditions on preferences hold, specific functional forms can be fit to v^A and v^B.

The weights w^A and w^B, by convention, are assumed to sum to 1. One straightforward approach for assessing w^A and w^B is as follows. Let x_0^A and x_1^A represent the worst and best levels, respectively, of x^A, and x_0^B and x_1^B represent the worst and best levels, respectively, of x^B. Then, ask the decision maker which outcome of (x_0^A, x_1^B) and (x_1^A, x_0^B) is preferred. Arbitrarily, imagine the decision maker judges (x_1^A, x_0^B) to be preferable. Then determine a level x_i^A such that the decision maker is indifferent between (x_i^A, x_1^B) and (x_1^A, x_0^B). If v^A has already been assessed, then $v^A(x_i^A)$ is known, and we can solve for the two weights. For more details on this type of approach for weight elicitation, see Keeney and Raiffa (1976) or Eisenfuhr et al. (2010).

4 Cardinal altruistic value functions

It is also possible to represent relative differences in preference between outcomes through a cardinal value function if additional information about preferences is specified. (Dyer and Sarin 1979 refer to these functions as measurable value functions). In particular, we define the binary preference difference relation $\succsim_A^*$ on the set of pairs of outcomes $x_1^A, x_2^A \in X^A$, which we will write as $X^A \times X^A$, with the interpretation that $x_1^A x_2^A \succsim_A^* x_3^A x_4^A$ means the difference in desirability between x_1^A and x_2^A is at least as great as the difference in desirability between x_3^A and x_4^A. $\succsim_B^*$ can be defined similarly. To ensure that $\succsim_A^*$ is consistent with the underlying preferences expressed by $\succsim_A$, we require that $x_1 \succsim_A x_2$ and $x_3 \succsim_A x_4$ implies $x_1 x_4 \succsim_A^* x_2 x_3$, and similarly for $\succsim_B^*$. We restrict the set of outcome pairs to those in which the first outcome is at least as desirable as the second outcome. That is, we do not allow "negative" differences between outcomes. Allowing negative differences would give rise to an algebraic difference structure, whereas we will use a positive difference structure. Similar results can be obtained using either approach. Corresponding strict and indifference relations for $\succsim_A^*$ and $\succsim_B^*$ can also be expressed as they were for the previous binary relations in this paper.

Just as conditions on $\succsim_A$ were needed to establish the existence of an ordinal value function representing $\succsim_A$, conditions on $\succsim_A^*$ are needed to establish the existence of a cardinal value function representing $\succsim_A^*$. To obtain such cardinal value functions, we must first define two properties of these relations:

Definition 8 $\succsim_A^*$ is weak monotonic if $x_1^A x_2^A \succsim_A^* x_4^A x_5^A$ and $x_2^A x_3^A \succsim_A^* x_5^A x_6^A$ implies $x_1^A x_3^A \succsim_A^* x_4^A x_6^A$ for all $x_1^A, \ldots, x_6^A \in X^A$.

Definition 9 $\succsim_A^*$ is solvable if for any $x_1^A, x_2^A, x_3^A, x_4^A \in X^A$ such that $x_1^A x_2^A \succsim_A^* x_3^A x_4^A$ and $x_3^A x_4^A \succsim_A^* x_1^A x_1^A$, there exist $x_5^A, x_6^A \in X^A$ such that $x_1^A x_5^A \sim_A^* x_3^A x_4^A$ and $x_6^A x_2^A \sim_A^* x_3^A x_4^A$.

Weak monotonicity ensures that preference differences can be concatenated. That is, a weak monotonic preference difference relation is modular; if a preference difference is split into smaller preference differences, the aggregate of these smaller differences cannot be greater or less than the original preference difference. Solvability ensures that a positive preference difference can be reproduced within a larger preference difference by another pair of outcomes. It is similar in nature to the indifference property in Sect. 3; intuitively, it asserts that an equivalent positive preference difference exists somewhere "between" a larger preference difference and a zero preference difference.

Let $\succsim_A^*$ be complete, transitive, weak monotonic, and solvable, and satisfy a technical condition called the Archimedean property (defined in the Appendix in the proof of Theorem 2). Then, based on Theorem 1 of Section 4.2 of Krantz et al. (1971)[1], there exists a value function $v^A(x^A)$ such that for any $x_1^A, x_2^A, x_3^A, x_4^A \in X^A$, $v^A(x_1^A) - v^A(x_2^A) \geq v^A(x_3^A) - v^A(x_4^A)$ iff $x_1^A x_2^A \succsim_A^* x_3^A x_4^A$. This value function is unique up to positive linear transformations.

Analogously, we can define altruistic preference difference relations $\succsim_{A'}^*$ and $\succsim_{B'}^*$ on $X \times X$. We will consider only $\succsim_{A'}^*$ henceforth; the exposition for $\succsim_{B'}^*$ is identical.

[1] The theorem from Krantz et al. (1971) requires two additional conditions that are met trivially by $\succsim_A^*$ and $\succsim_B^*$ due to their consistency with $\succsim_A$ and $\succsim_B$. These conditions are stated in the proof of Theorem 3.

Definition 10 $\succsim^*_{A'}$ satisfies the Pareto difference property if for any $x_1, x_2, x_3, x_4 \in X$, $x_1^A x_2^A \succsim^*_A x_3^A x_4^A$ and $x_1^B x_2^B \succsim^*_B x_3^B x_4^B$ implies $x_1 x_2 \succsim^*_{A'} x_3 x_4$.

The Pareto difference property ensures consistency with the difference orderings induced by $\succsim^*_A$ and $\succsim^*_B$; it is analogous to the Pareto property in Sect. 3.

Definition 11 $\succsim^*_{A'}$ is an altruistic preference difference relation iff $\succsim_{A'}$ satisfies the conditions of Theorem 1, $\succsim^*_{A'}$ satisfies the Pareto difference property, and $x_1 \succsim_{A'} x_2$ and $x_3 \succsim_{A'} x_4$ implies $x_1 x_4 \succsim^*_{A'} x_2 x_3$ (for consistency with the ordering induced by $\succsim_{A'}$).

Note that $\succsim_{A'}$ still satisfies the Pareto property. In that sense, $\succsim^*_{A'}$ reflects altruistic ordinal preferences even without the Pareto difference property. An altruistic preference difference relation, however, is also consistent with each individual's relative preference differences between outcomes.

As in the previous section, we would like to determine whether or not $\succsim^*_{A'}$ can be represented by a corresponding value function. Unfortunately, as before, the existence of cardinal v^A and v^B representing $\succsim^*_A$ and $\succsim^*_B$ does not guarantee the existence of cardinal $v^{A'}$ representing $\succsim^*_{A'}$, and therefore additional conditions are needed.

Similarly to the previous section, we do not assume that $\succsim^*_{A'}$ induces a complete ordering on the positive difference pairs in $X \times X$. We define two new relevant properties for $\succsim^*_{A'}$:

Definition 12 $\succsim^*_{A'}$ satisfies the cardinal substitution property if for any $x_1, x_2, x_3, x_4 \in X$ such that $x_1 x_2 \sim^*_{A'} x_3 x_4$, it holds for all $x_5, x_6 \in X$ that $x_1 x_2 \succsim^*_{A'} x_5 x_6$ iff $x_3 x_4 \succsim^*_{A'} x_5 x_6$, and $x_5 x_6 \succsim^*_{A'} x_1 x_2$ iff $x_5 x_6 \succsim^*_{A'} x_3 x_4$.

Definition 13 $\succsim^*_{A'}$ satisfies the cardinal indifference property if for all $x_1, x_2, x_3, x_4 \in X$ such that $x_1 x_2 \succsim^*_{A'} x_3 x_4$, there exist $x_5, x_6 \in X$ such that $x_1 x_5 \sim^*_{A'} x_3 x_4$ and $x_6 x_2 \sim^*_{A'} x_3 x_4$, where $x_1 x_2 \succsim^*_{A'} x_1 x_5$ and $x_1 x_2 \succsim^*_{A'} x_6 x_2$.

The cardinal substitution property is analogous to the substitution property used previously. The cardinal indifference property, while in the same spirit as the indifference property from Sect. 3, is actually more similar to the solvability property on $\succsim^*_A$ and $\succsim^*_B$, as it involves reproducing a preference difference within a larger preference difference.

Theorem 3 *Let $\succsim_{A'}$ satisfy the conditions of Theorem 1, and let $\succsim^*_A$ and $\succsim^*_B$ be reversible, weak monotonic, and solvable. Then the following two statements are equivalent:*

1. *$\succsim^*_{A'}$ is an altruistic preference difference relation satisfying the cardinal substitution and cardinal indifference properties.*
2. *There exists a continuous cardinal $v^{A'}(x)$ representing $\succsim^*_{A'}$ that is monotonically increasing in x^A and x^B.*

See the Appendix for the proof of Theorem 3. As for cardinal $v^A(x^A)$ and $v^B(x^B)$, $v^{A'}(x)$ is unique up to positive linear transformations. Given the existence of

$v^{A'}(x)$, it is also helpful to consider the possibility that $v^{A'}(x)$ may be additive. As in the case of ordinal value functions, we will use the Thomsen condition. As previously, the hexagon condition could also be used to obtain a similar result. Alternatively, Dyer and Sarin (1979) present a property called difference independence on a preference difference relation that could also be used to show the existence of an additive function representing $\succsim^*_{A'}$.

Theorem 4 *If the conditions of Theorem 3 hold, then the following two statements are equivalent:*

1. $\succsim_{A'}$ *satisfies the Thomsen condition.*
2. *There exists a continuous* $v^{A'}(x)$ *such that:*

$$v^{A'}(x) = w^A v^A\left(x^A\right) + w^B v^B\left(x^B\right), \tag{2}$$

where w^A *and* w^B *are positive constants, and* $v^{A'}$ *represents* $\succsim^*_{A'}$.

The proof is similar to the proof of Theorem 2.

The assessment methods for ordinal additive value functions described at the end of the previous section are applicable to cardinal altruistic value functions as well. However, there are additional methods taking advantage of interval scales that may be more accessible for decision makers. If a midvalue splitting approach is used, the midvalue of $[x_1^A, x_2^A]$ can now be described as the level x_{mid}^A such that $x_2^A x_{\text{mid}}^A \sim^*_{A'} x_{\text{mid}}^A x_1^A$. Weights can be determined using a swing weighting approach (von Winterfeldt and Edwards 1986). As previously, let x_0^A and x_1^A represent the worst and best levels, respectively, of x^A, and x_0^B and x_1^B represent the worst and best levels, respectively, of x^B. First, the decision maker is asked to consider the worst possible outcome, in this case (x_0^A, x_0^B). If $(x_1^A, x_0^B)(x_0^A, x_0^B) \succsim^*_{A'} (x_0^A, x_1^B)(x_0^A, x_0^B)$, then the decision maker is asked to identify the magnitude of the difference between (x_0^A, x_1^B) and (x_0^A, x_0^B) as a percentage of the difference between (x_1^A, x_0^B) and (x_0^A, x_0^B), which allows us to solve for the two weights.

5 Altruistic utility functions

It may also be desirable to examine an altruistic decision maker's preferences over gambles, i.e., decisions for which the realized outcome will be uncertain. In this section, we consider both intrinsic and altruistic preference relations over gambles, and provide representation theorems yielding altruistic utility functions. The main challenge is that the structure of the set of possible gambles differs from that of the set of possible outcomes used in previous sections, and therefore several of the earlier concepts and conditions will need to be modified accordingly.

Let G denote the set of all $\{p_1 x_1, p_2 x_2, \ldots\}$ such that $x_1, x_2, \ldots \in X$, $p_1, p_2, \ldots \geq 0$, and $\sum_i p_i = 1$. That is, G is the set of all possible gambles over elements of X. We assume common beliefs; the individuals do not disagree on the probabilities of the outcomes. Let g_1, g_2, g_3 represent arbitrary gambles, i.e., elements of G, and let $\succsim^G$ denote a preference relation over elements of G, with corresponding strict and indifference relations. The definitions of completeness, transitivity, and continuity can be updated accordingly:

Definition 14 $\succsim_A^G$ is complete if for any $g_1, g_2 \in G$, either $g_1 \succsim_A^G g_2$, $g_2 \succsim_A^G g_1$, or both.

Definition 15 $\succsim_A^G$ is transitive if for any $g_1, g_2, g_3 \in G$, $g_1 \succsim_A^G g_2$ and $g_2 \succsim_A^G g_3$ implies $g_1 \succsim_A^G g_3$.

Definition 16 $\succsim_A^G$ is continuous if for any $g_1, g_2, g_3 \in G$ for which $g_1 \succ_A^G g_2$ and $g_2 \succ_A^G g_3$, $\exists \varepsilon > 0$ such that $((1-\varepsilon)g_1 + \varepsilon g_3) \succ_A^G g_2$ and $g_2 \succ_A^G (\varepsilon g_1 + (1-\varepsilon)g_3)$.

There are many equivalent ways in which continuity of gambles can be expressed; the definition given above is sometimes referred to as the Archimedean property. It asserts that if g_1 is strictly preferred to g_2, a mixture of gambles in which an undesirable gamble g_3 occurs with infinitesimal probability and g_1 is taken in all other cases must also be strictly preferred to g_2, and similarly, g_2 must be strictly preferred to a mixture of gambles involving g_3 and an infinitesimal probability of g_1. Similarly to value functions, these three conditions are sufficient to establish the existence of a function over gambles representing $\succsim_A^G$. However, without additional conditions, the form of this function is unspecified. The standard approach to utility theory (von Neumann and Morgenstern 1947) is to assert one further condition:

Definition 17 $\succsim_A^G$ satisfies the independence property if for any $g_1, g_2 \in G$ for which $g_1 \succ_A^G g_2$, and any $g_3 \in G$, it must be true that for any $p \in (0, 1]$, $pg_1 + (1-p)g_3 \succ_A^G pg_2 + (1-p)g_3$.

That is, if g_1 is strictly preferred to g_2, a probability of g_1 must be strictly preferred to the same probability of g_2 (regardless of the common third gamble). Any of several equivalent conditions may be used to reach the same result. If this fourth condition also holds, then there exists a utility function $u^A(x)$, unique up to positive linear transformations, such that for any $g_1, g_2 \in G$, $g_1 \succsim_A^G g_2$ iff the expected value of $u^A(x)$ under g_1 is greater than or equal to the expected value of $u^A(x)$ under g_2. We assume that $\succsim_A^G$ and $\succsim_B^G$ satisfy completeness, transitivity, continuity, and independence.

Given the existence of $u^A(x)$ and $u^B(x)$ representing $\succsim_A^G$ and $\succsim_B^G$, respectively, we construct an altruistic preference relation $\succsim_{A'}^G$ over G, and develop conditions that establish the existence of an altruistic utility function $u^{A'}(x)$ representing $\succsim_{A'}^G$. As previously, we assume that $\succsim_{A'}^G$ satisfies the Pareto property; in this section, the Pareto property describes elements of G rather than elements of X. We do not assume completeness, transitivity, or continuity of $\succsim_{A'}^G$. Instead, two further conditions are needed for $\succsim_{A'}^G$; they are analogous to the indifference and substitution conditions used previously, describing gambles rather than certain outcomes:

Definition 18 $\succsim_{A'}^G$ satisfies the indifference property if for any $g_1, g_2, g_3 \in G$ for which $g_1 \succ_{A'}^G g_2$ and $g_2 \succ_{A'}^G g_3$, $\exists p \in [0, 1]$ such that $(pg_1 + (1-p)g_3) \sim_{A'}^G g_2$.

Definition 19 $\succsim_{A'}^G$ satisfies the substitution property if for any $g_1, g_2, g_3 \in G$ for which $g_1 \sim_{A'}^G g_2$, $g_1 \succsim_{A'}^G g_3$ iff $g_2 \succsim_{A'}^G g_3$.

The gamble indifference property ensures that, if g_1 is preferred to g_2 and g_2 is preferred to g_3, there must be a mixture of g_1 and g_3 such that individual A is indifferent between the mixture and g_2. The substitution property ensures that indifferent gambles can always be substituted for one another without affecting the truth of a comparison.

Theorem 5 *The following two statements are equivalent for given $\succsim_A^G$, $\succsim_B^G$, G, and a preference relation $\succsim_{A'}^G$ defined based on $\succsim_A^G$ and $\succsim_B^G$:*

1. *$\succsim_{A'}^G$ is an altruistic preference relation satisfying the independence, indifference and substitution properties.*
2. *There exists a von Neumann-Morgenstern utility function $u^{A'}(x)$ representing $\succsim_{A'}^G$ that is monotonically increasing in x^A and x^B.*

See the Appendix for the proof of Theorem 5.

As in the case of altruistic value functions, it is convenient to use an additive form for $u^{A'}$. There is a substantial body of existing work on additive multiattribute utility functions; see, for instance, Fleming (1952), Harsanyi (1955), and Fishburn (1965). Many possible conditions could be used to give rise to an additive altruistic utility function. We use the following condition, adapted from Fishburn (1965):

Definition 20 $\succsim_{A'}^G$ satisfies mutual independence if for any $g_1, g_2 \in G$ such that the probability of any given level of x^A is equal for g_1 and g_2, and the same holds for any given level of x^B, it must be true that $g_1 \sim_{A'}^G g_2$.

Mutual independence asserts that the only the marginal distributions of x^A and x^B are relevant, that is, it does not matter which of individual A's and individual B's outcomes occur together. Note that preferences for equity are likely to violate this condition.

Theorem 6 *The following two statements are equivalent for given $\succsim_A^G$, $\succsim_B^G$, G, and a preference relation $\succsim_{A'}^G$ satisfying the conditions of Theorem 5:*

1. *$\succsim_{A'}^G$ satisfies mutual independence.*
2. *There exists a von Neumann-Morgenstern utility function $u^{A'}(x)$ representing $\succsim_{A'}^G$ given by:*
$$u^{A'}(x) = w^A u^A(x) + w^B u^B(x), \tag{3}$$
where u^A and u^B represent $\succsim_A^G$ and $\succsim_B^G$, respectively, and w^A and w^B are positive constants.

Theorem 6 follows directly from Theorem 5 and Fishburn (1965).

$u^{A'}(x)$ can be assessed using any standard utility assessment procedures. A review of many such methods is provided by Farquhar (1984).

6 Discussion and conclusion

It is well-understood that the behavior of humans often reflects preferences which include altruistic components, and a great deal of research has been done to justify this behavior from different perspectives. Much work has also been done applying value and utility functions that capture altruistic preferences. This paper provides a basic theoretical framework to establish the existence of altruistic value and utility functions based on simple preference conditions, beginning with the development of what it means for preferences to be altruistic. Both ordinal and cardinal value functions are considered. The altruistic value and utility functions can be shown to exist without assuming explicitly that altruistic preference relations are complete, transitive, or continuous; weaker conditions of substitution and indifference will suffice. Weakening the required preference conditions is beneficial when aggregating the preferences of multiple individuals, as the conditions typically used for individual value and utility functions are substantially stronger in this case. Finally, this paper provides additional conditions that can be used to establish additive forms for the altruistic value and utility functions.

There are several possibilities for further research on this topic. One possibility is to consider altruistic preferences concerning more than two individuals, and to develop analogous preference conditions that would establish the existence of altruistic value or utility functions. The Pareto property can be expanded easily to capture more than two individuals' intrinsic preferences, as can many of the preference conditions. Some portions of the proofs in this paper, however, rely on there being only two individuals, and thus would have to be reworked accordingly.

Another possible direction of work is establishing the existence of value or utility functions based on interpersonal preference properties other than the Pareto property. For example, it might be possible to show that similar conditions can establish value or utility functions when individual A has preferences that are piecewise monotonic in individual B's outcome such that A is altruistic when doing "better" than B, but spiteful when doing "worse" than B.

Acknowledgments The author is grateful to Jon Baron, David Krantz, Cameron MacKenzie, an editor, and two anonymous reviewers for helpful comments on earlier versions of this paper.

Appendix

Proof (Theorem 1) We will first show that statement 1 implies statement 2.

Based on the results of Debreu (1954, 1964), there exists a $v^{A'}\left(x^A, x^B\right)$ such that $v^{A'}\left(x_1^A, x_1^B\right) \geq v^{A'}\left(x_2^A, x_2^B\right)$ iff $\left(x_1^A, x_1^B\right) \succsim_{A'} \left(x_2^A, x_2^B\right)$ provided the following conditions hold:

Completeness: Either $\left(x_1^A, x_1^B\right) \succsim_{A'} \left(x_2^A, x_2^B\right)$ or $\left(x_2^A, x_2^B\right) \succsim_{A'} \left(x_1^A, x_1^B\right)$, or both.

Transitivity: $\left(x_1^A, x_1^B\right) \succsim_{A'} \left(x_2^A, x_2^B\right)$ and $\left(x_2^A, x_2^B\right) \succsim_{A'} \left(x_3^A, x_3^B\right)$ implies $\left(x_1^A, x_1^B\right) \succsim_{A'} \left(x_3^A, x_3^B\right)$.

Continuity: For any $x_1, x_2 \in X$ for which $x_1 \succ_A x_2$, $\exists \Delta > 0$ such that for any $x_3 \in X$, $\max(|x_1^A - x_3^A|, |x_1^B - x_3^B|) < \Delta$ implies $x_3 \succ_{A'} x_2$, and $\max(|x_3^A - x_2^A|, |x_3^B - x_2^B|) < \Delta$ implies $x_1 \succ_A x_3$.

The completeness condition can be shown to hold by considering four possible cases, which are collectively exhaustive by completeness of $\succsim_A$ and $\succsim_B$. In case 1,

$x_1^A \succsim_A x_2^A$ and $x_1^B \succsim_B x_2^B$, and in case 2, $x_2^A \succsim_A x_1^A$ and $x_2^B \succsim_B x_1^B$. The Pareto property ensures that $x_1 \succsim_{A'} x_2$ in case 1, and $x_2 \succsim_{A'} x_1$ in case 2.

In case 3, $x_1^A \succsim_A x_2^A$ and $x_2^B \succsim_B x_1^B$. Consider the outcomes (x_1^A, x_2^B) and (x_2^A, x_1^B). By the Pareto property, $(x_1^A, x_2^B) \succsim_{A'} x_1$, and $x_1 \succsim_{A'} (x_2^A, x_1^B)$. Then, by the indifference property, there exists an outcome x_1' on the line segment between (x_1^A, x_2^B) and (x_2^A, x_1^B) such that $x_1' \sim_{A'} x_1$. By the same argument, there exists an outcome x_2' on the line segment between (x_1^A, x_2^B) and (x_2^A, x_1^B) such that $x_2' \sim_{A'} x_2$. For any pair of outcomes on this line segment, one is Pareto superior to the other. Hence, by the Pareto property, either $x_1' \succsim_{A'} x_2'$ or $x_2' \succsim_{A'} x_1'$. Applying the substitution property twice to both comparisons yields that either $x_1 \succsim_{A'} x_2$ or $x_2 \succsim_{A'} x_1$.

In case 4, $x_2^A \succsim_A x_1^A$ and $x_1^B \succsim_B x_2^B$. Using the same argument as for case 3, either $x_1 \succsim_{A'} x_2$ or $x_2 \succsim_{A'} x_1$. Hence, completeness of $\succsim_{A'}$ holds.

To establish the transitivity condition, consider $x_1, x_2, x_3 \in X$ such that $x_1 \succsim_{A'} x_2$ and $x_2 \succsim_{A'} x_3$. If x_1 is Pareto superior to x_2 and x_2 is Pareto superior to x_3, then $x_1 \succsim_{A'} x_3$ trivially by the Pareto property. We must consider the remaining three cases: x_1 is not Pareto superior to x_2, x_2 is not Pareto superior to x_3, and both.

First, consider the case where x_1 is not Pareto superior to x_2, but x_2 is Pareto superior to x_3. Arbitrarily, let $x_1^A \geq x_2^A$ and $x_1^B < x_2^B$. Then consider the outcome (x_1^A, x_2^B), which must be preferred to x_1 by the Pareto property. By the indifference property, there exists an outcome (z^A, x_2^B) such that $z^A > x_2^A$ and $(z^A, x_2^B) \sim_{A'} x_1$. Since x_2 is Pareto superior to x_3, (z^A, x_2^B) must also be Pareto superior to x_3, hence $(z^A, x_2^B) \succsim_{A'} x_3$ by the Pareto property. Then, by the substitution property, $x_1 \succsim_{A'} x_3$, establishing transitivity for this case. In the case where x_2 is not Pareto superior to x_3, but x_1 is Pareto superior to x_2, it can be shown in a similar manner that $x_1 \succsim_{A'} x_3$.

Finally, consider the case where x_1 is not Pareto superior to x_2, and x_2 is not Pareto superior to x_3. Assume that x_1 is not Pareto superior to x_3 (if it were, the Pareto property would establish trivially that $x_1 \succsim_{A'} x_3$). Arbitrarily, let $x_1^A \geq x_2^A$ and $x_1^B < x_2^B$, and let $x_2^A < x_3^A$ and $x_2^B \geq x_3^B$. Consider the outcome (x_1^A, x_2^B), which must be preferred to x_1 by the Pareto property. By the indifference property, there exists an outcome (z^A, x_2^B) such that $z^A > x_2^A$ and $(z^A, x_2^B) \sim_{A'} x_1$. Similarly, consider the outcome (x_2^A, x_3^B), over which x_3 must be preferred by the Pareto property. By the indifference property, there exists an outcome (x_2^A, z^B) such that $x_2^B > z^B$ and $(x_2^A, z^B) \sim_{A'} x_3$. Since $z^A > x_2^A$ and $x_2^B > z^B$, (z^A, x_2^B) is Pareto superior to (x_2^A, z^B), and by the Pareto property, $(z^A, x_2^B) \succsim_{A'} (x_2^A, z^B)$. Applying the substitution property twice yields $x_1 \succsim_{A'} x_3$, establishing transitivity.

To establish the continuity condition, first consider arbitrary $x_1, x_2 \in X$ such that $x_1 \succ_{A'} x_2$. If x_1 is strictly Pareto superior to x_2, i.e., $x_1^A > x_2^A$ and $x_1^B > x_2^B$, then continuity can be shown to hold easily; let Δ be any positive number less than the smaller of $x_1^A - x_2^A$ and $x_1^B - x_2^B$.

If x_1 is Pareto superior to x_2 with $x_1^B = x_2^B$ (the logic would be similar for the case where $x_1^A = x_2^A$), if there exists some $x_3^B < x_1^B, x_2^B$, then consider the point (x_2^A, x_3^B). By the indifference property, there exists an outcome x_4 on the line segment between (x_2^A, x_3^B) and x_1 such that $x_2 \sim_{A'} x_4$. Since x_1 is strictly Pareto superior to x_4, any positive Δ_1 less than the smaller of $x_1^A - x_4^A$ and $x_1^B - x_4^B$ will suffice for the comparison of outcomes close to x_1 with x_2. If there does not exist an $x_3^B < x_1^B, x_2^B$ (i.e., these outcomes are the least preferable according to $\succsim_B$), then such a Δ_1 exists trivially by continuity of $\succsim_A$ and the Pareto property. Similar logic can be used to obtain a positive Δ_2 for the comparison of outcomes close to x_2 with x_1. We can then let $\Delta = \min\{\Delta_1, \Delta_2\}$. Thus, if x_1 is Pareto superior to x_2, there exists a Δ as specified.

If x_1 is not Pareto superior to x_2, continuity can be established by showing that there is an outcome that is Pareto superior to x_2 and indifferent to x_1 and, correspondingly, an outcome to which x_1 is Pareto superior that is indifferent to x_2. We will show the former; the approach for the latter is similar. Arbitrarily, let $x_1^B \leq x_2^B$. Consider any outcome x_3 such that $x_2^A \leq x_3^A \leq x_1^A, x_3^B \geq x_2^B$, and $x_3 \succsim_{A'} x_1$. By the indifference property, there must be an outcome x_4 on the line segment between x_2 and x_3 such that $x_1 \sim_{A'} x_4$. Since x_4 is Pareto superior to x_2, a Δ satisfying the requirement exists as shown above. Thus, $\succsim_{A'}$ is continuous, establishing the existence of a $v^{A'}$ representing $\succsim_{A'}$.

Finally, we need to show that $v^{A'}(x^A, x^B)$ is monotonically increasing in x^A and x^B. This can be done easily by observing that v^A is monotonically increasing in (x^A) and unaffected by x^B, and v^B is monotonically increasing in (x^B) and unaffected by x^A. Consider $x_1, x_2 \in X$. If $x_1^A > x_2^A$ and $x_1^B = x_2^B$, then $x_1 \succ_A x_2$ and $x_1 \sim_B x_2$. By the Pareto property, $x_1 \succ_{A'} x_2$. Since $v^{A'}$ represents $\succsim_{A'}$, it follows that $v^{A'}(x_1) > v^{A'}(x_2)$. Similar logic applies if $x_1^A = x_2^A$ and $x_1^B > x_2^B$, or if both $x_1^A = x_2^A$ and $x_1^B = x_2^B$. Thus, $v^{A'}$ is monotonically increasing in x^A and x^B, establishing the first half of Theorem 1.

To establish the converse, assume $v^{A'}$ represents $\succsim_{A'}$, and is continuous and monotonically increasing in x^A and x^B. It will suffice to show that $\succsim_{A'}$ must satisfy the Pareto, substitution, and indifference properties.

To show that $\succsim_{A'}$ satisfies the Pareto property, consider $x_1, x_2 \in X$ such that $x_1 \succsim_A x_2$ and $x_1 \succsim_B x_2$. We have assumed that $\succsim_A$ and $\succsim_B$ are monotonic, so this is equivalent to stating that $x_1^A \geq x_2^A$ and $x_1^B \geq x_2^B$. Since $v^{A'}$ is monotonically increasing in x^A and x^B, it must be true that $v^{A'}(x_1) \geq v^{A'}(x_2)$, and because $v^{A'}$ represents $\succsim_{A'}$, it follows that $x_1 \succsim_{A'} x_2$. Similar logic can be used to establish the strict relation case.

To show that $\succsim_{A'}$ satisfies the substitution property, consider $x_1, x_2, x_3 \in X$ such that $x_1 \sim_{A'} x_2$. Because $v^{A'}$ represents $\succsim_{A'}$, it must be true that $v^{A'}(x_1) = v^{A'}(x_2)$. Consider $v^{A'}(x_3)$. Clearly, $v^{A'}(x_1) \geq v^{A'}(x_3)$ iff $v^{A'}(x_2) \geq v^{A'}(x_3)$, and $v^{A'}(x_3) \geq v^{A'}(x_1)$ iff $v^{A'}(x_3) \geq v^{A'}(x_2)$. Again, because $v^{A'}$ represents $\succsim_{A'}$, it must then be true that $x_1 \succsim_{A'} x_3$ iff $x_2 \succsim_{A'} x_3$, and $x_3 \succsim_{A'} x_1$ iff $x_3 \succsim_{A'} x_2$, establishing the substitution property.

To show that $\succsim_{A'}$ satisfies the indifference property, consider $x_1, x_2, x_3 \in X$ such that $x_1 \succsim_{A'} x_2$ and $x_2 \succsim_{A'} x_3$. Because $v^{A'}$ represents $\succsim_{A'}$, it must be true that $v^{A'}(x_1) \geq v^{A'}(x_2) \geq v^{A'}(x_3)$. Since $v^{A'}$ is continuous, the intermediate value theorem establishes the existence of a real number $k \in \{0, 1\}$, as described in the definition of indifference, such that $v^{A'}(x_1^A + km^A, x_1^B + km^B) = v^{A'}(x_2)$. Again, because $v^{A'}$ represents $\succsim_{A'}$, it must be true that $(x_1^A + km^A, x_1^B + km^B) \sim_{A'} x_2$, establishing the indifference property, and hence the converse direction of the Theorem. □

Proof (Theorem 2) We will first show that statement 1 implies statement 2. Based on Holman (1971) and Krantz et al. (1971), a relation on $X^A \times X^B$ can be represented by an additive function $\phi^A(x^A) + \phi^B(x^B)$ if the following conditions hold (we use

ϕ^A and ϕ^B to distinguish from the individually obtained v^A and v^B, for reasons given at the end of this direction of the proof):

Weak ordering: $\succsim_{A'}$ is complete and transitive.

Independence: $(x_1^A, x_0^B) \succsim_{A'} (x_2^A, x_0^B)$ for some x_0^B implies $(x_1^A, x^B) \succsim_{A'} (x_2^A, x^B)$ $\forall x^B \in X^B$, and $(x_0^A, x_1^B) \succsim_{A'} (x_0^A, x_2^B)$ for some x_0^A implies $(x^A, x_1^B) \succsim_{A'} (x^A, x_2^B) \forall x^A \in X^A$.

Thomsen condition.

Restricted solvability: Given $x_1^A, x_{\underline{2}}^A, x_{\overline{2}}^A \in X^A$ and $x_1^B, x_2^B \in X^B$ such that $(x_{\overline{2}}^A, x_2^B) \succsim_{A'} (x_1^A, x_1^B)$ and $(x_1^A, x_1^B) \succsim_{A'} (x_{\underline{2}}^A, x_2^B)$, $\exists x_2^B \in X^B$ such that $(x_1^A, x_1^B) \sim_{A'} (x_2^A, x_2^B)$, and the same property holds for X^B.

Archimedean property: Every strictly bounded standard sequence (defined below) is finite.

Each component is essential: There exist $x_1^A, x_2^A \in X^A$ and $x_0^B \in X^B$ such that $(x_1^A, x_0^B) \succ_{A'} (x_2^A, x_0^B)$, and similarly for X^B.

Several of these conditions hold trivially, as this set overlaps somewhat with the set used to establish the existence of $v^{A'}$. Weak ordering holds, since completeness and transitivity were derived to establish Theorem 1. Independence follows easily from the Pareto property. The Thomsen condition is stated as a condition for this Theorem. Restricted solvability is a special case of the indifference property. Because $\succsim_A$ and $\succsim_B$ are non-trivial, the Pareto property ensures that each component is essential.

The only remaining condition is the Archimedean property. A standard sequence on X^A is a sequence of x_i^A such that, for specific distinct x_1^B, x_2^B that are not indifferent to one another, $(x_{i+1}^A, x_2^B) \sim_{A'} (x_i^A, x_1^B)$. If $x_1^B \succ_B x_2^B$, then $(x_i^A, x_1^B) \succ_{A'} (x_i^A, x_2^B)$ and thus $(x_{i+1}^A, x_1^B) \succ_{A'} (x_i^A, x_1^B)$. An infinite strictly bounded sequence would imply that as $i \to \infty$, there exists a limiting x_L^A such that $(x_L^A, x_2^B) \succ_{A'} (x_i^A, x_1^B)$ for x_i^A infinitesimally close to x_L^A. By the Pareto property, it must be true that $(x_L^A, x_1^B) \succ_{A'} (x_L^A, x_2^B)$. Thus, by the indifference property, there would be an outcome on the line segment between (x_i^A, x_1^B) and (x_L^A, x_1^B), for any i, that is indifferent to (x_L^A, x_2^B). However, this creates a contradiction, because (x_L^A, x_2^B) is strictly preferred to all such points by definition of x_L^A. A similar argument can be made for standard sequences on X^B. Hence, the Archimedean property holds, and $\succsim_{A'}$ can be represented by an additive function $\phi^A(x^A) + \phi^B(x^B)$. If $\phi^{A'}(x^A) + \phi^{B'}(x^B)$ represents $\succsim_{A'}$, then it must be true that $\phi^{A'} = \alpha\phi^A + \beta_1$ and $\phi^{B'} = \alpha\phi^B + \beta_2$ (Krantz et al. 1971), and therefore we can obtain the functional form of Eq. 1.

By the Pareto property, it must be true that $\phi^A(x^A)$ and $\phi^B(x^B)$ represent $\succsim_A$ and $\succsim_B$, respectively, as do any $\phi^{A'}(x^A)$ and $\phi^{B'}(x^B)$ obtained by a positive linear transformation, establishing statement 2 of Theorem 2.

Note that $\phi^A(x^A)$, $\phi^B(x^B)$, $\phi^{A'}(x^A)$, and $\phi^{B'}(x^B)$ are positive monotonic transformations of any given v^A and v^B. However, Eq. 1 permits only positive linear transformations of v^A and v^B. Therefore, we cannot simply assess ordinal v^A and v^B individually and plug them into Eq. 1; the ϕ functions in this proof should be viewed as distinct from previously obtained ordinal v^A and v^B.

To establish the converse, assume $v^{A'}$ represents $\succsim_{A'}$, is continuous, and is given by $v^{A'}(x) = w^A v^A(x^A) + w^B v^B(x^B)$, as stated in the Theorem. It will suffice to show that $\succsim_{A'}$ satisfies the Thomsen condition.

Consider outcomes (x_1^A, x_2^B), (x_2^A, x_1^B), (x_2^A, x_3^B), $(x_3^A, x_2^B) \in X$ such that $(x_1^A, x_2^B) \sim_{A'} (x_2^A, x_1^B)$ and $(x_2^A, x_3^B) \sim_{A'} (x_3^A, x_2^B)$. Since $v^{A'}$ represents $\succsim_{A'}$, it must be true that $v^{A'}(x_1^A, x_2^B) = v^{A'}(x_2^A, x_1^B)$, and $v^{A'}(x_2^A, x_3^B) = v^{A'}(x_3^A, x_2^B)$. Writing out the additive form of the value function yields:

$$w^A v^A\left(x_1^A\right) + w^B v^B\left(x_2^B\right) = w^A v^A(x_2^A) + w^B v^B\left(x_1^B\right) \tag{4}$$

and

$$w^A v^A\left(x_2^A\right) + w^B v^B\left(x_3^B\right) = w^A v^A\left(x_3^A\right) + w^B v^B\left(x_2^B\right). \tag{5}$$

Adding Eqs. 4 and 5, the $w^A v^A(x_2^A)$ terms cancel out, as do the $w^B v^B(x_2^B)$ terms, resulting in:

$$w^A v^A\left(x_1^A\right) + w^B v^B\left(x_3^B\right) = w^A v^A\left(x_3^A\right) + w^B v^B\left(x_1^B\right). \tag{6}$$

Since this is precisely the additive form of $v^{A'}$, it is equivalent to stating that $v^{A'}(x_1^A, x_3^B) = v^{A'}(x_3^A, x_1^B)$, and thus that $(x_1^A, x_3^B) \sim_{A'} (x_1^A, x_3^B)$, which establishes the Thomsen condition, and the converse direction of the proof. □

Proof (Theorem 3) We will first show that statement 1 implies statement 2. Based on Krantz et al. (1971), it will suffice to show that $\succsim_{A'}^*$ induces a weak order on $X \times X$ (i.e., is complete and transitive), that if x_1x_2 and x_2x_3 are positive difference pairs, then x_1x_3 must be a positive difference pair such that $x_1x_3 \succsim_{A'}^* x_1x_2$ and $x_1x_3 \succsim_{A'}^* x_2x_3$, and that $\succsim_{A'}^*$ is weak monotonic, solvable, and satisfies the Archimedean property. Solvability is equivalent to the cardinal indifference property. The remaining properties are established as follows. Throughout the proof, let $x_1, \ldots, x_9$ represent outcomes in X.

To show completeness, we must demonstrate that either $x_1x_2 \succsim_{A'}^* x_3x_4$ or $x_3x_4 \succsim_{A'}^* x_1x_2$ (or both). Because $\succsim_{A'}^*$ must be consistent with $\succsim_{A'}$, if $x_1 \succsim_{A'} x_3$ and $x_4 \succsim_{A'} x_2$, then $x_1x_2 \succsim_{A'}^* x_3x_4$, and if $x_3 \succsim_{A'} x_1$ and $x_2 \succsim_{A'} x_4$, then $x_3x_4 \succsim_{A'}^* x_1x_2$. There are two remaining cases which can be handled similarly. Arbitrarily, let $x_1 \succ_{A'} x_3$ and $x_2 \succ_{A'} x_4$; the identical logic applies in the other case, where these comparisons are reversed. Consistency with $\succsim_{A'}$ implies that $x_1x_4 \succsim_{A'}^* x_1x_2$, and $x_1x_4 \succsim_{A'}^* x_3x_4$. By the cardinal indifference property, there exists x_5 such that $x_1x_4 \succsim_{A'}^* x_1x_5$ and $x_1x_5 \sim_{A'}^* x_1x_2$, and there exists x_6 such that $x_1x_4 \succsim_{A'}^* x_1x_6$ and $x_1x_6 \sim_{A'}^* x_3x_4$. By completeness of $\succsim_{A'}$ (established in the proof of Theorem 1), either $x_6 \succsim_{A'} x_5$ or $x_5 \succsim_{A'} x_6$. If $x_6 \succsim_{A'} x_5$, then $x_1x_5 \succsim_{A'}^* x_1x_6$, and applying the cardinal substitution property twice yields $x_1x_2 \succsim_{A'}^* x_3x_4$. If $x_5 \succsim_{A'} x_6$, then $x_1x_6 \succsim_{A'}^* x_1x_5$, and applying the cardinal substitution property twice yields $x_3x_4 \succsim_{A'}^* x_1x_2$. Hence, $\succsim_{A'}^*$ is complete.

To show transitivity, let $x_1x_2 \succsim_{A'}^* x_3x_4$ and $x_3x_4 \succsim_{A'}^* x_5x_6$. By the cardinal indifference property, there must exist x_7 such that $x_1x_2 \succsim_{A'}^* x_1x_7$, and $x_1x_7 \sim_{A'}^* x_3x_4$. Similarly, there must exist x_8 such that $x_3x_4 \succsim_{A'}^* x_3x_8$, and $x_3x_8 \sim_{A'}^* x_5x_6$. By two applications of the cardinal substitution property, $x_1x_7 \succsim^*_{A'} x_5x_6$. Applying the cardinal indifference property again, there must exist x_9 such that $x_1x_7 \succsim^*_{A'} x_1x_9$, and $x_1x_9 \sim^*_{A'} x_5x_6$. We now have $x_1x_2 \succsim^*_{A'} x_1x_7$, and $x_1x_7 \succsim^*_{A'} x_1x_9$. By transitivity of $\succsim_{A'}$ (established in the proof of Theorem 1) and consistency with $\succsim_{A'}$, $x_1x_2 \succsim^*_{A'} x_1x_9$, and finally, by the cardinal substitution property, $x_1x_2 \succsim^*_{A'} x_5x_6$, establishing transitivity of $\succsim^*_{A'}$.

It is trivial to show that if x_1x_2 and x_2x_3 are positive difference pairs, then x_1x_3 must be as well, since $\succsim^*_{A'}$ is consistent with $\succsim_{A'}$, and it is established in the proof of Theorem 1 that $\succsim_{A'}$ is transitive. Similarly, it is trivial to show based on consistency with $\succsim_{A'}$ that if x_1x_2 and x_2x_3 are positive difference pairs, then $x_1x_3 \succsim^*_{A'} x_1x_2$ and $x_1x_3 \succsim^*_{A'} x_2x_3$.

The Archimedean property asserts that every strictly bounded standard sequence is finite. In this context, a standard sequence on $X \times X$ is a sequence of $x_{i+1}x_i$ such that $x_{i+1}x_i \sim^*_{A'} x_ix_{i-1}$, and not $x_2x_1 \sim^*_{A'} x_1x_1$. An infinite strictly bounded sequence would imply that as $i \to \infty$, there must exist a non-empty set of limiting outcomes X_L such that for all $x_L \in X_L$, $x_L \succ_{A'} x_i$ or equivalently, $x_Lx_i \succ^*_{A'} x_{i+1}x_i$. By the cardinal indifference property, for some arbitrary i, there must exist x'_{iL} for each $x_L \in X_L$ such that $x_Lx_i \succ^*_{A'} x_Lx'_{iL}$ and $x_Lx'_{iL} \sim^*_{A'} x_{i+1}x_i$. By the definition of x_L, there cannot then exist some x_j in the sequence such that $x_j \succ_{A'} x'_{iL}$. However, by the Pareto property of $\succsim_{A'}$, there must be some x_L for which there are x_i infinitesimally close to x_L. This violates the continuity of $\succsim_{A'}$, because $x_L \succ_{A'} x'_{iL}$, and there exist x_j infinitesimally close to x_L. Thus, such an x_L cannot exist, and the Archimedean property holds, which establishes the existence of a continuous cardinal $v^{A'}$ representing $\succsim^*_{A'}$.

The Pareto difference property, and consistency of the preference difference relations with the original preference relations, ensure that because $v^{A'}$ represents $\succsim^*_{A'}$, it must be monotonic.

To establish the converse, assume $v^{A'}$ represents $\succsim^*_{A'}$, and is continuous and monotonically increasing in x^A and x^B. It will suffice to show that $\succsim^*_{A'}$ must satisfy the Pareto difference, cardinal substitution, and cardinal indifference properties. These are straightforward, and follow the same approach used to establish the required conditions in the converse direction of Theorem 1. □

Proof (Theorem 5) We will first show that statement 1 implies statement 2. It will suffice to show that $\succsim^G_{A'}$ satisfies completeness, transitivity, and continuity.

It will be useful to introduce g^0 and g^*, based on a given g_1, g_2, to represent gambles in G that both individuals consider equally preferable to their less preferred and more preferred, respectively, of g_1 and g_2. That is, g^0 is a gamble constructed such that g_1 and g_2 are Pareto superior to it, and g^* is a gamble constructed to be Pareto superior to g_1 and g_2. Intuitively, g^0 combines individual A's less preferable gamble over X^A with individual B's less preferable gamble over X^B, while g^* combines the individuals' more preferable gambles over X^A and X^B. Because X is the product set of $X^A \times X^B$, it is guaranteed that such gambles exist.

It will also be useful to note that, by the independence condition, $g_1 \succsim^G_{A'} g_2$ implies that $(p_1g_1 + (1 - p_1)g_2) \succsim^G_{A'} (p_2g_1 + (1 - p_2)g_2)$ iff $p_1 \geq p_2$. That is, whichever mixture of the two gambles has the higher probability of the preferred gamble must be preferred.

The completeness condition can be shown to hold by considering four possible cases, which are collectively exhaustive by completeness of $\succsim_A^G$ and $\succsim_B^G$. In case 1, $g_1 \succsim_A^G g_2$ and $g_1 \succsim_B^G g_2$, and in case 2, $g_2 \succsim_A^G g_1$ and $g_2 \succsim_B^G g_1$. The Pareto property ensures that $g_1 \succsim_{A'}^G g_2$ in case 1, and $g_2 \succsim_{A'}^G g_1$ in case 2.

In case 3, $g_1 \succ_A^G g_2$ and $g_2 \succ_B^G g_1$. Consider g^0 and g^* based on g_1 and g_2. By the Pareto property, $g^* \succsim_{A'}^G g_1, g_2$ and $g_1, g_2 \succsim_{A'}^G g^0$. Then, by the indifference property, there must exist p_1, p_2 between 0 and 1 such that $(p_1 g^* + (1 - p_1)g^0) \sim_{A'}^G g_1$ and $(p_2 g^* + (1 - p_2)g^0) \sim_{A'}^G g_2$. However, the independence condition implies that $(p_1 g^* + (1 - p_1)g^0) \succsim_{A'}^G (p_2 g^* + (1 - p_2)g^0)$ iff $p_1 \geq p_2$. Since either $p_1 \geq p_2$ or $p_2 \geq p_1$, it follows that either $(p_1 g^* + (1 - p_1)g^0) \succsim_{A'}^G (p_2 g^* + (1 - p_2)g^0)$ or $(p_2 g^* + (1 - p_2)g^0) \succsim_{A'}^G (p_1 g^* + (1 - p_1)g^0)$. Applying the substitution property yields that either $g_1 \succsim_{A'}^G g_2$ or $g_2 \succsim_{A'}^G g_1$. Case 4, where $g_2 \succ_A^G g_1$ and $g_1 \succ_B^G g_2$, can be handled in a similar manner, which establishes completeness.

To establish transitivity, let $g_1 \succsim_{A'}^G g_2$ and $g_2 \succsim_{A'}^G g_3$. If either $g_1 \sim_{A'}^G g_2$ or $g_2 \sim_{A'}^G g_3$, then the substitution property implies that $g_1 \succsim_{A'}^G g_3$. Thus, we need only consider the case where $g_1 \succ_{A'}^G g_2$ and $g_2 \succ_{A'}^G g_3$, which will be shown by contradiction. Assume it is not true that $g_1 \succsim_{A'}^G g_3$, or equivalently, assume that $g_3 \succ_{A'}^G g_1$. Then, by the indifference property, there exists $p \in [0, 1]$ such that $(pg_3 + (1 - p)g_2) \sim_{A'}^G g_1$. If $p = 0$ or $p = 1$, this indifference relationship would contradict one of the initial strict preferences. Thus, we can let $p \in (0, 1)$. The substitution property allows us to state that $(pg_3 + (1 - p)g_2) \succ_{A'}^G g_2$. The independence condition implies that for this to hold, it must be true that $g_3 \succ_{A'}^G g_2$. However, this contradicts our initial premise that $g_2 \succsim_{A'}^G g_3$. Hence, it must be true that $g_1 \succsim_{A'}^G g_3$, which establishes transitivity.

To establish continuity, let $g_1 \succ_{A'}^G g_2$ and $g_2 \succ_{A'}^G g_3$. By the indifference property, there exists $p \in [0, 1]$ such that $(pg_1 + (1 - p)g_3) \sim_{A'}^G g_2$. It is clear that p cannot equal 0 or 1, given the strict preference relationships stated. Consider any $\varepsilon \in (0, min(p, 1 - p))$. By the independence property, the mixture $((1 - \varepsilon)g_1 + \varepsilon g_3) \succ_{A'}^G g_2$, and $g_2 \succ_{A'}^G (\varepsilon g_1 + 1 - (\varepsilon)g_3)$, which establishes continuity.

The converse direction of the theorem follows directly from properties of $u^{A'}(x)$, using an approach analogous to that used for the converse directions of the earlier proofs.

References

Alexander, R. D. (1987). *The biology of moral systems*. New York: Aldine de Gruyter.

Alger, I., & Weibull, J. W. (2010). Kinship, incentives, and evolution. *American Economic Review*, *100*, 1727–1760.

Andreoni, J. (1990). Impure altruism and donations to public goods: A theory of warm-glow giving. *The Economic Journal*, *100*(401), 464–477.

Arrow, K. J. (1951). *Social choice and individual values*. New York: Wiley.

Arrow, K. J. (1975). Gifts and exchanges. In E. S. Phelps (Ed.), *Altruism, morality and economic theory*. New York: Russell Sage Foundation.

Batson, C. D., Batson, J. G., Griffitt, C. A., Barrientos, S., Brandt, J. R., Sprengelmeyer, P., et al. (1989). Negative-state relief and the enpathy-altruism hypothesis. *Journal of Personality and Social Psychology*, *56*(6), 922–933.

Bell, D. E., & Keeney, R. L. (2009). Altruistic utility functions for joint decisions. In *The Mathematics of Preference, Choice, and Order: Essays in Honor of Peter C. Fishburn*. Berlin, Heidelberg: Springer.

Bergstrom, T. C., & Stark, O. (1993). How altruism can prevail in an evolutionary environment. *The American Economic Review*, *83*(2), 149–155.

Cialdini, R. B., Schaller, M., Houlihan, D., Arps, K., Fultz, J., & Beaman, A. L. (1987). Empathy-based helping: Is it selflessly or selfishly motivated? *Journal of Personality and Social Psychology*, *52*(4), 749–758.

Condorcet, M. D. (1785). *Essai sur l'Application de l'Analyse a la Probabilite des Decisions rendues a la Pluralite des Voix*. Paris: Imprimerie Royale.

Debreu, G. (1954). Representation of a preference ordering by a numerical function. In R. M. Thrall, C. H. Coombs, & R. L. D (Eds.), *Decision processes* (pp. 159–165). New York: Wiley.

Debreu, G. (1964). Continuity properties of paretian utility. *International Economic Review*, *5*(3), 285–293.

Dyer, J. S., & Sarin, R. K. (1979). Measurable multiattribute value functions. *Operations Research*, *27*(4), 810–822.

Eisenfuhr, F., Weber, M., & Langer, T. (2010). *Rational decision making*. Berlin: Springer.

Farquhar, P. H. (1984). State of the art utility assessment methods. *Management Science*, *30*(11), 1283–1300.

Fishburn, P. C. (1965). Independence in utility theory with whole product sets. *Operations Research*, *13*(1), 28–45.

Fishburn, P. C. (1970). *Utility theory for decision making*. New York: Wiley.

Fleming, M. (1952). A cardinal concept of welfare. *The Quarterly Journal of Economics*, *66*(3), 366–384.

Hamilton, W. D. (1963). The evolution of altruistic behavior. *The American Naturalist*, *97*(896), 354–356.

Harsanyi, J. C. (1955). Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility. *The Journal of Political Economy*, *63*(4), 309–321.

Holman, E. W. (1971). A note on additive conjoint measurement. *Journal of Mathematical Psychology*, *8*, 489–494.

Hori, H., & Kanaya, S. (1989). Utility functionals with nonpaternalistic intergenerational altruism. *Journal of Economic Theory*, *49*(2), 241–265.

Karni, E., & Safra, Z. (1998). The hexagon condition and additive representation for two dimensions: An algebraic approach. *Journal of Mathematical Psychology*, *42*(4), 393–399.

Keeney, R. L., & Raiffa, H. (1976). *Decisions with multiple objectives: Preferences and value tradeoffs*. New York: Wiley.

Kirkwood, C. (1997). *Strategic decision making: Multiobjective decision analysis with spreadsheets*. Belmont: Duxbury Press.

Krantz, D. H., Luce, R. D., Suppes, P., & Tversky, A. (1971). *Foundations of measurement. Additive and polynomial representations* (Vol. 1). New York and London: Academic Press.

McCardle, K. F., Rajaram, K., & Tang, C. S. (2009). A decision analysis tool for evaluating fundraising tiers. *Decision Analysis*, *6*(1), 4–13.

Montero, M. (2008). Altruism, spite and competition in bargaining games. *Theory and Decision*, *65*(2), 125–151.

Nowak, M. A., & Sigmund, K. (2005). Evolution of indirect reciprocity. *Nature*, *437*, 1291–1298.

Phelps, E., & Pollak, R. A. (1968). On second-best national saving and game-equilibrium growth. *The Review of Economic Studies*, *35*(2), 185–199.

Rader, T. (1963). The existence of a utility function to represent preferences. *The Review of Economic Studies*, *30*(3), 229–232.

Ray, D. (1987). Nonpaternalistic intergenerational altruism. *Journal of Economic Theory*, *41*, 112–132.

Reyniers, D., & Bhalla, R. (2013). Reluctant altruism and peer pressure in charitable giving. *Judgment and Decision Making*, *8*(1), 7–15.

Saez-Marti, M., & Weibull, J. W. (2005). Discounting and altruism to future decision makers. *Journal of Economic Theory*, *122*, 254–266.

Sen, A. K. (1970). The impossibility of a paretian liberal. *Journal of Political Economy*, *78*(1), 152–157.

Shubik, M. (1982). *Game theory in the social sciences*. Cambridge: MIT Press.

Simon, H. A. (1990). A mechanism for social selection and successful altruism. *Science*, *250*(4988), 1665–1668.

Simon, J., Kirkwood, C. W., & Keller, L. R. (2014). Decision analysis with geographically varying outcomes: Preference models and illustrative applications. *Operations Research*, *62*(1), 182–194.

Trivers, R. L. (1971). The evolution of reciprocal altruism. *The Quarterly Review of Biology*, *46*(1), 35–57.

von Neumann, J., & Morgenstern, O. (1947). *Theory of games and economic behavior (2nd rev.)*. Princeton: Princeton University Press.

von Winterfeldt, D., & Edwards, W. (1986). *Decision analysis and behavioral research*. New York: Cambridge University Press.

Vostroknutov, A. (2013). Preferences over consumption and status. *Theory and Decision*, *74*(4), 509–537.