Theoretical Economics 13 (2018), 579–606 1555-7561/20180579

# The transitive core: Inference of welfare from nontransitive preference relations

Hiroki Nishimura
Department of Economics, University of California, Riverside

In this paper, we study methods of inferring a decision maker's true preference relation when observed choice data reveal a nontransitive preference relation due to choice mistakes. We propose some sensible properties of such methods and show that these properties characterize a unique rule of inference, called the transitive core. This rule is applied to a variety of nontransitive preference models, such as semiorders on the commodity space, relative discounting time preferences, justifiable preferences over ambiguous acts, regret preferences over risky prospects, and collective preferences induced by majority voting. We show that the transitive core offers a nontrivial and reasonable inference of the decision maker's true preference relation in these contexts.

Keywords. Nontransitive preferences, welfare, bounded rationality.
JEL classification. D11, D60.

## 1. Introduction

Many behavioral models postulate boundedly rational or heuristic choice procedures rather than the standard utility maximization. For example, satisficing, limited cognition, short-list methods, and framing effects are behavioral choice models that attract significant attention in the literature. (See, for instance, Bernheim and Rangel 2007, Caplin et al. 2011, Manzini and Mariotti 2007, Masatlioglu et al. 2012, Rubinstein and Salant 2006, Salant 2011, Salant and Rubinstein 2008, Tyson 2008 for recent developments.)

However, when a decision maker follows a behavioral choice procedure, it is unclear how we can infer the values of alternatives for the decision maker. In this paper, we consider a decision maker endowed with a complete and transitive preference relation, which is revealed through the observed choice data. If the decision maker is capable

Hiroki Nishimura: hiroki.nishimura@ucr.edu
I am deeply indebted to Efe A. Ok for his continuous support and guidance. I am also grateful to Debraj Ray and David Pearce for their help and encouragement. Last, I thank Simone Cerreia-Vioglio, Yoram Halevy, Hao Li, Fabio Maccheroni, Michael Peters, Michael Richter, Gil Riella, Ariel Rubinstein, Yuval Salant, Ennio Stacchetti, seminar participants at Bocconi University, Kellogg School of Management at Northwestern University, New York University, Princeton University, University of California Riverside, University of California San Diego, University of California Santa Barbara, University of Southern California, Washington University in St. Louis, Yokohama National University, Bounded Rationality in Choice Workshop, Canadian Economic Association Annual Meeting, Canadian Economic Theory Conference, the Co-editor, and anonymous referees for helpful comments and discussion. The financial supports from C.V. Starr Center at New York University and from University of California Riverside are gratefully acknowledged. All errors are my own.

 https://doi.org/10.3982/TE1769

of making consistent choices from each choice problem, then the observed choice data will reveal her "true" preference relation. However, if the decision maker tends to make mistakes when she makes choices, then the revealed preference relation may be distorted and different from her true preference relation. (For example, a revealed indifference between $x$ and $y$ might only mean that she mistakenly chose $y$ over $x$ on some occasions when she actually preferred $x$ to $y$.) This is problematic from both the positive and the normative perspectives. For example, given a set of alternatives, our predictions of the best alternatives in this set may not coincides with those according to her true preference relation. What is even worse is that we may not be able to identify the best alternatives, as the revealed preference relation could fail transitivity. In this paper, to deal with these difficulties, we study methods of inferring the decision maker's true preference relation from observation of a revealed preference relation.

To be more specific, we assume observability of the decision maker's choices from every binary choice problem. Under this assumption, the observed choice data reveal a complete preference relation. But when the decision maker makes choice mistakes, the revealed preference relation may fail transitivity. Therefore, we can model the method of inference as a function that maps every complete, but possibly nontransitive, binary relation to a reflexive and transitive binary relation. This function is referred to as a welfare evaluation rule (WER). Since a WER should make a best guess of the decision maker's true preference relation, we require inferred preference relations to be transitive. However, when the observed choice data are too conflicted, we may not be able to fully recover the decision maker's true preference relation. To accommodate this impossibility of inference, we allow inferred preference relations to be incomplete.

In this paper, we study a series of properties of WERs. For example, suppose that the decision maker chooses an alternative $x$ over another alternative $y$ and, thus, $x$ is revealed to be preferred over $y$. Also, suppose that this preference of $x$ over $y$ does not conflict with the observed choice data, in the sense that the revealed preference relation has no cycle that involves the preference of $x$ over $y$. In this case, it seems reasonable to say that the decision maker did not choose $x$ by mistake, and the revealed preference of $x$ over $y$ reflects the decision maker's true preference. We formulate such criteria for inference as properties of WERs and discuss their plausibility.

We use these properties to refine the class of WERs and to obtain sensible methods of inferring the decision maker's true preference relation. In fact, we show that the proposed properties characterize a unique welfare evaluation rule $\mathrm{c}(\cdot)$ such that

$$x\,\mathrm{c}(\succsim)\,y \quad \text{if and only if} \quad \begin{cases} z \succsim x & \text{implies} \quad z \succsim y \\ y \succsim z & \text{implies} \quad x \succsim z \end{cases} \quad \text{for every } z \in X$$

for any complete preference relation $\succsim$ on a nonempty set $X$. The inferred preference relation $\mathrm{c}(\succsim)$ is transitive even if a revealed preference relation $\succsim$ is not transitive. We refer to the relation $\mathrm{c}(\succsim)$ as the *transitive core* of a revealed preference relation $\succsim$.

The experimental literature provides ample evidence of cyclical choices, and many models of nontransitive preferences have been developed to explain such observations. For example, Kahneman and Tversky (1979) present laboratory choice data on uncertain prospects that exhibit pairwise choice cycles. They argue that the data are well explained by a nonexpected utility representation that entails cyclical evaluations of prospects. In the same context, Loomes and Sugden (1982) develop a model of nontransitive preferences that accounts for the experience of regret. Their model accommodates choice anomalies such as the certainty effect, the common ratio effect, and the common consequence effect (also known as the Allais paradox). For intertemporal choice problems, Read (2001) and Roelofsma and Read (2000) provide experimental data that show that subjects make consistent cyclical choices over intertemporal outcomes. Following this observation, Ok and Masatlioglu (2007), Read (2001), and Rubinstein (2003) study alternative discounting models that induce nontransitive intertemporal preferences. In the theory of consumer preferences, Armstrong (1939, 1948, 1950, 1951) argues that a decision maker exhibits nontransitive indifference when alternatives are too similar to be discerned. This intuitive argument led to the introduction of semiorders (Luce 1956) and interval orders (Fishburn 1970). Each of these works motivates the central question of this paper, and, in turn, we study the implications of the transitive core for these models (see Section 4).

This paper is not the first to study methods of inferring the decision maker's preferences, or welfare, when observed choice data are not rationalizable. Bernheim and Rangel (2007, 2009) study a choice environment in which a decision maker's choice is affected by ancillary conditions. They develop a method of inferring an unambiguous welfare improvement relation over alternatives.[1] Rubinstein and Salant (2012) study an individual decision maker whose behavior is affected by choice frames. They assume a set of preference relations as observable data (each of which accounts for the decision maker's behavior under some frame) and seek an unobservable welfare order that underlies her behavior. Kőszegi and Rabin (2007) discuss a method to identify choice observations that are made by the decision maker's mistakes. Using an example of the gambler's fallacy, they argue that understanding the decision maker's mistakes helps us analyze her welfare. Chambers and Hayashi (2012) examine welfare criteria that apply to random choice data. With the formulation of welfare inference rules as mappings from random choice data into weak orders, they study the implications of normative properties for such mappings. However, to the best of my knowledge, there is no systematic study in the literature that examines and axiomatizes the methods of welfare inference from complete but nontransitive preference relations.

Last, we note that the definition of the transitive core is similar to *the covering order* studied by Fishburn (1977) and Miller (1980) in social choice theory. For any preference relation $\succsim$ on a set $X$, we say that $x$ *covers* $y$ if $y \succsim z$ implies $x \succsim z$ for all $z$ in $X$. By comparing this with the above definition of the transitive core, we can easily observe the similarity between these two rules. However, the seemingly minor difference between the two rules yields quite different implications. Within the context of social

[1] In fact, we can find their unambiguous welfare improvement order from observed choice data assumed in this paper. However, when we observe that choice data consist of the decision maker's choices from only binary choice problems, the unambiguous welfare improvement order is not guaranteed to be transitive.

choice theory, Fishburn (1970) shows that the covering order satisfies the exclusive Condorcet principle, while the transitive core does not. (We verify this in Section 4.5.) Outside the context of social choice theory, the covering order does not obtain results similar to Proposition 2, Theorem 3, and Proposition 4.

The reminder of the paper is structured as follows. Section 2 introduces the notation and terminologies used throughout the paper. In Section 3, we formally define welfare evaluation rules and discuss their properties. Then we characterize the transitive core as a unique welfare evaluation rule that satisfies all the proposed properties. Section 4 provides case studies of the transitive core. These include its applications to semiorders, time preferences, preferences over ambiguous acts, preferences over risky prospects, and collective preferences. Section 5 concludes the paper. All proofs and supplementary results are given in the Appendix.

## 2. Preliminaries

For any set $X$, *a binary relation* on $X$ is a subset of $X \times X$ with a generic notation $\succsim$. As usual, we write $x \succsim y$ to mean that $(x, y) \in \succsim$. The notations $\succ$ and $\sim$ represent the strict part and the symmetric part, respectively, of $\succsim$. A binary relation $\succsim$ on $X$ is *reflexive* if $x \succsim x$ for all $x \in X$, is *complete* if either $x \succsim y$ or $y \succsim x$ for any $x, y \in X$, is *transitive* if $x \succsim y$ and $y \succsim z$ imply that $x \succsim z$ for any $x, y, z \in X$, and is *antisymmetric* if $x \succsim y$ and $y \succsim x$ imply that $x = y$ for any $x, y \in X$. A reflexive and transitive binary relation on $X$ is called *a preorder* on $X$. A preorder on $X$ is called *a weak order* if it is complete, *a partial order* if it is antisymmetric, and *a linear order* if it is complete and antisymmetric. The *diagonal order* on $X$ is the trivial partial order $\Delta_X = \{(x, x) : x \in X\}$. Throughout the paper, preference relations on finite sets are depicted by graphs, as in Figures 1 and 2. In these figures, vertices represent alternatives, and directed arrows are depicted from strictly preferred alternatives to less preferred alternatives. Undirected lines between pairs of alternatives represent indifference of the corresponding pairs.

For any preference relation $\succsim$ on a set $X$, *a cycle* of $\succsim$ is a finite sequence $(z_l)_{l=1}^k$ of distinct points in $X$ such that $z_1 \succsim z_2 \succsim \cdots \succsim z_k \succsim z_1$ with at least one strict preference. An ordered pair $(x, y)$ is *involved* in a cycle $(z_l)_{l=1}^k$ of $\succsim$ if $z_l = x$ and $z_{l+1} = y$ for some $l < k$ or if $z_k = x$ and $z_1 = y$. A preference relation $\succsim$ is called *cyclic* if it has at least one cycle and is *acyclic* otherwise. We identify two cycles of a preference relation if they are identical upon rotation. For example, a preference relation in Figure 1 has three cycles, $(y, z, w)$, $(z, w, y)$, and $(w, y, z)$, but they are viewed as the same cycle.

Let $\succsim$ be any preference relation on a set $X$ and let $\pi$ be a permutation on $X$. With abuse of notation, we denote by $\pi(\succsim)$ the binary relation $\{(\pi(x), \pi(y)) : x \succsim y\}$. The *inverse* of a preference relation $\succsim$ is the binary relation $\text{inv}(\succsim) = \{(y, x) : x \succsim y\}$. Given a subset $S$ of $X$, the *restriction* of $\succsim$ on $S$ is the binary relation $\succsim_S = \{(x, y) \in S \times S : x \succsim y\} \cup \Delta_X$. Note that the restriction $\succsim_S$ makes no comparison for distinct alternatives in $X \setminus S$, but it is reflexive on $X$. When $S$ is a finite set, say $S = \{x, y, z\}$, we write $\succsim_{xyz}$ instead of $\succsim_{\{x,y,z\}}$ for brevity.

## 3. The transitive core

Let $X$ be the set of all conceivable alternatives of interest with $|X| > 3$. We assume that there is a decision maker who has a complete and transitive preference relation on $X$, which is revealed through the observed choice data. In this paper, we allow for the possibility that the observed choice data may not be rich enough to reveal the decision maker's preference relation on the entire $X$. Instead, we assume that there is a subset $S$ of $X$ such that we observe the decision maker's choices from each binary choice problem of alternatives in $S$.[2,3] We can interpret the set $S$ as a technical restriction on the data set. For example, $S$ may consist of only goods available today, while $X$ is a set of goods available today or at some time $t > 0$ in the future.

If the decision maker is capable of making consistent choices from each choice problem, then the observed choice data will fully reveal her "true" preference relation on the set $S$. However, if the decision maker tends to make mistakes when she makes choices, then the revealed preference relation may be distorted and different from her true preference relation. (For example, a revealed indifference between $x$ and $y$ might only mean that she mistakenly chose $y$ over $x$ on some occasions when she actually preferred $x$ to $y$.) This is problematic from both the positive and normative perspectives since, given a set of alternatives, our predictions of the best alternatives in this set may not coincide with those according to her true preference relation. What is even worse is that we may not be able to identify the best alternatives, as the revealed preference relation could fail transitivity.

In this paper, we study methods of inferring the decision maker's true preference relation from a revealed preference relation. To this end, we model these methods as a function that maps each revealed preference relation to a best guess of a true preference relation. Let $\mathcal{P}$ be the set of all revealed preference relations that we may obtain from the observed choice data. Then, due to our observability assumption, each $\succsim \in \mathcal{P}$ is complete on some subset $S$ of $X$. However, if the decision maker makes choice errors, then the revealed preference relation may fail transitivity. Therefore, we define

$$\mathcal{P} = \{\succsim : \succsim = \succsim_S \text{ is complete on some subset } S \subseteq X\}$$

without requiring transitivity of preference relations in this set.

We then define *a welfare evaluation rule* (WER) as a mapping $\sigma$ on $\mathcal{P}$ that associates each $\succsim \in \mathcal{P}$ with a preorder $\sigma(\succsim)$ on $X$. A WER represents a method of inferring the

[2]We may observe the decision maker's choices from the same choice problem on more than one occasion. If she chooses one alternative on some occasions and another on others, then we reveal indifference between these two alternatives.

[3]This assumption does not exclude the possibility that we observe the decision maker's choices from nonbinary choice problems. For example, suppose that the decision maker is endowed with a utility function, but she fails to distinguish small differences in utility when she make choices. In this case, the decision maker's choice behavior is explained by maximization of a semiorder (see Section 4.1). Even though a semiorder is nontransitive, maximization of a semiorder always has solutions under appropriate topological assumptions. In fact, we can reveal this semiorder through the observation of choice data from *non*binary choice sets. An essential assumption of the paper is that the observed choice data are explained by some choice model under a complete but nontransitive preference relation and that this preference relation is revealed through the data.

decision maker's true preference relation from each revealed preference relation. For $\sigma(\succsim)$ as the best guess of the true preference relation (which is transitive), we require an inferred preference relation $\sigma(\succsim)$ to be transitive. However, if the observed data are too conflicted, then we may not be able to fully recover the decision maker's true preference relation. To accommodate this impossibility of inference, we allow an inferred preference relation $\sigma(\succsim)$ to be incomplete.[4]

While WERs represent methods of inferring the decision maker's true preference relation, the concept itself does not offer a reasonable way to do so. In fact, the following examples present two extreme welfare evaluation rules.

Example 1 (Universal incomparability). The universally incomparable WER is a map $\sigma_0$ that assigns the diagonal order $\Delta_X$ for every revealed preference relation in $\mathcal{P}$. This rule only says that the decision maker's true preference is reflexive and makes no further inference. While the rule does not make an incorrect inference, it is likely useless. ◊

Example 2 (Universal indifference). The universally indifferent WER is a map $\sigma_1$ that assigns the trivial weak order $X \times X$ for every preference relation in $\mathcal{P}$. This rule assumes that the decision maker is indifferent between any pair of alternatives, regardless of the observed choice data. Even when the decision maker makes no choice mistakes, and, thus, the observed data reveal her true preference relation, the rule ignores it. The universally indifferent WER is most certainly unattractive. ◊

These examples suggest that the concept of WERs itself does not offer sensible methods for inferring the decision maker's true preference relation. So as to find a useful method, we examine some properties of WERs below.

### 3.1 *Properties of welfare evaluation rules*

In this section, we present six axioms of welfare evaluation rules. These axioms apply for any preference relation $\succsim$ in $\mathcal{P}$, any subsets $S$ and $T$ of $X$, and any permutation $\pi$ on $X$.

Axiom 1 (Prudence). *For any $x, y \in X$, $x\, \sigma(\succsim)\, y$ implies that $x \succsim y$.*

Axiom 2 (Principle of revealed preferences). *If $x \succsim y$ and no cycle of $\succsim$ involves $(x, y)$, then $x\, \sigma(\succsim)\, y$.*

Axiom 3 (Neutrality). *We have $\sigma \circ \pi(\succsim) = \pi \circ \sigma(\succsim)$.*

Axiom 4 (Inverse). *We have $\sigma \circ \mathrm{inv}(\succsim) = \mathrm{inv} \circ \sigma(\succsim)$.*

Axiom 5 (Downward consistency). *If $x\, \sigma(\succsim_T)\, y$ and $x, y \in S \subseteq T$, then $x\, \sigma(\succsim_S)\, y$.*

Axiom 6 (Upward consistency). *If $x\, \sigma(\succsim_{xyz})\, y$ for all $z \in X$, then $x\, \sigma(\succsim)\, y$.*

[4]As any preorder is identified as an intersection of complete and transitive binary relations, we could interpret $\sigma(\succsim)$ as the set of all possible candidates for the decision maker's true preference relation.

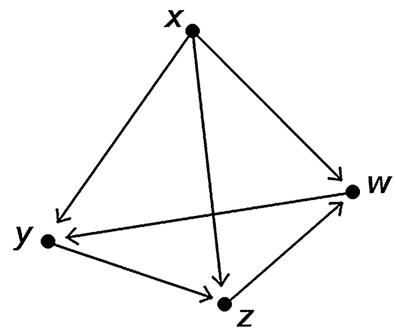


Figure 1. A dominating alternative.

The first axiom, *prudence*, says that a welfare evaluation rule should draw an inference about the decision maker's true preference only when the observed choice data at least reveal the same preference. In other words, if the decision maker never chooses $x$ over $y$ and, therefore, $x$ is not revealed to be preferred to $y$, then a WER should not infer that the decision maker actually prefers $x$ to $y$.

In revealed preference theory, we assume that the observed choice data should, in principle, reveal the decision maker's true preference relation.[5] Axiom 2 requires WERs to respect the same principle whenever we do not observe inconsistency in the choice data. For example, consider a revealed preference relation $\succsim$, depicted in Figure 1. While this preference relation is cyclic, the decision maker consistently chooses an alternative $x$ whenever it is available. Consequently, the revealed preference of $x$ over the other alternatives is not involved in any preference cycles. Under Axiom 2, WERs infer the decision maker's true preference of $x$ over the other alternatives in this case.

Note that Axioms 1 and 2 imply that $\sigma(\succsim) = \succsim$ when a revealed preference relation $\succsim$ is transitive. Indeed, if the decision maker never makes choice mistakes, then the observed choice data reveal the decision maker's true preference relation, and, hence, the revealed preference relation is transitive. Identifying the revealed preference relation with the true preference relation when the decision maker makes no choice mistakes is a reasonable property of welfare evaluation rules implied by the two axioms.

The neutrality axiom requires that the inference of true preference relations be independent of labeling of alternatives. Suppose that two preference relations, $\succsim$ and $\succsim'$, on $X$ are identical upon relabeling of alternatives. (Formally, this implies the existence of a permutation $\pi$ on $X$ such that $x \succsim y$ if and only if $\pi(x) \succsim' \pi(y)$ for any $x$ and $y$ in $X$. Figure 2 provides an example of such preference relations.) If a WER $\sigma$ is independent of labeling, then the true preference relations $\sigma(\succsim)$ and $\sigma(\succsim')$ inferred from these revealed preference relations must be identical upon the same relabeling. Axiom 3 requires this by imposing that $x\, \sigma(\succsim)\, y$ if and only if $\pi(x)\, \sigma(\succsim')\, \pi(y)$ for all $x$ and $y$ in $X$.

[5]Though this is the standard assumption in welfare economics, it is not adopted in contexts in which the decision maker's choice is affected by habits and/or temptations. See Gul and Pesendorfer (2001) and Rozen (2010) for examples.

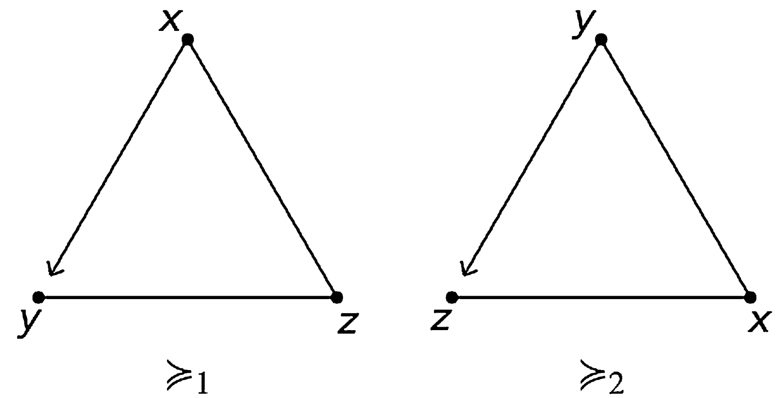


FIGURE 2. Identical relations upon relabeling.

The inverse axiom states that if two revealed preference relations are inverse to each other, then the true preference relations inferred from these revealed preference relations should be inverse to each other too. The inverse relation $\mathrm{inv}(\succsim)$ of a preference relation $\succsim$ is such that $x \succsim y$ if and only if $y \ \mathrm{inv}(\succsim)\ x$ for any $x$ and $y$. Axiom 4 requires that $\sigma(\mathrm{inv}(\succsim))$ be inverse of $\sigma(\succsim)$, that is, $\sigma(\mathrm{inv}(\succsim)) = \mathrm{inv}(\sigma(\succsim))$.

The downward consistency axiom imposes consistency of inference of the true preference relation from two sets of choice data, where one set of data is a superset of the other. For simplicity of exposition, suppose that there are comprehensive choice data that reveal a complete preference relation $\succsim$ on $X$. Suppose, also, that $x \succsim y$ for some $x$ and $y$ in $X$. The decision maker may make some choice mistakes, and, therefore, there might exist some preference cycles that involve $(x, y)$. Note that the observation of such inconsistency with the preference $x \succsim y$ reduces when our observation of choice data is limited. If we observe only the choice data for alternatives in a set $S \subseteq X$, then these limited choice data will reveal a preference relation $\succsim_S$. Any preference cycle of $\succsim_S$ that involves $(x, y)$ is a preference cycle of $\succsim$, but not vice versa. (The extreme case is $S = \{x, y\}$, in which case we have no means to observe inconsistency with the preference $x \succsim y$.) Axiom 5 requires that if a WER infers the true preference of $x$ over $y$ when we observe more inconsistencies with this preference, then the rule must infer the same when we observe less.

The upward consistency axiom is also a consistency requirement on the inference from two sets of choice data. It claims that if a WER infers the true preference of $x$ over $y$ whenever we observe choice data on three alternatives $\{x, y, z\}$ for some $z \in X$, then the rule should draw the same inference when we observe the comprehensive choice data.

To illustrate the implications of this axiom, let $X = \{1, 2, \dots, N\}$ for some $N > 3$, and suppose that the comprehensive choice data would reveal a complete preference relation $\succsim$ on $X$ such that

$$n \succsim m \quad \text{if and only if} \quad \left[n \geq m \text{ and } (n, m) \neq (N, 1)\right] \quad \text{or} \quad (n, m) = (1, N) \tag{1}$$

for any $n, m \in X$. The preference relation $\succsim$ is identical to the Euclidean order $\geq$, except that $1 \succ N$. Notice that *every* preference of this relation is involved in some preference cycles. For example, a revealed preference of 3 over 2 ($3 \succ 2$) is involved in a cycle $(3, 2, 1, N)$. So, Axiom 2 offers no implication for this preference relation $\succsim$. However, our intuition may suggest that the preference $1 \succ N$ is the main cause of preference cycles. The preference relation $\succsim$ is, after all, "almost" identical to the transitive order $\geq$, except for the preference of 1 over $N$, which is perhaps due to a choice mistake. Therefore, we may wish to infer that the revealed preference $3 \succ 2$, for example, should identify the decision maker's true preference.

Note that we can observe inconsistency of the revealed preference relation $\succsim$ only when we observe the decision maker's choice from $\{1, N\}$. If the observed choice data are restricted on $\{x, y, z\}$ for any $x, y, z \in X$ with $N - 1 \geq x, y \geq 2$, then a revealed preference relation from such data is always transitive. We can show that if a WER $\sigma$ satisfies Axiom 6 as well as Axiom 2, then the rule must infer the true preference of $x$ over $y$ for any $x$ and $y$ with $N - 1 \geq x \geq y \geq 2$. The upward consistency axiom allows WERs to make nontrivial implications by identifying "less conflicting" revealed preferences.

In the Appendix, we prove that the proposed axioms are mutually independent by presenting welfare evaluation rules that satisfy all except one axiom. For example, the universally incomparable WER $\sigma_0$ admits all but Axiom 2, and the universally indifferent WER $\sigma_1$ admits all but Axiom 1. Note that even these extreme welfare evaluation rules satisfy almost all the axioms.

### 3.2 *The transitive core*

In this section, we introduce a certain WER called the transitive core. The transitive core checks consistency of a revealed preference relation using an arbitrary reference point. For example, consider a revealed preference relation $\succsim$ on $X$ such that $x \succsim y$ for some alternatives $x$ and $y$. The transitive core views this revealed preference $x \succsim y$ to be conflicting with a third alternative $z \in X$ if either $y \succsim z \succ x$ or $y \succ z \succsim x$. The transitive core implies that the revealed preference $x \succsim y$ identifies the decision maker's true preference if no third alternative $z \in X$ conflicts with the preference $x \succsim y$.

Formally, the *transitive core* is a welfare evaluation rule $\mathrm{c}(\cdot)$ such that

$$x \,\mathrm{c}(\succsim)\, y \quad \text{if and only if} \quad \begin{cases} z \succsim x & \text{implies} \quad z \succsim y \\ y \succsim z & \text{implies} \quad x \succsim z \end{cases} \quad \text{for every } z \in X \tag{2}$$

for any preference relation $\succsim$ in $\mathcal{P}$ and any $x$ and $y$ in $X$. It is straightforward to show that the transitive core maps every preference relation in $\mathcal{P}$ to a preorder on $X$.

The more important observation regarding the transitive core is that it satisfies all of the axioms discussed above. Therefore, if we adopt the axioms as desired properties of WERs, then the transitive core is a possible candidate for methods of inferring the decision maker's true preference relation. In fact, we can show that the rule offers a reasonable inference of true preferences in many contexts. The next example identifies the transitive core of a revealed preference relation defined by (1). In Section 4, we study implications of the transitive core for other models of nontransitive preference relations.

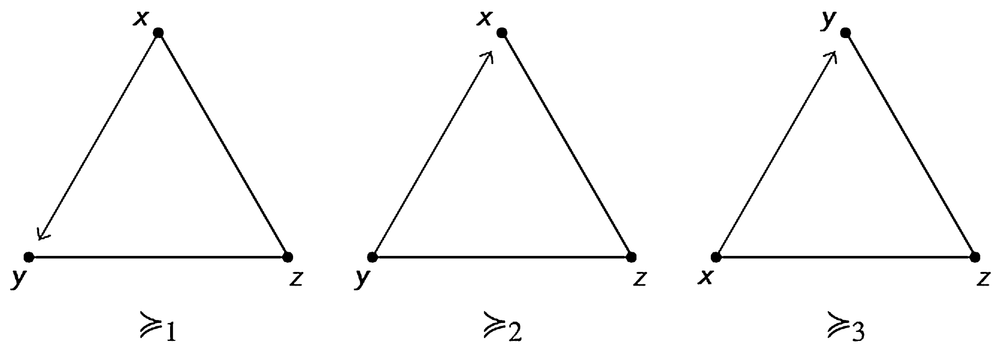


FIGURE 3. Illustration of Theorem 1.

EXAMPLE 3. Recall the preference relation $\succsim$ on $X = \{1, 2, \dots, N\}$ defined by (1). As we observed earlier, if $N - 1 \geq x \geq y \geq 2$, then the preference relation is transitive when it is restricted to three alternatives $\{x, y, z\}$ for any $z \in X$. This means that no $z$ satisfies $y \succsim z \succ x$ or $y \succ z \succsim x$, and, therefore, no third alternative conflicts with the preference $x \succsim y$. Thus, the transitive core infers that the revealed preference $x \succsim y$ identifies the decision maker's true preference. In contrast, the preference $1 \succ N$ conflicts with some (in fact, all) third alternatives. So, the rule drops this revealed preference from the decision maker's true preference relation. After all, the transitive core infers the true preference relation $\mathrm{c}(\succsim)$ such that

$$x \,\mathrm{c}(\succsim)\, y \quad \text{if and only if} \quad N - 1 \geq x \geq y \geq 2$$

for any $x, y$ in $X$. Therefore, while every revealed preference belongs to some preference cycles, the transitive core drops only the most conflicting preferences. Notice that this example also verifies the difference between the transitive core and the removal of preference cycles. ◊

Below, the main result of this paper shows that the transitive core is not only a WER that satisfies Axioms 1–6, but also a unique WER that does so. Therefore, these axioms characterize the transitive core.

THEOREM 1. *A WER satisfies Axioms 1–6 if and only if it is the transitive core.*

The theorem implies that if a WER satisfies all axioms, then we can uniquely identify the true preference relation inferred by the rule for each revealed preference relation. Indeed, it is worthwhile to take a look at how the axioms pin down the true preference relations for some examples. We already verified above that Axioms 1 and 2 imply that $\sigma(\succsim) = \succsim$ for all transitive preference relations $\succsim$ in $\mathcal{P}$. Below, we consider a revealed preference relation given by Figure 3.

EXAMPLE 4. Let $\sigma$ be a WER that satisfies Axioms 1–6, and suppose that the observed choice data reveal a preference relation $\succsim_1$ in Figure 3. Theorem 1 implies that $\sigma(\succsim_1) =$

$c(\succsim_1)$, which we verify in this example. First, observe that preferences $x \succsim_1 z$ and $z \succsim_1 y$ are not involved in any preference cycle of $\succsim_1$. So, Axiom 2 implies that $x\,\sigma(\succsim_1)\,z$ and $z\,\sigma(\succsim_1)\,y$, which, in turn, imply $x\,\sigma(\succsim_1)\,y$ by transitivity. Now, consider $\succsim_2$ and $\succsim_3$ in Figure 3. The relation $\succsim_2$ is the inverse of the preference relation $\succsim_1$. The relation $\succsim_3$ is identical to $\succsim_2$ upon relabeling of alternatives. Therefore, it follows from Axiom 3 and Axiom 4 that

$$z\,\sigma(\succsim_1)\,x \quad \overset{\text{Inverse}}{\Longleftrightarrow} \quad x\,\sigma(\succsim_2)\,z \quad \overset{\text{Neutrality}}{\Longleftrightarrow} \quad y\,\sigma(\succsim_3)\,z.$$

However, observe that $\succsim_1$ and $\succsim_3$ are the same preference relation. So, if $z\,\sigma(\succsim_1)\,x$, then we must have $y\,\sigma(\succsim_1)\,z$ and, thus, $y\,\sigma(\succsim_1)\,x$ by transitivity. Since the inference $y\,\sigma(\succsim_1)\,x$ would contradict Axiom 1, we cannot have $z\,\sigma(\succsim_1)\,x$. The same argument also proves that $y\,\sigma(\succsim_1)\,z$ does not hold. This concludes the characterization of $\sigma(\succsim_1)$ by $\sigma(\succsim_1) = \{(x, y), (x, z), (z, y)\}$. It is straightforward via (2) to check that the transitive core $c(\succsim_1)$ obtains the same order. $\lozenge$

Before proceeding, we note that the assumption of completeness on the revealed preference relation is crucial for Theorem 1. In general, if we apply the transitive core to incomplete preference relations by the same rule (2), then it does not even satisfy Axiom 2. (For example, the transitive core maps some acyclic preference relations to the diagonal order.) In addition, if we extend the domain of welfare evaluation rules by including incomplete preference relations, then there is no welfare evaluation rule that satisfies all the axioms discussed above. Finding a reliable method of inferring the decision maker's true preference relation when the observed choice data reveals only an incomplete preference relation is an open question.

## 4. Applications

In the previous section, we show that the transitive core is the unique welfare evaluation rule that satisfies Axioms 1–6. While this characterization provides some justification for the rule, it does not necessarily imply that the rule is a useful concept. In this section, we examine implications of the transitive core by applying it to the models of nontransitive preference relations of economic interest.

### 4.1 *Semiorders: Imperfect ability of discrimination*

Nontransitive indifference due to imperfect ability of discrimination has been studied in the literature. Armstrong (1939, 1948, 1950, 1951) poses a question on the assumption of transitive indifference and first introduces a utility model of imperfect discrimination. Luce (1956) brings a notion of semiorders into economics and provides its axiomatic foundation. Subsequently, many generalizations of semiorders, such as interval orders by Fishburn (1970), are developed in search of descriptive models of nontransitive indifference. In this section, we take $X$ as a connected metric space and consider the following representation of semiorders introduced by Luce (1956).

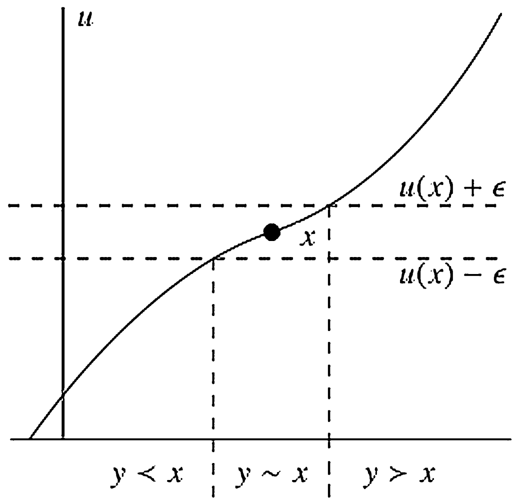


FIGURE 4. A semiorder representation.

DEFINITION 1. A semiorder is a binary relation $\succsim$ on $X$ for which there is a pair $(u, \epsilon)$ of a continuous function $u : X \to \mathbb{R}$ and a nonnegative number $\epsilon \geq 0$ such that

$$x \succsim y \quad \text{if and only if} \quad u(x) \geq u(y) - \epsilon$$

holds for all $x, y \in X$.

Obviously, if $\epsilon = 0$, the representation reduces to the standard utility representation. When $\epsilon > 0$, however, a semiorder entails preference cycles. Note that the representation implies that $x \sim y$ if and only if $|u(x) - u(y)| \leq \epsilon$. So, a semiorder shows an indifference between alternatives $x$ and $y$ even when their utility values are different. Luce refers to the coefficient $\epsilon$ as the "just noticeable difference." Figure 4 illustrates the regions of preferred, indifferent, and less preferred alternatives to a given $x \in X$ when $X$ is a real line.

While a semiorder is cyclic in general, the true preference relation for the decision maker seems obvious in this context. Nontransitivity of the semiorder is induced by an imperfect perception. If it were possible to eliminate the perception error $\epsilon$, the decision maker would consistently evaluate alternatives by the utility function $u$. A ranking induced by the utility function is transitive and represents the decision maker's evaluations of alternatives. It turns out that the transitive core precisely infers this utility order for each semiorder.

PROPOSITION 2. *Let $\succsim$ be a semiorder on $X$ with a representation $(u, \epsilon)$. Then*

$$x \operatorname{c}(\succsim)\, y \quad \textit{if and only if} \quad u(x) \geq u(y)$$

*for any $x, y \in X$, provided that* $\sup |u(x) - u(y)| > 2\epsilon$.[6]

[6]The added condition is a necessary and sufficient condition for uniqueness of the utility order. For example, suppose that $\sup |u(x) - u(y)| < \epsilon$. Then, since the semiorder is indifferent for every pair of alternatives, any function $u' : X \to \mathbb{R}$ with $\sup |u'(x) - u'(y)| < \epsilon$ would represent the same semiorder. If the given condition is satisfied, an ordinal ranking of the underlying utility function is uniquely determined.

Note that the utility function is not directly observable. The proposition implies that the transitive core infers an unobservable utility order from observation of a nontransitive revealed preference relation. Also, the inferred preference relation turns out to be complete in this case.

### 4.2 *Time preferences*

Let $Z$ be a nonempty open interval in $\mathbb{R}_+$, and let $X = Z \times [0, \infty)$. In this section, a generic member $(x, t)$ of $X$ is interpreted as a dated outcome, where the decision maker receives a prize of $x$ dollars at time $t$. Correspondingly, we refer to a complete preference relation on $X$ as a *time preference*. While there are many models of time preferences, the class of *absolute discounting time preferences* attracts particular interest in the literature. These preference relations are represented as

$$(x, t) \succsim (y, s) \quad \text{if and only if} \quad \delta(t)u(x) \geq \delta(s)u(y)$$

for each $(x, t)$, $(y, s)$ in $X$ under some discounting function $\delta : [0, \infty) \to [0, 1]$ and utility function $u : Z \to \mathbb{R}$. The exponential discounting and hyperbolic discounting models are special cases of this class. Note that any absolute discounting time preference is transitive, as it has a utility representation $(x, t) \mapsto \delta(t)u(x)$.

Read (2001) finds that a large fraction of subjects in an experiment appear inconsistent with the model of absolute discounting and are better explained by that of subadditive discounting. The subadditive discounting model postulates that a discount factor for a delay is larger than the product of discount factors for subdelays that partition the original delay. To be more specific, consider the following general representation of time preferences studied by Ok and Masatlioglu (2007).

Definition 2. A time preference $\succsim$ is a relative discounting time preference if there exist continuous functions $u : Z \to \mathbb{R}_{++}$ and $\eta : \mathbb{R}^2_+ \to \mathbb{R}_{++}$ that satisfy

$$(x, t) \succsim (y, s) \quad \text{if and only if} \quad u(x) \geq \eta(s, t)u(y)$$

for each $(x, t)$, $(y, s)$ in $X$, where $u$ is an increasing homeomorphism, $\eta(\cdot, t)$ is decreasing with $\eta(\infty, t) = 0$, and $\eta(t, s) = \eta(s, t)^{-1}$ for any $t, s \geq 0$.

The function $\eta$, called a relative discounting function, measures a relative discount factor for delays between any two points in time. So, the decision maker discounts a utility value of a prize at time $s$ by the factor $\eta(s, t)$ so as to compare it with that of another prize at time $t$. If $\eta(s, t) = \delta(s)/\delta(t)$, the relative discounting model reduces to the absolute discounting model. Given the representation, we can write the subadditive discounting model as a property on $\eta$:

$$\eta(r, t) \geq \eta(r, s)\eta(s, t) \quad \text{for every } r \geq s \geq t. \tag{3}$$

Note that absolute discounting time preferences satisfy (3) with equality. (Indeed, it is a characterization of the absolute discounting model.) Read's finding suggests that allowing (3) to be an inequality provides a significant improvement in the fit of the data.

For another alternative to the absolute discounting model, Rubinstein (1980) proposes *a similarity-based time preference*. This model postulates that the decision maker follows up to three steps of heuristic procedures so as to compare a pair of dated outcomes. In the first step, the decision maker looks for dominance of dated outcomes. So, if $x > y$ and $t < s$, then the dated outcome $(x, t)$ is preferred to $(y, s)$. Provided that there is no dominance between them, the decision maker next looks for similarity in the dated outcomes either in delivery dates or in prizes. For example, if the delivery dates are similar but the prizes are not, the decision maker chooses one that offers the larger prize. Last, if neither of the first two steps resolves her choice, another criterion is applied to compare $(x, t)$ and $(y, s)$. The author argues that this model explains observed data from an experiment better than the hyperbolic discounting model and is more intuitive as a description of the decision maker's reasoning process. Ok and Masatlioglu (2007) prove that a variety of similarity-based time preferences are, in fact, relative discounting time preferences.

We can easily see that some relative discounting time preferences are not transitive. For example, consider a preference relation $\succsim$ represented under a relative discounting function $\eta$ such that

$$\eta(s,t) = \begin{cases} \delta^{s-t} & \text{if } T \geq s-t \geq 0, \\ \delta^{T} & \text{if } s-t > T \end{cases}$$

for any $s$ and $t$, where $T > 0$ is some constant. The decision maker exponentially discounts a delay if it is shorter than $T$. However, she perceives any delays longer than $T$ as the same and applies a constant discount factor for them. (This is an example of subadditive discounting. Note that $\eta$ satisfies (3).) Then, whenever $x$, $y$, and $z$ are prizes such that $(x, 0) \sim (y, T) \sim (z, 2T)$, we have $(z, 2T) \succ (x, 0)$.

This example of a nontransitive time preference is not a special case. Indeed, a relative discounting time preference turns out to be transitive if and only if it is an absolute discounting time preference (Ok and Masatlioglu 2007, Corollary 1). Therefore, any empirical implications of the relative discounting model beyond that of the absolute discounting model are attributed to the nontransitivity of time preferences. For better explanations of the observed data offered by the subadditive discounting model or the similarity-based time preferences, allowing nontransitivity of preference relations is essential.

While the relative discounting model generalizes the class of time preferences to the extent that it contains cyclic relations, we can show that the decision maker's true preference relation inferred by the transitive core has a representation under absolute discounting.

Theorem 3. *Let $\succsim$ be a relative discounting time preference with a representation $(u, \eta)$. Then there is a set $\mathcal{D}$ of continuous functions $\delta : \mathbb{R}_+ \to \mathbb{R}_{++}$ such that*

$$(x,t)\, \mathrm{c}(\succsim)\, (y,s) \quad \textit{if and only if} \quad \delta(t)u(x) \geq \delta(s)u(y) \quad \textit{for all } \delta \in \mathcal{D}$$

*whenever $(x, t)$ and $(y, s)$ are dated outcomes in $X$.*

The theorem shows that the transitive core of a relative discounting time preference is represented by *multiple* absolute discounting functions. In the proof of the theorem, we show that the collection $\mathcal{D}$ consists of functions of the form $\eta(\,\cdot\,, r)$ for arbitrary fixed points $r \geq 0$ in time. Therefore, the transitive core infers the true preference of a dated outcome $(x, t)$ over another $(y, s)$ if and only if the former has a higher discounted utility value than the latter, regardless of the time at which they are evaluated.

### 4.3 *Justifiable preferences*

Let $\Omega$ be a finite nonempty set of states of the world, and let $Y = \Delta(\mathbb{R})$ be the set of all lotteries over real prizes. An arbitrary function $f : \Omega \to Y$ mapping each state to a lottery is referred to as an Anscombe and Aumann (1963) *act*. We denote the collection of all acts by $X$. An act is a description of state-contingent prize schedules, where the objective likelihood of each state is not known.[7]

Consider a group of agents, each of whom has a subjective belief about the likelihood of the states. To study justifiable collective decision making in this context, Lehrer and Teper (2011) propose a model in which the group prefers one act over another if and only if at least one agent in the group has higher expected utility from the former act than the latter. We say that a preference relation $\succsim$ on $X$ is *justifiable* if there exist a continuous affine monotone utility function $u : Y \to \mathbb{R}$ and a nonempty closed convex set $P$ of probability distributions over $\Omega$ such that

$$f \succsim g \quad \text{if and only if} \quad \exists p \in P \quad \text{such that} \quad \sum_{\omega\in\Omega} p(\omega)u\big(f(\omega)\big) \geq \sum_{\omega\in\Omega} p(\omega)u\big(g(\omega)\big)$$

for any two acts $f$ and $g$.[8]

While the agents share the same utility function $u$ over lotteries, they can disagree on evaluations of acts due to different subjective beliefs about the states of the world. In particular, it is possible that, for three acts $f, g, h \in X$, some agents prefer $f$ over $g$ and $g$ over $h$, whereas the others prefer $h$ over $f$ and $f$ over $g$. The justifiable preference of this group is such that $f \sim h \sim g$ and $f \succ g$. Therefore, a justifiable preference relation is not transitive in general (Lehrer and Teper 2011, p. 763).

Lehrer and Teper (2011) contrast justifiable preferences with the Knightian preferences of Bewley (2002). Under Knightian preferences, the group prefers one act over another if and only if *all* agents in the group have higher expected utility from the former act than the latter. Formally, a preference relation $\succsim$ on $X$ is called *Knightian* if there exist a continuous affine monotone utility function $u : Y \to \mathbb{R}$ and a nonempty closed

[7]Throughout this section, a Borel probability measure on $\mathbb{R}$ is called a lottery. We endow the set $Y$ of lotteries with the weak topology and the set $X$ with the product topology. A function $u : Y \to \mathbb{R}$ is said to be monotone if $u(\mu) > u(\nu)$ whenever $\mu$ first-order stochastically dominates $\nu$.

[8]The model of justifiable preferences studied in this section differs from the original work by Lehrer and Teper on two points. First, we take the set $\mathbb{R}$ of real prizes for the outcome space of lotteries. Second, the utility function is assumed to be monotone. For the result in this section, we can weaken these assumptions as long as the utility function $u$ remains locally nonsatiable.

convex set $P$ of probability distributions over $\Omega$ such that

$$f \succsim g \quad \text{if and only if} \quad \sum_{\omega \in \Omega} p(\omega) u\big(f(\omega)\big) \geq \sum_{\omega \in \Omega} p(\omega) u\big(g(\omega)\big) \quad \forall p \in P$$

for any two acts $f$ and $g$. It is straightforward to show that a Knightian preference relation is not necessarily complete but is always transitive. Lehrer and Teper (2011) remark that a justifiable preference relation is a completion of a Knightian preference relation: if $\succsim_J$ and $\succsim_K$ are, respectively, justifiable and Knightian preferences associated with the same pair $(u, P)$ of a utility function and a set of subjective beliefs, then $f \succsim_K g$ implies that $f \succsim_J g$.

A justifiable preference relation may have preference cycles, but it offers complete comparisons of acts and never contradicts a Knightian preference relation. A Knightian preference relation provides an incomplete but transitive ranking over acts unanimously supported by the agents. The transitive core associates these two models in an intuitive way.

Proposition 4. *The transitive core of a justifiable preference relation is a Knightian preference relation under the same pair* $(u, P)$ *of a utility function and a set of beliefs.*

If we observe a justifiable preference relation from a group of agents, the welfare order of this group may not be clear due to the presence of preference cycles. The transitive core suggests that an act will improve the group's welfare over another act when all of the agents in the group unanimously prefer the former act to the latter. Also notice that we do not have to know a representing pair $(u, P)$ of an observed preference relation to apply the transitive core. As long as we observe the group's preference relation, the transitive core infers the unanimity ranking by the agents.

### 4.4 *Regret theory*

Let $\{1, \ldots, n\}$ be a finite set of states of the world with $n \geq 3$. Suppose that there is a *nature* that resolves a state according to a probability distribution $p$ such that $p_i > 0$ for every $i \in \{1, \ldots, n\}$ and $\sum_{i=1}^{n} p_i = 1$. A *prospect* is a real valued function on the set of the states, and it is interpreted as a state-contingent prize schedule delivered to a decision maker. Let $X = \mathbb{R}^n$ be the set of all prospects, and we consider preference relations on $X$ in this section.

The main body of economic analysis under uncertainty relies on the expected utility theory developed by von Neumann and Morgenstern (1944). It postulates that prospects are compared by their expected utility values $\mathbb{E}(u \circ x)$ for some real function $u : \mathbb{R} \to \mathbb{R}$. This theory is acknowledged as the model of rational decision making under uncertainty and is justified from the normative perspective.

However, experimental studies find disparities between observed behavior and the predictions of expected utility theory. The celebrated work by Kahneman and Tversky (1979), for example, provides extensive evidence that subjects violate the expected utility hypothesis in consistent ways. Some of these violations are known as the certainty effect, the common consequences effect (or the Allais paradox), and the isolation effect.

| State | $u(x_i)$ | $u(y_i)$ | $u(z_i)$ |
|---|---|---|---|
| 1 | 2 | 0 | 1 |
| 2 | 1 | 2 | 0 |
| 3 | 0 | 1 | 2 |

Table 1. Utility for $i = 1, 2, 3$.

To accommodate the observed violations of expected utility theory, Loomes and Sugden (1982) propose an alternative theory of decision making that reflects the experience of *regret*. The regret theory takes into account that the decision maker may regret or rejoice in the chosen prospect upon realization of a state. Specifically, when a state is resolved, the decision maker may regret (rejoice) if the outcome of the chosen prospect happens to be worse (better) than that of the alternative prospect. The theory postulates that the psychological factor, as such, affects ex ante tastes over prospects by introspection. In this section, we consider the following representation of a regret preference by Loomes and Sugden.

Definition 3. A regret preference is a preference relation on $X$ for which there exist two continuous functions $u : \mathbb{R} \to \mathbb{R}$ and $Q : \mathbb{R} \to \mathbb{R}$ that satisfy

$$x \succsim y \quad \text{if and only if} \quad \sum_{i=1}^{n} p_i Q\big(u(x_i) - u(y_i)\big) \geq 0$$

for every $x, y \in X$, where $u$ is increasing homeomorphism with $u(0) = 0$, and $Q$ is convex and strictly increasing and satisfies $Q(-a) = -Q(a)$ for all $a \geq 0$.

The function $Q$ measures how much the decision maker regrets/rejoices due to the difference between the realized prizes of the two prospects. The model of regret preferences reduces to an expected utility representation when the function $Q$ is linear. Loomes and Sugden show that regret preferences robustly explain the observed choice anomalies under an assumption of strict convexity of the function $Q$.

We can show that a regret preference is not transitive whenever an associated function $Q$ is strictly convex. For example, let $n = 3$ and $p_1 = p_2 = p_3 = 1/3$, and consider three prospects $x$, $y$, $z$ that give utility values as in Table 1. Then

$$\sum_{i=1}^{3} p_i Q\big(u(x_i) - u(y_i)\big) = \frac{1}{3}Q(2) - \frac{1}{3}Q(1) - \frac{1}{3}Q(1) > 0,$$

and, thus, $x \succ y$. The decision maker prefers $x$ to $y$ since the degree to which she rejoices in the choice of $x$ over $y$ at state 1 is more than enough to compensate for her smaller regrets at states 2 and 3. The same argument applies to pairs $(y, z)$ and $(z, x)$ by symmetry, resulting in a preference cycle $x \succ y \succ z \succ x$.

Indeed, this observation regarding nontransitivity of regret preferences holds in general. Bikhchandani and Segal (2011) show that a regret preference is transitive if and only if it admits an expected utility representation. Therefore, any descriptive power of regret theory beyond that of the expected utility theory is inseparable from the presence of preference cycles.

When the observed choice data reveal a regret preference, the decision maker's true preference relation is unclear due to nontransitivity of the revealed preference relation. Then we might find the transitive core useful for inferring the decision maker's true evaluations of alternatives. Indeed, the rule consistently infers the true preference of one prospect over another whenever the former statewise dominates the latter.

Proposition 5. *Let $\succsim$ be a regret preference on $X$. If $x_i \geq y_i$ for every state $i$, then $x\,\mathrm{c}(\succsim)\,y$. If, in addition, $x_i > y_i$ for some state $i$, then $y\,\mathrm{c}(\succsim)\,x$ does not hold.*

This proposition shows that the transitive core infers the true preference for a pair of prospects when one prospect dominates the other. In general, we can show that the transitive core also infers the true preference for two prospects that do not dominate each other. However, a characterization of the transitive core for regret preferences is not known and is left as an open question.

### 4.5 *Majority voting*

Let $n$ be an arbitrary natural number representing the number of voters in a society, and let $X$ be a set of policies to be chosen. In this section, we consider a society as a representative decision maker and examine its preference relation induced by the majority voting rule. Suppose that an individual voter $i$ for each $i \in \{1, \dots, n\}$ evaluates policies according to a linear order $\succsim_i$ on $X$ and has an equal share of votes. It is well known that a social preference induced by the majority criterion fails transitivity in general. Arrow (1951) gives the celebrated impossibility theorem that proves that there is no voting system that satisfies all of certain desirable criteria at the same time. Given the inevitable ambiguity in policy evaluations, social choice theory has been extensively studied in the literature. The purpose of this section is to examine implications of the transitive core for social preferences. (To this end, we interpret preference relations inferred by the transitive core as social welfare rankings of policy alternatives.)

Definition 4. A majority preference is a preference relation $\succsim$ on $X$ defined by

$$x \succsim y \quad \text{if and only if} \quad \big|\{i : x \succsim_i y\}\big| \geq \big|\{i : y \succsim_i x\}\big|$$

for every $x$ and $y$ in $X$.

Note that a majority preference can also be interpreted as a preference relation of an *individual* decision maker over alternatives with $n$ many attributes. For every $i \in \{1, \dots, n\}$, the decision maker has a ranking $\succsim_i$ of the alternatives according to the $i$th attribute. The majority preference arises when she chooses one alternative over another if and only if there are more attributes in which the former ranks higher than the latter. In this case, the transitive core applied to the majority preference infers the individual's true preference relation.

| Ranking | $\succsim_1$ | $\succsim_2$ | $\succsim_3$ |
|---|---|---|---|
| 1 | $x$ | $w$ | $z$ |
| 2 | $y$ | $x$ | $w$ |
| 3 | $z$ | $y$ | $x$ |
| 4 | $w$ | $z$ | $y$ |

TABLE 2. Individual preferences.

A majority preference is cyclic in general. Table 2 gives an example of voters' preference relations when $n = 3$ and $X = \{x, y, z, w\}$. A majority preference of this society has a preference cycle $x \succ z \succ w \succ x$, making the desirability of each policy unclear. Many voting criteria have been proposed to study an optimal choice for the society. Some of them are listed below.

**Pareto criterion**. We say that a policy $x$ is *a Pareto improvement* over a policy $y$ if every voter in the society prefers $x$ over $y$. The Pareto criterion requires that $y$ not be chosen over $x$ when $x$ is unanimously preferred to $y$.

**Condorcet principle**. A *Condorcet winner* is a policy $x \in X$ that beats every other policy under pairwise majority voting. The principle claims that a Condorcet winner is an optimal choice for the society. The Condorcet winner may not exist, but it is always unique if it does exist.

**Smith's principle (Smith 1973)**. Let $S$ and $T$ be a partition of $X$ such that every policy in $S$ beats every policy in $T$ under pairwise majority voting. This principle suggests that an optimal social choice can be made from the set $S$.

**Exclusive Condorcet principle**. The set $\{x \in X : y \succ x \text{ for no } y \in X\}$ is the set of all undominated policies under pairwise majority voting. The exclusive Condorcet principle assumes that the optimal social choice can be found in this set of policies if it is nonempty.

The Pareto criterion and the Condorcet principle are commonly viewed as desirable properties of social choice. Indeed, these criteria are often used as cornerstones for evaluating voting systems. Smith's principle and the exclusive Condorcet principle imply, but are not implied by, the Condorcet principle. With regard to Smith's principle, Fishburn (1977) states, "I find it hard to imagine an argument against Smith's Condorcet principle that would not also be an argument against Condorcet's principle." The social welfare order inferred by the transitive core turns out to satisfy all principles except for the exclusive Condorcet principle. In the proposition below, we denote by $\mathrm{c}^*(\succsim)$ the strict part of the transitive core, so that $x\,\mathrm{c}^*(\succsim)\,y$ means $x\,\mathrm{c}(\succsim)\,y$ and not $y\,\mathrm{c}(\succsim)\,x$ for any $x, y \in X$.

PROPOSITION 6. *The transitive core of a majority preference satisfies following concepts (but not the exclusive Condorcet principle):*

*Pareto criterion. If $x$ is a Pareto improvement over $y$, then $x\,\mathrm{c}^*(\succsim)\,y$.*

| Ranking | | | | | |
|---|---|---|---|---|---|
| $(N)$ | $x$ | $a_1$ | $\dots$ | $a_k$ | $y$ |
| $(N+1)$ | $y$ | $x$ | $a_1$ | $\dots$ | $a_k$ |
| $(1)$ | $a_k$ | $\dots$ | $a_1$ | $y$ | $x$ |

TABLE 3. Exclusive Condorcet principle.

*Condorcet principle. If $x$ is a Condorcet winner, then $x\,\mathrm{c}^*(\succsim)\,y$ for all $y \in X \setminus \{x\}$.*

*Smith's principle. If $S$ and $T$ are a partition of $X$ such that $x \succsim y$ for every $x \in S$ and $y \in T$, then $x\,\mathrm{c}^*(\succsim)\,y$ for any $x \in S$ and $y \in T$.*

The next example verifies that the transitive core violates the exclusive Condorcet principle in general. While there are many reasons to support the criterion, the example suggests that the exclusive Condorcet principle might exclude the choice of an attractive policy in some cases.

EXAMPLE 5. Consider a society in which a policy is chosen from $\{x, y, a_1, \dots, a_k\}$ for $2N+2$ voters (with a large number $N$). Assume that voters' preferences in this society are distributed as in Table 3. We can observe that everyone except one voter ranks a policy $x$ second or higher, whereas $y$ is a controversial policy that splits the society about in half. Under the exclusive Condorcet principle, the social choice is uniquely determined by $y$, eliminating the possibility of choosing an attractive "second-best" policy $x$. The transitive core reserves the welfare comparison between $x$ and $y$ and, thus, leaves room to choose the policy $x$. ♢

In the context of social choice, many alternative methods of welfare inference are proposed in the literature, including those that offer *complete* welfare rankings over policies. For example, Rubinstein (1980) studies a ranking of policies induced by *the point system*. We can define the point system as a WER $\sigma$ such that

$$x\,\sigma(\succsim)\,y \quad \text{if and only if} \quad \big|\{z : x \succsim z\}\big| \geq \big|\{z : y \succsim z\}\big|$$

for every $x$ and $y$ in $X$, assuming that $X$ is finite. Then $\sigma(\succsim)$ is clearly a complete and transitive order on $X$. Moreover, $\sigma(\succsim)$ is a completion of the transitive core so that $x\,\mathrm{c}(\succsim)\,y$ implies $x\,\sigma(\succsim)\,y$ and that $x\,\mathrm{c}^*(\succsim)\,y$ implies $x\,\sigma^*(\succsim)\,y$, where $\mathrm{c}^*(\succsim)$ and $\sigma^*(\succsim)$ denote the strict parts of $\mathrm{c}(\succsim)$ and $\sigma(\succsim)$, respectively. Therefore, the point system always agrees with welfare comparisons by the transitive core, while it offers a complete welfare ranking at the same time. However, this does not necessarily imply that the transitive core is an inferior method of welfare inference. For example, the point system infers that the controversial policy $y$ is a strict welfare improvement over the second-best policy $x$ in the example of Table 3. We might still find the transitive core useful, as it identifies a more reliable part of welfare comparisons by the point system.

## 5. Concluding remarks

In this paper, we study methods of inferring the decision maker's true preference relation when the observed choice data reveal a complete but nontransitive preference relation. We formulate such methods as a concept that we call welfare evaluation rules, that is, functions mapping complete binary relations to transitive and reflexive binary relations. We discuss certain properties of welfare evaluation rules, and the transitive core is characterized as a unique welfare evaluation rule that satisfies all of the properties. The transitive core infers the decision maker's true preference of an alternative $x$ over another alternative $y$ when no third alternative conflicts with the revealed preference $x \succsim y$.

We study implications of the transitive core for various models of nontransitive preference relations. We show that the transitive core fully recovers an underlying utility function that measures the values of alternatives when it is applied to a semiorder. For preferences over intertemporal outcomes, the transitive core of relative discounting time preferences admits a representation by multiple absolute discounting functions. In other contexts, we study applications of the rule for preferences over ambiguous acts (justifiable preference relations), uncertain prospects (regret preference relations), and policy alternatives (majority preferences of the society). These examinations verify that the transitive core offers nontrivial and reasonable inference of the preference relation.

We note that this paper focuses on methods of inferring the decision maker's true preference relation from observation of a nontransitive revealed preference relation. This feature distinguishes this paper from the works of Bernheim and Rangel (2009) and others discussed in Section 1. However, the existing literature on behavioral welfare economics does not, by far, cover all the positive models of behavioral decision making. Finding intuitive and reliable methods of inferring evaluations of alternatives for boundedly rational decision makers remains an open problem of interest.

## Appendix A: Proofs

Proof of Theorem 1. (If) We show that the transitive core meets the axioms. Let $\succsim$ be a complete preference relation on $X$, let $S$ be a subset of $X$, and let $\pi$ be a permutation on $X$.

*Prudence*. Take any $x, y \in X$. As $\succsim_S$ is a reflexive binary relation on $X$, we have $x \succsim_S x$. Then, by (2), $x\, \mathrm{c}(\succsim_S)\, y$ implies $x \succsim_S y$.

*Principle of revealed preferences*. Let $(x, y)$ be a preference of $\succsim$ that is involved in none of its cycles. To the contrary, suppose that $x\, \mathrm{c}(\succsim)\, y$ does not hold. Then there is a $z \in X$ that meets either $y \succ z \succsim x$ or $y \succsim z \succ x$. In both cases, the cycle $(x, y, z)$ involves the preference $(x, y)$, a contradiction.

*Neutrality*. Recall that we denote by $\pi(R)$ the binary relation $\{(\pi(x), \pi(y)) : x\, R\, y\}$ for any binary relation $R$ on $X$. So, $x\, \pi(R)\, y$ if and only if $\pi^{-1}(x)\, R\, \pi^{-1}(y)$ for any $x$ and $y$

in $X$. Then a statement

$$\begin{cases} z\,\pi(\succsim)\,x & \text{implies} \quad z\,\pi(\succsim)\,y \\ y\,\pi(\succsim)\,z & \text{implies} \quad x\,\pi(\succsim)\,z \end{cases} \quad \text{for every } z \in X \tag{4}$$

is equivalent to a statement

$$\begin{cases} \pi^{-1}(z) \succsim \pi^{-1}(x) & \text{implies} \quad \pi^{-1}(z) \succsim \pi^{-1}(y) \\ \pi^{-1}(y) \succsim \pi^{-1}(z) & \text{implies} \quad \pi^{-1}(x) \succsim \pi^{-1}(z) \end{cases} \quad \text{for every } z \in X. \tag{5}$$

But, since $\pi^{-1}$ is bijective, the statement (5) is equivalent to

$$\begin{cases} z \succsim \pi^{-1}(x) & \text{implies} \quad z \succsim \pi^{-1}(y) \\ \pi^{-1}(y) \succsim z & \text{implies} \quad \pi^{-1}(x) \succsim z \end{cases} \quad \text{for every } z \in X. \tag{6}$$

Then $x\ \mathrm{c} \circ \pi(\succsim)\, y \Leftrightarrow (4) \Leftrightarrow (6) \Leftrightarrow \pi^{-1}(x)\,\mathrm{c}(\succsim)\,\pi^{-1}(y) \Leftrightarrow x\,\pi \circ \mathrm{c}(\succsim)\,y$, as desired.

*Inverse.* For any $x$ and $y$ in $X$, a statement

$$\begin{cases} z\ \mathrm{inv}(\succsim)\,x & \text{implies} \quad z\ \mathrm{inv}(\succsim)\,y \\ y\ \mathrm{inv}(\succsim)\,z & \text{implies} \quad x\ \mathrm{inv}(\succsim)\,z \end{cases} \quad \text{for every } z \in X \tag{7}$$

is equivalent to a statement

$$\begin{cases} x \succsim z & \text{implies} \quad y \succsim z \\ z \succsim y & \text{implies} \quad z \succsim x \end{cases} \quad \text{for every } z \in X. \tag{8}$$

Then $x\ \mathrm{c} \circ \mathrm{inv}(\succsim)\,y \Leftrightarrow (7) \Leftrightarrow (8) \Leftrightarrow y\,\mathrm{c}(\succsim)\,x \Leftrightarrow x\,\mathrm{inv} \circ \mathrm{c}(\succsim)\,y$.

*Downward consistency.* Suppose that $x\,\mathrm{c}(\succsim_T)\,y$ for some $x, y \in S \subseteq T$. Let $z \in X$ be such that $z \succsim_S x$. Then, by definition of the restricted relation $\succsim_S$, we have $z \in S$ and $z \succsim_T x$. As $x\,\mathrm{c}(\succsim_T)\,y$, it follows that $z \succsim_T y$ and $z \succsim_S y$. A similar argument shows that any $z \in X$ with $y \succsim_S z$ satisfies $x \succsim_S z$. So $x\,\mathrm{c}(\succsim_S)\,y$.

*Upward consistency.* Suppose that $x\,\mathrm{c}(\succsim_{xyz})\,y$ for every $z \in S$. Let $z \in X$ be such that $z \succsim x$. Then $z \succsim_{xyz} x$ and, thus, the hypothesis implies $z \succsim_{xyz} y$ and $z \succsim y$. Similarly, we can show that $y \succsim z$ implies $x \succsim z$ for any $z \in X$. So $x\,\mathrm{c}(\succsim)\,y$.

(Only if) Above we have shown that the transitive core is a WER that satisfies Axioms 1–6. For this part, it suffices to prove uniqueness of such a WER. We first note that a WER is characterized by its inference from revealed preference relations restricted on domains of three alternatives if it meets all the axioms.

Claim 1. *Let $\sigma$ and $\sigma'$ be two welfare evaluation rules that meet Axioms 1–6. If $\sigma(\succsim_{xyz}) = \sigma'(\succsim_{xyz})$ holds for any complete preference relation $\succsim$ on $X$ and any $x, y, z \in X$, then $\sigma = \sigma'$.*

Proof of Claim 1. Let $\sigma$ and $\sigma'$ be two WERs that meet the hypothesis. Take any complete preference relation $\succsim$ on $X$ and any set $S \subseteq X$. Define a complete binary relation $\trianglerighteq := \succsim_S \cup (X \times X \setminus S)$ on $X$, and observe that (a) $\trianglerighteq_S = \succsim_S$ and (b) $\trianglerighteq_{xyz}$ is transitive if $x, y \in S$ and $z \in X \setminus S$. For any $x, y \in S$, we can follow the implications

$$\begin{aligned} x\,\sigma(\succsim_S)\,y \quad &\Rightarrow \quad x\,\sigma(\succsim_{xyz})\,y \quad \text{for any } z \in S \quad &&\text{by Axiom 5} \\ &\Rightarrow \quad x\,\sigma(\trianglerighteq_{xyz})\,y \quad \text{for any } z \in S \quad &&\text{by (a)} \\ &\Rightarrow \quad x\,\sigma(\trianglerighteq_{xyz})\,y \quad \text{for any } z \in X \quad &&\text{by Axiom 2 and (b)} \\ &\Rightarrow \quad x\,\sigma(\trianglerighteq)\,y \quad &&\text{by Axiom 6} \\ &\Rightarrow \quad x\,\sigma(\trianglerighteq_S)\,y \quad &&\text{by Axiom 5} \\ &\Rightarrow \quad x\,\sigma(\succsim_S)\,y \quad &&\text{by (a)} \end{aligned}$$

to prove an equivalence $x\,\sigma(\succsim_S)\,y \Leftrightarrow x\,\sigma(\succsim_{xyz})\,y$ for any $z \in S$. Also, by replacing $\sigma$ with $\sigma'$ in the argument above, a similar equivalence holds for $\sigma'$. Then, by the hypothesis, it follows that $x\,\sigma(\succsim_S)\,y$ if and only if $x\,\sigma'(\succsim_S)\,y$ for every $x, y \in S$. As $\sigma$ and $\sigma'$ both admit Axiom 1, this shows that $\sigma(\succsim_S)$ and $\sigma'(\succsim_S)$ are identical. $\lhd$

By Claim 1, all we need to show is that the axioms uniquely determine $\sigma(\succsim_{xyz})$ for any complete preference relation $\succsim$ and any $x, y, z \in X$. By Axioms 1 and 2, if $\succsim_{xyz}$ is transitive, then we immediately have $\sigma(\succsim_{xyz}) = \succsim_{xyz}$. Also, we proved that $\sigma(\succsim_{xyz}) = \{(x, y), (x, z), (z, y)\}$ when $\succsim_{xyz}$ is given as $\succsim_1$ in Figure 3 (Section 3.2). If $\succsim_{xyz} = \{(x, y), (y, z), (z, x)\} \cup \Delta_X$, we can readily show that Axioms 1 and 3 imply $\sigma(\succsim_{xyz}) = \Delta_X$. The rest of the proof covers a case where $\succsim_{xyz} = \{(x, y), (y, z), (z, x), (x, z)\} \cup \Delta_X$. This exhausts all cyclic preference relations $\succsim_{xyz}$ on three alternatives by symmetry. Figure 5 presents all binary relations used below. For ease of reference, I denote by $\succsim_{ij}$ a preference relation of the $i$th row and the $j$th column in the figure. We consider a case where $\succsim_{xyz} = \succsim_{11}$, and the proof will show that $\sigma(\succsim_{11}) = \succsim_{14}$.

First, we show that $\sigma(\succsim_{11})$ must be either $\succsim_{14}$, $\succsim_{21}$, or $\succsim_{22}$. To see this, note that a preference $(x, z)$ of $\succsim_{11}$ is not involved in any cycles of $\succsim_{11}$, and, hence, Axiom 2 implies $x\,\sigma(\succsim_{11})\,z$. Also, observe that

$$x\,\sigma(\succsim_{11})\,y \quad \overset{\text{Inverse}}{\Longleftrightarrow} \quad y\,\sigma(\succsim_{12})\,x \quad \overset{\text{Neutrality}}{\Longleftrightarrow} \quad y\,\sigma(\succsim_{13})\,z$$

by Axioms 3 and 4. But $\succsim_{11} = \succsim_{13}$. So, $\sigma(\succsim_{11})$ either contains both $(x, y)$ and $(y, z)$ or contains neither of them. These two observations along with Axiom 1 imply that $\sigma(\succsim_{11})$ is either of $\succsim_{14}$, $\succsim_{21}$, or $\succsim_{22}$.

Next, we show that $\sigma(\succsim_{11})$ cannot be $\succsim_{21}$. Assume to the contrary. Take any $w$ in $X$ distinct from $x$, $y$, $z$, and let us consider a preference relation $\succsim_{23}$. Then the hypothesis, Axiom 1, Axiom 2, and Axiom 6 together imply $z\,\sigma(\succsim_{23})\,x$ and $x\,\sigma(\succsim_{23})\,w$. Since $\sigma(\succsim_{23})$ is transitive, these also imply $z\,\sigma(\succsim_{23})\,w$. We then have $z\,\sigma(\trianglerighteq)\,w$ by Axiom 5, where $\trianglerighteq$ is the restriction of $\succsim_{23}$ on $\{y, z, w\}$. This contradicts our previous observation $\sigma(\trianglerighteq) = \Delta_X$.

We show that $\sigma(\succsim_{11})$ is not $\succsim_{22}$ either. Assume to the contrary. Take any $w$ in $X$ distinct from $x, y, z$. Let us first prove that $\sigma(\succsim_{32}) = \{(z, x), (z, w), (w, x)\}$. For this, consider $\succsim_{24}$. Then the hypothesis, Axiom 1, Axiom 2, and Axiom 6 together imply that

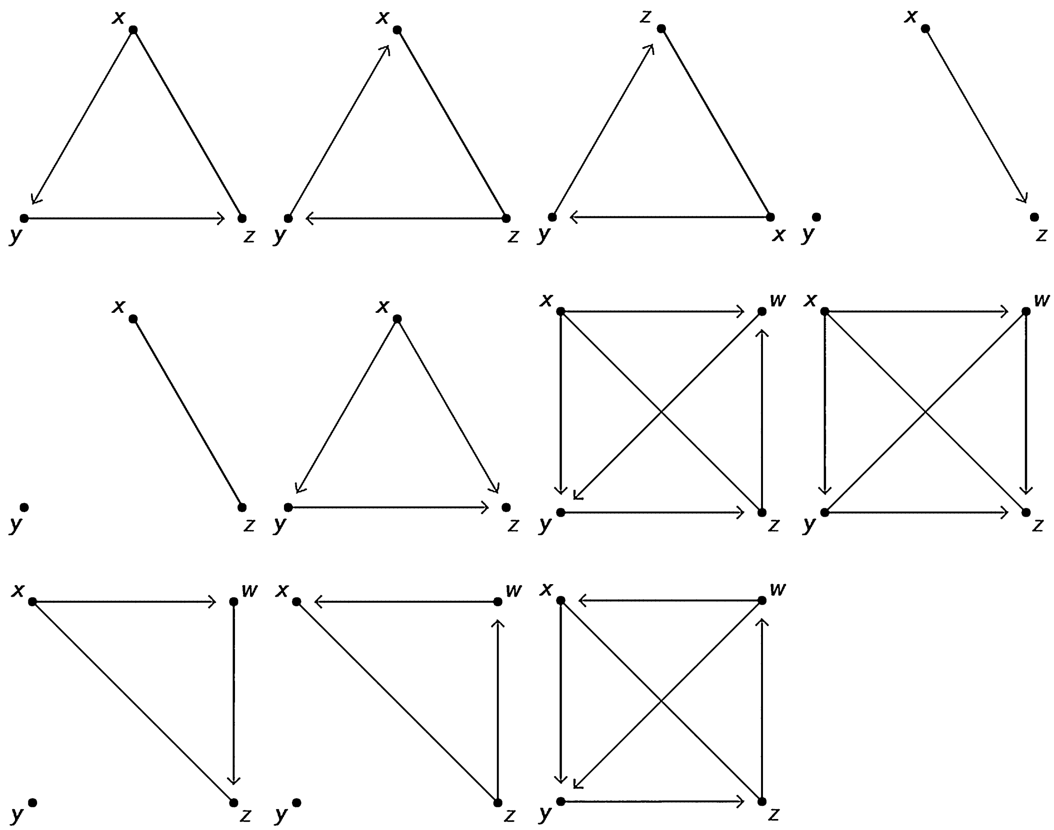


FIGURE 5. Proof of Theorem 1.

$x\,\sigma(\succsim_{24})\,y$ and $y\,\sigma(\succsim_{24})\,z$. These then imply that $x\,\sigma(\succsim_{24})\,w$ and $w\,\sigma(\succsim_{24})\,z$ by Axiom 3. Restricting $\succsim_{24}$ on $\{x, z, w\}$, Axiom 5 therefore implies that $x\,\sigma(\succsim_{31})\,w$ and $w\,\sigma(\succsim_{31})\,z$. As $\succsim_{32}$ is inverse to $\succsim_{31}$, it follows that $w\,\sigma(\succsim_{32})\,x$ and $z\,\sigma(\succsim_{32})\,w$ by Axiom 4. By transitivity of $\sigma(\succsim_{32})$ and Axiom 1, this completes to show that $\sigma(\succsim_{32}) = \{(z, x), (z, w), (w, x)\}$. Now consider $\succsim_{33}$. Then we have $w\,\sigma(\succsim_{33})\,x$ and $x\,\sigma(\succsim_{33})\,y$ by Axiom 6 and, thus, $w\,\sigma(\succsim_{33})\,y$ by transitivity. Axiom 5 hence implies $w\,\sigma(\trianglerighteq)\,y$, where $\trianglerighteq$ is the restriction of $\succsim_{33}$ on $\{y, z, w\}$. This contradicts our previous observation $\sigma(\trianglerighteq) = \Delta_X$. A conclusion is $\sigma(\succsim_{11}) = \succsim_{14}$.

We have shown that the axioms uniquely determine $\sigma(\succsim_{xyz})$ for any complete preference relation $\succsim$ on $X$ and any $x, y, z \in X$. By Claim 1, this proves uniqueness of a welfare evaluation rule that satisfies Axioms 1–6. The proof of Theorem 1 is now complete. □

PROOF OF PROPOSITION 2. Let $X$ be a connected metric space and let $\succsim$ be a semiorder on $X$ with a representation $(u, \epsilon)$, where $\sup |u(x) - u(y)| > 2\epsilon$. Take any $x, y \in X$ with $u(x) \geq u(y)$. If $z \in X$ is such that $z \succsim x$, then $u(z) \geq u(x) - \epsilon \geq u(y) - \epsilon$ and, thus, $z \succsim y$. If $z \in X$ is such that $y \succsim z$, then $u(x) \geq u(y) \geq u(z) - \epsilon$ and, thus, $x \succsim z$. So $x\,\mathrm{c}(\succsim)\,y$. For the converse, we prove the contrapositive. Take any $x, y \in X$ with $u(y) > u(x)$, and we show that $x \, \mathrm{c}(\succsim) \, y$ does not hold. Note that, by the hypothesis, there exists a $z \in X$ that meets either $u(z) > u(x) + \epsilon$ or $u(y) - \epsilon > u(z)$. Assume the existence of $z$ with the former inequality. (The proof is similar if the latter holds.) Then we can let $u(y) + \epsilon > u(z) > u(x) + \epsilon$ without loss of generality, for $u(X)$ is an interval by continuity of $u$. So $y \succsim z \succ x$, negating $x \, \mathrm{c}(\succsim) \, y$. □

Proof of Theorem 3. Let $\succsim$ be a relative discounting time preference with an associated representation $(u, \eta)$. Define $\mathcal{D} := \{\eta(\cdot, r) : r \in [0, \infty)\}$. Then every $\delta \in \mathcal{D}$ is a continuous function from $\mathbb{R}_+$ to $\mathbb{R}_{++}$. Suppose that $(x, t), (y, s) \in X$ are dated outcomes such that $(x, t) \, \mathrm{c}(\succsim) \, (y, s)$. Take any $\delta \in \mathcal{D}$ and let $r \in [0, \infty)$ be such that $\delta(\cdot) = \eta(\cdot, r)$. As $u$ is a homeomorphism from $Z$ to $\mathbb{R}_{++}$, there exists a $z \in Z$ such that $u(z) = \eta(s, r)u(y)$. Then we have $(y, s) \sim (z, r)$ by the representation and $(x, t) \succsim (z, r)$ by the hypothesis. So $\delta(t)u(x) = \eta(t, r)u(x) \geq u(z) = \eta(s, r)u(y) = \delta(s)u(y)$, as desired. For the converse, let $(x, t), (y, s) \in X$ be dated outcomes such that $\delta(t)u(x) \geq \delta(s)u(y)$ for any $\delta \in \mathcal{D}$. Take any $(z, r) \in X$ with $(y, s) \succsim (z, r)$. We have $\eta(s, r)u(y) \geq u(z)$ by the representation and $\eta(t, r)u(x) \geq \eta(s, r)u(y)$ by the hypothesis. Then $\eta(t, r)u(x) \geq u(z)$ and, thus, $(x, t) \succsim (z, r)$. We can similarly show that any $(z, r)$ with $(z, r) \succsim (x, t)$ satisfies $(z, r) \succsim (y, s)$. So $(x, t) \, \mathrm{c}(\succsim) \, (y, s)$. □

Proof of Proposition 4. Let $\succsim_J$ and $\succsim_K$ be a justifiable preference relation and a Knightian preference relation on $X$, respectively, with a representing pair $(u, P)$ of a utility function and a set of priors. Below I write $\varphi(f, p) = \sum_{s \in S} p(s)u(f(s))$ for any $f \in X$ and $p \in P$. We wish to show that $\mathrm{c}(\succsim_J) = \succsim_K$. First, take any $f, g \in X$ with $f \succsim_K g$, that is, $\varphi(f, p) \geq \varphi(g, p)$ for any $p \in P$. If $h \in X$ is such that $h \succsim_J f$, then there is a $p \in P$ with $\varphi(h, p) \geq \varphi(f, p) \geq \varphi(g, p)$ and, thus, $h \succsim_J g$. If $h \in X$ is such that $g \succsim_J h$, then there is a $p \in P$ with $\varphi(f, p) \geq \varphi(g, p) \geq \varphi(h, p)$ and, hence, $f \succsim_J g$. So $f \, \mathrm{c}(\succsim_J) \, g$. To show the converse, take any $f, g \in X$ with $f \, \mathrm{c}(\succsim_J) \, g$ and suppose that $f \succsim_K g$ does not hold by contradiction. Then there is a $p^* \in P$ such that $\varphi(g, p^*) > \varphi(f, p^*)$. For any positive real number $\epsilon > 0$, let $f^\epsilon$ be an act that gives a prize larger than that of $f$ by $\epsilon$ at any realization of state $s \in S$ and any resolution of lottery $f(s)$. By continuity of $u$, we can pick an $\epsilon > 0$ small enough so that $\varphi(g, p^*) > \varphi(f^\epsilon, p^*)$. But, for such an $\epsilon$, $\varphi(f^\epsilon, p) > \varphi(f, p)$ for all $p \in P$ as $f^\epsilon(s)$ first-order stochastically dominates $f(s)$ for all states $s \in S$. So $g \succsim_J f^\epsilon \succ_J f$, contradicting $f \, \mathrm{c}(\succsim_J) \, g$. □

Proof of Proposition 5. Let $\succsim$ be a regret preference on $X$ with an associated representation $(u, Q)$. Let $x, y \in X$ be two prospects such that $x_i \geq y_i$ for all $i$. If $z \in X$ is such that $y \succsim z$, then by monotonicity of $u$ and $Q$,

$$\sum_{i=1}^{n} p_i Q\big(u(x_i) - u(z_i)\big) \geq \sum_{i=1}^{n} p_i Q\big(u(y_i) - u(z_i)\big) \geq 0$$

and, thus, $x \succsim z$. If $z \succsim x$, $\sum_{i=1}^n p_i Q(u(z_i) - u(y_i)) \geq \sum_{i=1}^n p_i Q(u(z_i) - u(x_i)) \geq 0$ and, hence, $z \succsim y$. So $x \, \mathrm{c}(\succsim) \, y$. If, in addition, $x_i > y_i$ for some $i$, then $\sum_{i=1}^n p_i Q(u(x_i) - u(y_i)) > 0$, as $u$ and $Q$ are strictly increasing and $Q(0) = 0$. Hence, we have $x \succ y$, implying that $y \, \mathrm{c}(\succsim) \, x$ does not hold. □

Proof of Proposition 6. Let $X$ be an arbitrary nonempty set and let $\succsim$ be a majority preference induced by a set of linear orders $\succsim_i$ on $X$ for $i = 1, \dots, n$. To verify the Pareto criterion, let $x, y \in X$ be two policies such that $x \succ_i y$ for all $i$. Observe that, for an arbitrary policy $z \in X$, $\{i : y \succsim_i z\} \subseteq \{i : x \succsim_i z\}$ and $\{i : z \succsim_i x\} \subseteq \{i : z \succsim_i y\}$ as each $\succsim_i$ is transitive. So if $y \succsim z$, then

$$\left|\{i : x \succsim_i z\}\right| \geq \left|\{i : y \succsim_i z\}\right| \geq \left|\{i : z \succsim_i y\}\right| \geq \left|\{i : z \succsim_i x\}\right|$$

and, thus, $x \succsim z$. We can similarly show that $z \succsim x$ implies $z \succsim y$. So $x\, \mathrm{c}(\succsim)\, y$. Also, the hypothesis implies that $x \succ y$ and, thus, $y\, \mathrm{c}(\succsim)\, x$ does not hold. Next, to check Smith's principle, let $S$ and $T$ be a partition of $X$ such that $x \succ y$ for each $x \in S$ and $y \in T$. Fix any $x \in S$ and $y \in T$. If $z \in X$ is such that $y \succsim z$, then $z$ must be a member of $T$ and, thus, $x \succ z$. Similarly, if $z \in X$ is such that $z \succsim x$, then $z \in S$ and $z \succ y$. So $x\, \mathrm{c}(\succsim)\, y$. Of course, $x \succ y$ and, hence, $y\, \mathrm{c}(\succsim)\, x$ does not hold. The Condorcet principle is implied by Smith's principle. □

## Appendix B: Supplementary results

*Independence of the axioms*. We show that the axioms introduced in Section 3.1 are mutually independent. As remarked above, the universally incomparable WER satisfies all but Axiom 2, and the universally indifferent WER $\sigma_1$ satisfies all but Axiom 1. Also, the covering order discussed in Section 1 is a WER that satisfies all but Axiom 4. For Axiom 3, fix any two alternatives $x^*$ and $y^*$ in $X$, and define a WER $\sigma$ by the following rule. Set $\sigma(\succsim_{xyz}) = \mathrm{c}(\succsim_{xyz})$ for any complete preference relation $\succsim$ on $X$, and $x, y, z \in X$ unless $x \succ y \succ z \succ x$ and $\{x^*, y^*\} \subseteq \{x, y, z\}$, in which case we set $\sigma(\succsim) = \succsim \cap \{(x^*, y^*), (y^*, x^*)\}$. In turn, for any complete preference relation $\succsim$ on $X$ and any $S \subseteq X$ with $|S| > 3$, we define $\sigma(\succsim_S)$ by $x\, \sigma(\succsim_S)\, y$ if and only if $x\, \sigma(\succsim_{xyz})\, y$ for all $z \in X$. Then $\sigma$ meets all the axioms except Axiom 3. For Axiom 5, define a WER $\sigma$ by $\sigma(\succsim_S) = \mathrm{c}(\succsim_S)$ for any complete preference relation $\succsim$ on $X$ and any $S \subseteq X$ unless $\succsim_S = \succsim_{33}$ in Figure 5 for some $x, y, z, w \in X$. In the excluded case, let $\sigma(\succsim_S) = \{(x, z), (z, x), (y, w), (w, y)\}$. Then $\sigma$ satisfies all but Axiom 5. Last, for any binary relation $\trianglerighteq$ on $X$, define a binary relation $\trianglerighteq^\circ$ by $x \trianglerighteq^\circ y$ if and only if $x \trianglerighteq y$ and no cycle of $\trianglerighteq$ involves $(x, y)$. In turn, define a welfare evaluation rule $\sigma$ by mapping $\succsim_S$ to the transitive closure of $\succsim_S^\circ$ for any complete preference relation $\succsim$ on $X$ and any $S \subseteq X$. Then $\sigma$ satisfies all but Axiom 6.

## References

Anscombe, F. J. and R. J. Aumann (1963), "A definition of subjective probability." *Annals of Mathematical Statistics*, 34, 199–205. [593]

Armstrong, W. E. (1939), "The determinateness of the utility function." *Economic Journal*, 49, 453–467. [581, 589]

Armstrong, W. E. (1948), "Uncertainty and the utility function." *Economic Journal*, 58, 1–10. [581, 589]

Armstrong, W. E. (1950), "A note on the theory of consumer's behaviour." *Oxford Economic Papers, New Series*, 2, 119–122. [581, 589]

Armstrong, W. E. (1951), "Utility and the theory of welfare." *Oxford Economic Papers, New Series*, 3, 259–271. [581, 589]

Arrow, Kenneth J. (1951), *Social Choice and Individual Values*. Wiley, New York. [596]

Bernheim, B. Douglas and Antonio Rangel (2007), "Toward choice-theoretic foundations for behavioral welfare economics." *American Economic Review*, 97, 464–470. [579, 581]

Bernheim, B. Douglas and Antonio Rangel (2009), "Beyond revealed preference: Choice-theoretic foundations for behavioral welfare economics." *Quarterly Journal of Economics*, 124, 51–104. [581, 599]

Bewley, Truman F. (2002), "Knightian decision theory. Part I." *Decisions in Economics and Finance*, 25, 79–110. [593]

Bikhchandani, Sushil and Uzi Segal (2011), "Transitive regret." *Theoretical Economics*, 6, 95–108. [595]

Caplin, Andrew, Mark Dean, and Daniel Martin (2011), "Search and satisficing." *American Economic Review*, 101, 2899–2922. [579]

Chambers, Christopher P. and Takashi Hayashi (2012), "Choice and individual welfare." *Journal of Economic Theory*, 147, 1818–1849. [581]

Fishburn, Peter C. (1970), "Intransitive indifference with unequal indifference intervals." *Journal of Mathematical Psychology*, 7, 144–149. [581, 582, 589]

Fishburn, Peter C. (1977), "Condorcet social functions." *SIAM Journal on Applied Mathematics*, 33, 469–489. [581, 597]

Gul, Faruk and Wolfgang Pesendorfer (2001), "Temptation and self-control." *Econometrica*, 69, 1403–1435. [585]

Kahneman, Daniel and Amos Tversky (1979), "Prospect theory: An analysis of decision under risk." *Econometrica*, 47, 263–292. [580, 594]

Kőszegi, Botond and Matthew Rabin (2007), "Mistakes in choice-based welfare analysis." *American Economic Review*, 97, 477–481. [581]

Lehrer, Ehud and Roee Teper (2011), "Justifiable preferences." *Journal of Economic Theory*, 146, 762–774. [593, 594]

Loomes, Graham and Robert Sugden (1982), "Regret theory: An alternative theory of rational choice under uncertainty." *Economic Journal*, 92, 805–824. [581, 595]

Luce, R. Duncan (1956), "Semiorders and a theory of utility discrimination." *Econometrica*, 24, 178–191. [581, 589]

Manzini, Paola and Marco Mariotti (2007), "Sequentially rationalizable choice." *American Economic Review*, 97, 1824–1839. [579]

Masatlioglu, Yusufcan, Daisuke Nakajima, and Erkut Y. Ozbay (2012), "Revealed attention." *American Economic Review*, 102, 2183–2205. [579]

Miller, Nicholas R. (1980), "A new solution set for tournaments and majority voting: Further graph-theoretical approaches to the theory of voting." *American Journal of Political Science*, 24, 68–96. [581]

Ok, Efe A. and Yusufcan Masatlioglu (2007), "A theory of (relative) discounting." *Journal of Economic Theory*, 137, 214–245. [581, 591, 592]

Read, Daniel (2001), "Is time-discounting hyperbolic or subadditive?" *Journal of Risk and Uncertainty*, 23, 5–32. [581, 591]

Roelofsma, Peter H. M. P. and Daniel Read (2000), "Intransitive intertemporal choice." *Journal of Behavioral Decision Making*, 13, 161–177. [581]

Rozen, Kareen (2010), "Foundations of intrinsic habit formation." *Econometrica*, 78, 1341–1373. [585]

Rubinstein, Ariel (1980), "Ranking the participants in a tournament." *SIAM Journal on Applied Mathematics*, 38, 108–111. [592, 598]

Rubinstein, Ariel (2003), "'Economics and psychology'? The case of hyperbolic discounting." *International Economic Review*, 44, 1207–1216. [581]

Rubinstein, Ariel and Yuval Salant (2006), "A model of choice from lists." *Theoretical Economics*, 1, 3–17. [579]

Rubinstein, Ariel and Yuval Salant (2012), "Eliciting welfare preferences from behavioural data sets." *Review of Economic Studies*, 79, 375–387. [581]

Salant, Yuval (2011), "Procedural analysis of choice rules with applications to bounded rationality." *American Economics Review*, 101, 724–748. [579]

Salant, Yuval and Ariel Rubinstein (2008), "$(A, f)$: Choice with frames." *Review of Economic Studies*, 75, 1287–1296. [579]

Smith, John H. (1973), "Aggregation of preferences with variable electorate." *Econometrica*, 41, 1027–1041. [597]

Tyson, Christopher J. (2008), "Cognitive constraints, contraction consistency, and the satisficing criterion." *Journal of Economic Theory*, 138, 51–70. [579]

von Neumann, John and Oskar Morgenstern (1944), *Theory of Games and Economic Behavior*. Princeton University Press, Princeton, New Jersey. [594]

---

Co-editor Faruk Gul handled this manuscript.

Manuscript received 22 January, 2014; final version accepted 29 January, 2017; available online 28 June, 2017.