School of Economics and Finance Discussion Papers

University of St Andrews

Preference intensity representation and revelation

Georgios Gerasimou

School of Economics and Finance Discussion Paper No. 1716
4 Dec 2017 (revised 1 Jun 2019)
JEL Classification:
Keywords:

School of Economics and Finance
Online Discussion Paper Series
issn 2055-303X
http://ideas.repec.org/s/san/wpecon.html
info: econ@st-andrews.ac.uk

Preference Intensity Representation and Revelation

Georgios Gerasimou*†
University of St Andrews

June 1, 2019

First draft: 4th December 2017

Abstract

This paper revisits the problem of preference intensity modelling by proposing and analysing the novel concept of *preference intensity functions*. These retain what are argued to be the most essential properties expected from the consistent numerical representation of preference intensity orderings and the ordinary preferences induced by them. Their existence is characterized by simple and clearly interpretable axioms, while they are also shown to be genuinely ordinally unique and more general than utility functions associated with utility-difference representations. The empirical content of this model is pinned down by means of weak necessary and sufficient condition on observable behavioural data that include choices and additional observables with intensity-revealing potential such as response times, willingness to pay or survey ratings. Finally, a particular normalization of the model allows for simple ordinal interpersonal comparisons of preference intensities to be made, which in turn invite a novel notion of *intensity efficiency* that is shown to be well-defined and to refine Pareto efficiency by discarding allocations that are dominated on intensity-difference grounds.

*I am grateful to Jean Baccelli and John Quah for very useful comments on a previous version. I also thank Tore Ellingsen, Ian Krajbich, Peter Wakker and audiences at the North American Summer Meetings of the Econometric Society (2018, Davis) and the Symposium on the Analysis of Revealed Preference (2018, Oxford) for helpful feedback. Any errors are my own.

†Email address: gg26@st-andrews.ac.uk

1 Introduction

The purpose of this study is to revisit and contribute to the foundations of preference-intensity modelling and its empirical and welfare-theoretic applications by raising and providing positive answers to the following questions:

1. Can a decision maker's preferences and preference-intensity comparisons on a finite set of arbitrary choice alternatives (e.g. indivisible goods) admit a simple numerical representation that features: (a) transparent and minimally restrictive behavioural underpinnings; (b) no confounds between intensity differences and attitudes to risk; (c) genuine ordinal uniqueness properties, and hence intensity comparisons that are not required to be *"as precise as length comparisons made by precision instruments"*[1]; (d) a generalization over existing utility-difference representations; (e) a foundation for purely ordinal interpersonal comparability of preference intensities?
2. For a decision maker whose preference intensity ordering belongs to this general class, what is the empirical content of statements like *"I prefer my child enrolling in School A than in School B more than I prefer it enrolling in School C than in School D"*? More generally, what kind of observable data and testable conditions on such data are required for such a decision maker's preferences and preference intensities to be recovered from his behaviour?
3. Does the ordinal interpersonal comparability of preference intensities that was claimed above allow for a well-defined and generally applicable refinement of Pareto efficiency to be introduced in such environments, thereby enabling the analyst/policy maker to discard those Pareto efficient allocations that are dominated on intensity-difference grounds?

Our analytical approach towards answering the first question –and hence the two that follow it– takes Samuelson's (1938) bivariate formulation of neoclassical cardinal utility representations as a starting point and extends it in a natural and but hitherto unexplored direction. The resulting representation combines simplicity, generality and applicability, and also addresses conceptual and analytical concerns about cardinal and more general utility-difference approaches to preference-intensity modelling that were expressed formally by Samuelson himself and were complemented by several other authors in the years that followed. A detailed summary of what might be considered some of the most significant challenges of the utility-difference approach to intensity modelling –for which the opening remarks above provide a quick preview– is given in the next section.

In response to these challenges, we propose and analyze the general class of *preference intensity functions/representations* that overcomes them while retaining what are argued to be the key ingredients of a satisfactory numerical representation of preference intensity comparisons and the ordinary preferences induced by them. Preference intensity functions and representations mimic *ordinal* utility ones by associating each *pair* of alternatives with a numerical value in a way that preserves the preference-intensity ordering. In particular, if a is preferred to b more than c is to d, then a higher value $s(a, b)$ is associated with the pair (a, b) than with (c, d). The second property, skew-symmetry, offers a convenient normalization, which is shown to be without loss of generality. The most important input of this normalization is that it allows the analyst to determine whether the first or

[1] Fishburn (1970, p. 82)

second alternative at a given pair is weakly preferred to the other by checking whether the function's sign at that pair is weakly positive or negative, respectively. The key novelty of preference intensity representations that formally distinguishes them from utility-difference representations –or bivariate representations that are ultimately formally equivalent to those– is a property that we refer to as *lateral consistency*. This requires that whenever a is preferred to b and b to c, then a is preferred to c more than a is to b and b is to c. It ensures conceptual harmony between and within the preference intensity and induced preference relations without forcing –as is done by utility-difference models– that the value capturing the intensity difference between a and c be the *sum* of the values for the intensity differences between a and b and between b and c. By not imposing this requirement, the model is characterized by standard Weak Order and Reversal together with either a mild Consistency axiom or, equivalently under the former two, standard Separability.

The proposed class of functions are the first in the literature that offer a genuinely ordinally unique representation of preference intensity relations. For comparison, cardinal utility models portray the decision maker as if he was able to think of a as being preferred to b exactly r many times as c is preferred to b. This degree of precision is obviously unrealistic. Other, non-cardinal utility-difference models do not make this prediction but are not invariant with respect to arbitrary strictly increasing transformations either. By contrast, in their most general formulation –which is shown to be formally equivalent to the normalized, skew-symmetric one that we use in most of the paper– preference intensity functions are simply unique up to arbitrary strictly increasing transformations, hence ordinally unique in the standard sense. This property is important because it clarifies that the decision maker's intensity comparisons in this model are not assumed to have any structure other than to reflect simple and intuitively realistic statements such as *"I prefer a to b more than I prefer c to d"*.

Our first theoretical application of the model leads to a significant generalization of the recent analysis in Echenique and Saito (2017). These authors proposed a notion of utility-difference rationalizability for choice and response-time data, assuming –intuitively, and also in line with recent experimental and theoretical work– a monotone relationship between response times and utility-differences. For finite and possibly incomplete datasets comprising binary choices and an additional menu-specific observable with intensity-revealing potential (e.g. response times; willingness to pay for the chosen alternative at a given menu; intensity rating of the chosen alternative relative to the non-chosen one), the *Congruent Monotonicity* axiom that we introduce is shown to be necessary and sufficient for behaviour to be *preference-intensity rationalizable*. When this is the case, the relevant dataset can be thought of as being generated by the behaviour of a decision maker whose preferences and preference-intensity comparisons are representable by a preference intensity function that is strictly monotonic in the observed resource for all relevant binary menus. Congruent Monotonicity combines the acyclicity restriction of the well-known Congruence axiom (Richter, 1966) with a cross-modal consistency requirement whereby the value of the intensity-revealing resource is monotonic in a natural sense for every sequence of alternatives where any two consecutive elements are related by revealed preference. By using preference intensity functions instead of difference-preserving utility functions as its building block, this more general notion of preference-intensity prationalizability –and hence the axiom that characterizes it– can account for substantially more behaviour without sacrificing completeness, transitivity or conceptual harmony between the revealed preference and preference intensity relations.

Our second theoretical application of the model is welfare-theoretic and leads to an intuitive refinement of Pareto efficiency that takes into account interpersonal differences in preference intensities. Specifically, we first make note of a particularly convenient *canonical* normalization of the proposed numerical representation whereby all agents' (strict) can be assumed to have preference intensity functions whose range is the same set of consecutive integers. This canonical normalization is analogous to the one made in relative utilitarianism studies[2] where the agents' cardinal (in fact, von Neumann-Morgenstern) utility functions are assumed to have the same range –typically the $[0,1]$ interval. As we show constructively in the proof of Theorem 1, such a canonical normalization is guaranteed to be possible under the above assumptions on the agents' intensity relations. Its usefulness in the context of the present welfare-theoretic problem lies in the fact that it allows for the statement $s_i(a,b) > s_j(a,b) > 0$ with respect to agents i,j and alternatives a,b to be interpreted as revealing that i prefers a to b more than j does. Thus, it makes ordinal interpersonal comparisons of preference intensities possible *without requiring interpersonally comparable utilities*. Building on this normalization –which is generally impossible under the utility-difference approach– we then introduce a notion of *intensity-efficiency* that is shown to be well-defined and to refine Pareto efficiency by discarding allocations that are unappealing once intensity-differences are also taken into account. In short, the novel part of the underlying notion of *intensity-dominance* through which intensity efficiency is defined postulates that whenever two allocations assign the same two alternatives to a given pair of agents but in opposite ways, then the dominant allocation is the one that assigns the commonly preferred object to the agent who prefers it most, other things equal.

The remainder of the paper is structured as follows. The next section introduces preference intensity relations and their utility-difference representations, and also discusses some of the challenges that are associated with this modelling approach. Motivated by this discussion, it then proceeds to the introduction and axiomatic characterization of preference intensity functions, clarifying why such representations are genuinely ordinally unique and how they include utility-difference representations as special cases. Novel connections are then made between these models and their stochastic-choice analogs, i.e. the *Fechnerian* and *scalable* models. The section following this analysis provides the revealed-preference foundations of the model, while the one after it investigates the welfare-theoretic problem mentioned above. The last section concludes. All proofs appear in the Appendix.

2 The Representation Problem

2.1 Preliminaries: Preference Intensity Relations and Utility Differences

Assumed throughout is a set X of general choice alternatives and a binary relation $\succsim$ on $X \times X$, to be thought of as a *preference intensity relation* on X. A concrete behavioural interpretation of the comparison $(a,b) \succsim (c,d)$ is that if the decision maker was to imagine being endowed with alternatives b and d and having to decide between moving away from b towards a and also from d towards c, then he would weakly prefer the former transition. Although we will return to this issue

[2]For example, Dhillon and Mertens (1999).

in more detail below, we can state from the outset that such the comparison $(a,b) \gtrsim (c,d)$ could be defined by one of the following:

1. Surveys: a is stated to be preferred to b more than c is to d;
2. Willingness to pay: higher for a at $\{a,b\}$ than for c at $\{c,d\}$;
3. Choice probabilities: a is chosen more frequently at $\{a,b\}$ than c is at $\{c,d\}$;
4. Response times: decision faster at $\{a,b\}$ than at $\{c,d\}$.

An ordinary preference relation $\gtrapprox$ on X is *induced* by a preference intensity relation $\gtrsim$ when the former is defined by[3]

$$a \gtrapprox b \iff (a,b) \gtrsim (b,a). \tag{1}$$

The asymmetric and symmetric parts of $\gtrsim$ and $\gtrapprox$ will be denoted $\succ, \sim$ and $\ggcurly, \approx$, respectively.

Definition 1
A binary relation $\gtrsim$ *on a set* $X \times X$ *admits a* utility-difference representation *if there exists a function* $u : X \to \mathbb{R}$ *such that, for all* $a,b,c,d \in X$,

$$(a,b) \gtrsim (c,d) \iff u(a) - u(b) \geqslant u(c) - u(d). \tag{2}$$

In addition, $\gtrsim$ *admits a* cardinal utility representation *if such a* u *exists and is unique up to a positive affine transformation.*

Cardinal-utility and utility-difference representations have a special place in the history of economic thought and, often under different names, have been at the heart of much inter-disciplinary research in the theory of measurement.[4] Some of the critical insights that have been offered by their extended study over the past several decades and which are particularly relevant for the present paper's motivation and focus are put together and briefly discussed below.

Riskless neo-classical cardinal utility

1. The neo-classical cardinal utility model implies that the agent's preference intensity comparisons are not only well-defined and consistent, but also precise to a behaviourally questionable degree. In particular, if such a relation $\gtrsim$ is represented by a utility function u of this kind, then cardinality of u implies that the statement $(a,b) \gtrsim (c,d)$ leads to a utility-difference ratio $\frac{u(a)-u(b)}{u(c)-u(d)} := r$ that is invariant with respect to all permissible transformations of u. Assuming, for simplicity, that both utility differences are positive here, this in turn translates into the claim that *"a is preferred to b* exactly *r times as much as c is preferred to d"*.

[3]Alternatively, $\gtrapprox$ could be defined by $a \gtrapprox b \Leftrightarrow (a,b) \gtrsim (c,c)$ for any $c \in X$. In our environment, the two definitions will turn out to be equivalent.

[4]For detailed and complementary accounts of these models, the debates they have been associated with and their implications for economic analysis, the reader is referred to Luce, Krantz, Suppes, and Tversky (1990), Krantz, Luce, Suppes, and Tversky (1971), Suppes, Krantz, Luce, and Tversky (1989), Hammond (1991), Ellingsen (1994), Mandler (1999), Falmagne (2002), Bossert and Weymark (2004), Fleurbaey and Hammond (2004), Abdellaoui, Barrios, and Wakker (2007), Baccelli and Mongin (2016) and Moscati (2018).

2. Axiomatizations of the neo-classical cardinal utility model typically impose restrictions on the preference intensity relation that force the underlying set of choice alternatives to be infinite. On the other hand, when this set is specifically required to be finite, the only known axiomatization of the model involves an "equal-spacing" axiom that requires preference intensity differences between any two pairs of consecutive –in the induced preference ranking– alternatives to be equivalent[5]. This assumption, however, effectively makes the model non-operational in such domains.

von Neumann-Morgenstern cardinal utility

It is not uncommon for the cardinally unique utility differences between pairs of riskless alternatives that can be obtained by lottery comparisons under the von Neumann and Morgenstern (1947) expected-utility model to be interpreted as indicating differences in preference intensities between the relevant riskless outcomes. However, it has been known for a long time that such utility differences confound the decision maker's preference intensities over the underlying set of riskless alternatives with his attitudes toward risk (Luce and Raiffa, 1957; Hammond, 1991; Ellingsen, 1994; Baccelli and Mongin, 2016). On these grounds, expected-utility theory provides an unsatisfactory approach to preference-intensity modelling.

Cardinal utility from stochastic choices

Debreu (1958) provided an axiomatization of cardinal utility that is based on comparisons of binary choice probabilities. As in the case of the riskless model, the set of alternatives here must again be infinite to avoid the triviality that was mentioned in the context of that model. In addition to this limitation, however, some conceptual concerns have also been raised about this approach to modelling intensity. First, as pointed out in Davidson and Marschak (1959), this model effectively does not allow for choice probabilities that are either 0 or 1. If that was the case, for example, one would have to accept the model's unrealistic prediction that whenever –as is intuitively expected to hold– the probability of choosing \$5 over \$0 and \$5000 over \$0 are both equal to 1, then $u(\$5) = u(\$5000)$[6]. This difficulty is circumvented by employing the *Positivity* axiom that assumes strictly positive probabilities throughout. As Baccelli and Mongin (2016) recently pointed out, however, the conceptual concern that emerges now is that such genuine choice stochasticity rules out the possibility of the decision maker behaving like an ordinal utility maximizer, and yet the model portrays him as one whose behaviour is describable by a utility function that has even more refined uniqueness properties.

Non-cardinally-unique utility differences

1. In contrast to neo-classical and cardinal utility from stochastic-choice data, basic utility-difference representations are applicable on finite domains in non-trivial ways. While this is clearly a virtue, an important drawback of this model is that its general behavioural content is unclear. Specifically, although sufficient conditions for a preference intensity relation on a finite set to be represented in this way were laid out in Scott and Suppes (1958) and complete characterizations

[5]Krantz, Luce, Suppes, and Tversky (1971; Theorem 5, p. 168).

[6]This is a simple variation of the example in Davidson and Marschak (1959, p. 237) that was phrased in terms of equal utility differences.

were independently given in Scott (1964) and Adams (1965), these characterizations involve a complicated Cancellation axiom. Scott's version of the axiom reads as follows:

Cancellation
For all $(a_1, b_1), \ldots, (a_n, b_n)$ *in* $X \times X$ *and all permutations* π, σ *of* $\{1, \ldots, n\}$, *if* $(a_i, b_i) \gtrsim (a_{\pi(i)}, b_{\sigma(i)})$ *for all* $1 < i \leqslant n$, *then* $(a_{\pi(1)}, b_{\sigma(1)}) \gtrsim (a_1, b_1)$.

Commenting on it, Luce and Suppes (1965, p. 277) wrote: *"The difficulty of this axiom from a psychological standpoint is that there seems to be no simple way of summarizing what is says about choice behaviour, but this we take to be an inherent complexity of the structural relations that must hold between elements of any finite set in order to guarantee the existence of a utility function that preserves the order of utility differences."*

We also note that, in order to cover finite sets of arbitrary cardinalities, the number n capturing the width of the relevant sequences in the statement of the Cancellation axiom is unbounded. As such, the utility-difference model is not finitely axiomatizable (Scott and Suppes, 1958). As shown in Fishburn (2001), however, for a fixed finite set X n is bounded above by the cardinality of X by $2 \cdot |X| - 1$.

2. Samuelson (1938) observed that utility-difference representations –cardinal or otherwise– necessitate the following condition on the agent's preference intensity relation:

Concatenation
If $(a, b) \gtrsim (a', b')$ *and* $(b, c) \gtrsim (b', c')$, *then* $(a, c) \gtrsim (a', c')$.

From the point of view of physical distance measurement between earthly objects lying on a straight line, Concatenation is clearly a descriptively relevant property. It requires, for example, that if the length between such points a and b equals that between a' and b', and the length between b and c equals that between b' and c', then a is distanced from c exactly as much as a' is from c'. Its appeal in the context of preference intensity modelling is more questionable, however. Samuelson himself challenged its relevance by noting (p. 70) that *"there is absolutely no a priori reason why the individual's* [preference intensity relation] *should obey this arbitrary restriction"*. As will be formally shown below through an example, Concatenation is a necessary condition in the context of that model in order to ensure transitivity of utility differences. Crucially, however, transitivity of utility differences is *not* implied by transitivity of the preference intensity relation. Taken together, this point and the preceding remark about the uninterpretability of the Cancellation axiom in turn suggest that the utility-difference model comes with more "baggage" than perhaps is necessary for the problem at hand.

3. If a preference intensity relation $\gtrsim$ on a finite set X admits a utility-difference representation by means of some non-cardinally unique function u, then u has the uniqueness property of additive utility representations on finite sets[7]. This uniqueness property is between cardinal

[7]This is formally pinned down in Krantz, Luce, Suppes, and Tversky (1971; Theorem 2, p. 431).

and ordinal[8], but with an intuitive interpretation of it being elusive. Importantly, however, since –by lack of cardinal uniqueness– the ratio of utility differences $\frac{u(a)-u(b)}{u(c)-u(d)}$ is no longer invariant with respect to the model's permissible transformations, there is no hope of the utility function possibly acting as a unit of preference-intensity measurement. But if precise measurement must be given up, does that not invite the development of simpler and more transparent models?

The alternative modelling approach that is laid out in the next section is informed by –and reflects– the preceding critical discussion of both utility-difference approaches. In summary, the proposed model:

(i) offers the first genuinely ordinal representation of a preference intensity relation, and does so without dropping any of the so-called rationality conditions, either on that relation or on the ordinary preferences induced by it;

(ii) is axiomatized on a finite set of general alternatives by means of simple and behaviourally interpretable axioms that do *not* imply Cancellation, or even Concatenation;

(iii) includes the utility-difference model as a special case in a way that is made precise[9].

2.2 Preference Intensity Functions

From now on the analysis will revolve around the following concept:

Definition 2
A binary relation $\succsim$ on a set $X \times X$ is representable by a preference intensity function *if there exists a mapping $s : X \times X \to \mathbb{R}$ such that, for all $a, b, c, d \in X$,*

$$(a,b) \succsim (c,d) \iff s(a,b) \geqslant s(c,d) \tag{3a}$$
$$s(a,b) = -s(b,a) \tag{3b}$$
$$s(a,b), s(b,c) \geqslant 0 \implies s(a,c) \geqslant s(a,b), s(b,c), \tag{3c}$$

where s is unique up to an odd[10] and strictly increasing transformation *in the sense that $t : X \times X \to \mathbb{R}$ also represents $\succsim$ as in (3) if and only if $t = f \circ s$ for some $f : \mathbb{R} \to \mathbb{R}$ that is odd and strictly increasing in $s(X \times X)$.*

[8]Specifically, all positive affine transformations of u also represent $\succsim$, but not all strictly increasing transformations do so. The set of permissible transformations of u therefore lies between these two extreme polar cases in a $\succsim$-dependent way. There is a conceptual analogy between this and the set of transformations that preserve decreasing marginal utility, which too was shown in Mandler (2006) to lie between these two extreme cases.

[9]An alternative approach to preference-intensity modelling is to take as primitive a *semi-order* P on X that is representable by a pair (u, δ) where $u : X \to \mathbb{R}$ and $\delta > 0$ are such that $aPb \Leftrightarrow u(a) - u(b) > \delta$ (Luce, 1956; Scott and Suppes, 1958; Scott, 1964; Fishburn, 1970b). A semi-order is a special case of a strict partial order that features an incomplete strict preference relation and an incomparability relation, with the latter typically being interpreted as an intransitive indifference relation. The uniqueness properties of such representations are also between ordinal and cardinal. Although the primitive of this model is a transitive binary relation on X rather than on $X \times X$, the differences in the values of u that are featured in the representation can also be used to define a binary relation on $X \times X$ by $(a,b)\widehat{P}(c,d) \Leftrightarrow u(a) - u(b) > u(c) - u(d) > \delta$. While this foundational decision-theoretic model has also proved fruitful in important welfare-theoretic applications (e.g. axiomatizations of weighted utilitarianism (Ng, 1975; Argenziano and Gilboa, 2019)) where its preference-intensity interpretation has been used, unlike utility-difference representations or those proposed below, it assumes that the decision maker's preference and preference-intensity comparisons are incomplete under one interpretation or complete and intransitive under another.

[10]A function $f : \mathbb{R} \to \mathbb{R}$ is odd in $A \subseteq \mathbb{R}$ if $f(-z) = -f(z)$ holds for all $z \in A$.

As captured by the order-preservation requirement (3a), a preference intensity function's first property is to represent a binary relation on the set of *pairs* of alternatives in exactly the same way that an *ordinal* utility function represents a binary relation on the set of *alternatives*. This analogy also motivates its bivariate nature. Under (3a), the *skew-symmetry* condition (3b) means that one need only look at the sign of $s(a, b)$ to infer whether the agent strictly prefers a to b (positive sign), if the opposite is true (negative sign), or if he is indifferent between a and b (zero value). Therefore, a preference intensity function that represents such a relation also represents the ordinary preferences induced by it in the sense of (1). Moreover, in the special case where the preference intensity relation also admits a utility-difference representation by means of a function u, skew-symmetry allows s to be defined by $s(a, b) := u(a) - u(b)$. Apart from these implications of the particularly convenient normalization that is afforded by skew-symmetry, it will be formally established below that this condition is otherwise void of behavioural content.

The last defining property of preference intensity functions, (3c), will be referred to as *lateral consistency*. Under the maintained assumption that (3a) and (3b) are in place, it is interpretable as requiring that if a is weakly preferred to b and b to c, then a is weakly preferred to c at least as much as a is to b and b is to c. Lateral consistency therefore imposes the intuitive restriction that the preference and preference-intensity comparisons are in conceptual harmony in the sense that as the decision maker goes down his preference ranking from a to b and from b to c, his preference intensity between the "remote" alternatives a and c in this ranking is higher than that between the "proximal" alternatives a, b and b, c. In addition, it ensures that the preferences induced by the preference intensity relation represented by s are transitive, which is not implied by (3a) and (3b) alone.[11]

Finally, preference intensity functions are essentially ordinally unique, with strictly increasing transformations also required to be odd only in order to preserve the normalization offered by skew-symmetry. We will return to this point shortly.

The next concept helps towards clarifying that the *values* and *not* the value differences of a preference intensity function convey all the relevant information about the underlying comparisons.

Definition 3
A binary relation $\succsim$ *on a finite set* $X \times X$ *is representable by a* canonical *preference intensity function* $s : X \times X \to \mathbb{R}$ *if, in addition to satisfying* (3)*, its range* $s(X \times X)$ *is a symmetric set of consecutive integers* $\{-k, \ldots, -1, 0, 1, \ldots, k\}$.

As is shown below, in addition to their clarifying role, canonical representations allow for meaningful *ordinal* interpersonal comparisons of preference intensities where the inequality $s_i(a, b) > s_j(a, b) > 0$ for two agents i, j and alternatives a, b is legitimately interpretable as suggesting that i prefers a to b more than j does. Specifically, assuming in such a context that all agents' intensity relations are representable by canonical preference intensity functions that have the same range (for example, by assuming that all agents' preference intensity relations are strict), interpersonal comparisons in the values of these functions could be thought of as constituting the

[11] Formally, (3a) and (3b) imply that, for all $a, b \in X$, $a \gtrapprox b \Leftrightarrow (a, b) \succsim (b, a) \Leftrightarrow s(a, b) \geqslant 0 \geqslant s(b, a)$, while it then follows from this and (3c) that $a \gtrapprox b \gtrapprox c \Rightarrow s(a, c) \geqslant s(c, a) \Leftrightarrow (a, c) \succsim (c, a) \Leftrightarrow a \gtrapprox c$.

ordinal analogue to the *relative* utilitarianism assumption that rests on interpersonal comparisons of normalized von Neumann-Morgenstern utility functions whose range is the unit interval[12]. It is worth stressing, in particular, that although canonical preference intensity representations can deliver an analogously unique and well-defined normalization that allows for interpersonal comparisons of preference intensities, it follows from well-known results [see Bossert and Weymark (2004) and Fleurbaey and Hammond (2004)] that this is not the case with non-cardinal utility-difference representations. We elaborate on these points in Section 4.

Towards establishing the behavioural irrelevance of both skew-symmetry and odd transformations in this model, the following class of functions is now introduced where both these properties are relaxed and lateral consistency is slightly modified.

Definition 4
A binary relation $\succsim$ on a set $X \times X$ is representable by a general *preference intensity function if there exists a mapping $g : X \times X \to \mathbb{R}$, unique up to a strictly increasing transformation, such that, for all $a, b, c, d \in X$,*

$$(a, b) \succsim (c, d) \iff g(a, b) \geqslant g(c, d) \tag{4a}$$

$$g(a, b) \geqslant g(c, d) \implies g(d, c) \geqslant g(b, a) \tag{4b}$$

$$g(a, b), g(b, c) \geqslant g(a, a) \implies g(a, c) \geqslant g(a, b), g(b, c). \tag{4c}$$

For a function g that satisfies (4) but not (3), skew-symmetry is replaced by a weaker condition that retains the same ordering restriction without imposing a sign requirement on the function's values. As a consequence, $a \gtrapprox b \Leftrightarrow s(a, b) \geqslant 0$ in (3) becomes $a \gtrapprox b \Leftrightarrow s(a, b) \geqslant s(b, a)$ in (4). Consistent with this observation, the comparison of (3c) and (4c) suggests that the real "zero" in a general preference intensity function is its value at any point on the diagonal of $X \times X$ where, by definition, there is zero preference intensity difference in moving from the second point in the pair to itself. Finally, without a need for skew-symmetry to be accounted for in intensity-preserving transformations of a function satisfying (4), the uniqueness property in the class of general preference intensity functions coincides with that derived by standard ordinal transformations.

We now turn to the axioms that will be imposed on the relation $\succsim$.

Weak Order
For all $a, b, c, d \in X$, $(a, b) \succsim (c, d)$ or $(c, d) \succsim (a, b)$.
For all $a, b, c, d, e, f \in X$, $(a, b) \succsim (c, d) \succsim (e, f)$ implies $(a, b) \succsim (e, f)$.

Weak Order requires the decision maker to be able to make preference comparisons universally and consistently. While both these standard assumptions are known to be challenged descriptively, especially as complexity of the decision task increases, they are retained here in order for an alternative baseline model to be developed and compared to that of utility differences (which imposes these

[12]See, for example, Dhillon and Mertens (1999). Hammond (1991) traces the origins of this assumption in multi-agent environments in Isbell (1959).

axioms too). It is envisaged that both parts of the Weak Order axiom will be relaxed in future work.

Reversal
For all $a, b, c, d \in X$, $(a, b) \succsim (c, d)$ implies $(d, c) \succsim (b, a)$.

Reversal is also a standard condition and allows for the relation $\succsim$ to be interpreted as a preference intensity relation by requiring that whenever the transition from b to a is more desirable than that from d to c, then the transition from c to d also be preferable to that from a to b. If a is preferred to b more than c is to d, for example, then since the intensity difference between c and d is smaller than that between a and b, the transition from c to d should be associated with a smaller psychological cost than that from a to b, and therefore, intuitively, $(d, c) \succsim (b, a)$ should hold.

Consistency
For all $a, b, c \in X$, $(a, c) \succsim (b, c)$ implies $(a, b) \succsim (b, a)$.

Consistency requires that whenever the agent prefers the transition from c to a more than that from c to b, then he also prefers a to b. As is shown below, it turns out that, under Weak Order and Reversal, Consistency is equivalent to the familiar

Separability
For all $a, b, c, d \in X$, $(a, c) \succsim (b, c)$ implies $(a, d) \succsim (b, d)$.

We note that, alongside Weak Order, Reversal and several additional axioms, a stronger version of Separability appears in the first axiomatization of (cardinal, neoclassical) utility-difference models that was given in Alt (1936, 1971).

Theorem 1
The following are equivalent for a binary relation $\succsim$ on a finite set $X \times X$:

1. *$\succsim$ satisfies Weak Order, Reversal and Consistency.*
2. *$\succsim$ satisfies Weak Order, Reversal and Separability.*
3. *$\succsim$ is representable by a unique canonical preference intensity function.*
4. *$\succsim$ is representable by an odd-ordinally unique preference intensity function.*
5. *$\succsim$ is representable by an ordinally unique general preference intensity function.*

The example below illustrates Theorem 1 and the power of preference intensity functions by presenting two relations that are preference-intensity but not utility-difference representable, due to failures of Concatenation –hence Cancellation– in the first case and of Cancellation alone in the second case.

Example 1

Suppose that $X = \{a, b, c, d\}$ and consider the intensity relations $\succsim_1$[13] and $\succsim_2$ on X such that

$$\begin{array}{ccccccccccc} (a,d) & \succ_1 & (b,d) & \succ_1 & (a,c) & \succ_1 & (a,b) & \succ_1 & (b,c) & \succ_1 & (c,d) \\ (a,d) & \succ_2 & (b,d) & \succ_2 & (a,c) & \succ_2 & (b,c) & \succ_2 & (a,b) & \succ_2 & (c,d) \end{array}$$

Moreover, as per the Reversal axiom, assume that $(a', b') \succsim_i (c', d')$ implies $(d', c') \succsim_i (b', a')$ for $i = 1, 2$. This, in particular, means that the pair (c, d) in both orderings is followed by all pairs (z, z) that lie on the diagonal of $X \times X$, which are then followed by the pair (d, c), etc. Thus extended, both relations satisfy Weak Order, Reversal and Consistency/Separability.

Suppose to the contrary that there are $u_1, u_2 : X \to \mathbb{R}$ that represent $\succsim_1$ and $\succsim_2$, respectively, as in (2). In the first case we have

$$\begin{array}{ccccccc} (a,b) & \succ_1 & (b,c) & \succ_1 & (c,d) & \iff \\ u_1(a) - u_1(b) & > & u_1(b) - u_1(c) & > & u_1(c) - u_1(d) & \iff \\ & & u_1(a) - u_1(c) & > & u_1(b) - u_1(d) & \iff \\ & & (a,c) & \succ_1 & (b,d), & \end{array}$$

which contradicts the postulate $(b, d) \succ_1 (a, c)$. This shows that $\succsim_1$ violates Concatenation, and that violations of Concatenation lead to intransitive utility differences even though the relation $\succsim_1$ is actually transitive.

In the second case on the other hand, it follows from $(a, b) \succ_2 (c, d)$ and $(b, d) \succ_2 (a, c)$ that

$$\begin{array}{ccc} u_2(a) - u_2(b) & > & u_2(c) - u_2(d) \\ u_2(a) - u_2(c) & < & u_2(b) - u_2(d), \end{array}$$

and subtracting the second inequality from the first yields

$$u_2(c) - u_2(b) \quad > \quad u_2(c) - u_2(b).$$

This contradiction stems from a different violation of the Cancellation axiom that, unlike the one above, does not admit a straightforward behavioural interpretation. Indeed, Cancellation together with the following comparisons (rephrased in the permutation language of the Cancellation axiom) in $\succsim_2$

$$\begin{array}{ccccccc} (a_2, b_2) & := & (a,b) & \succ_2 & (c,d) & := & (a_{\pi(2)}, b_{\sigma(2)}) \\ (a_3, b_3) & := & (b,d) & \succ_2 & (a,c) & := & (a_{\pi(3)}, b_{\sigma(3)}) \\ (a_4, b_4) & := & (c,a) & \succ_2 & (d,b) & := & (a_{\pi(4)}, b_{\sigma(4)}) \end{array}$$

imply

$$(a_1, b_1) \;:=\; (d,c) \;\precsim_2\; (b,a) \;:=\; (a_{\pi(1)}, b_{\sigma(1)}),$$

which contradicts the $(d, c) \succ_2 (b, a)$ postulate.

Notice, finally, that despite the failure of the utility-difference model to account for these perfectly

[13] This relation example is also presented in Köbberling (2006).

plausible intensity orderings, both are representable by a canonical preference intensity function that sets $s_i(a,d) = 6$, $s_i(b,d) = 5$, …, $s_i(c,d) = 1$ and $s_i(a',b') = -s_i(b',a')$ for $i = 1,2$, and with their only difference being their values at the pairs (a,b), (b,c) and (b,a), (c,b). ◊

Theorem 1 offers a characterization of preference intensity functions by means of standard, easily interpretable and collectively weak behavioural conditions, which, as shown in Example 1, do not imply the demanding and uninterpretable Cancellation axiom –or the simpler but still quite demanding Concatenation axiom– that is necessary in the utility-difference model. In addition, it establishes that the skew-symmetric and non-skew-symmetric versions of the model are in fact formally equivalent. Therefore, the convenient normalization offered by skew-symmetry –the version of the model that will be used in the sequel– is without loss of generality. Also without loss, finally, is to assume a canonical representation whenever such an assumption is useful in some application of the model such as the welfare-theoretic one that we purse in Section 4.

The next concept will be helpful in clarifying the relationship between preference-intensity and utility-difference representations.

Definition 5
A binary relation $\succsim$ on a set $X \times X$ is representable by a triangularly additive *preference intensity function $s : X \times X \to \mathbb{R}$ if, for all $a, b, c \in X$,*

$$s(a,c) = s(a,b) + s(b,c). \tag{5}$$

Triangular additivity is generally not satisfied by preference intensity functions, and both sub-additive and super-additive deviations generally occur within the context of the *same* such representation.[14] When this condition *is* satisfied, however, it implies both (3b) and (3c). Triangularly additive preference intensity functions are therefore characterized by (3a) and (5) only. This concept –unnamed, and accompanied by a critical discussion– first appeared in Samuelson (1938) in his bivariate reformulation of neoclassical cardinal utility functions.

Corollary 1
The following are equivalent for a binary relation $\succsim$ on a finite set $X \times X$:

1. *$\succsim$ satisfies Completeness, Reversal and Cancellation.*
2. *$\succsim$ is utility-difference representable.*
3. *$\succsim$ is representable by a triangularly additive preference intensity function.*

The equivalence between the first two statements in this corollary is due to Scott (1964, Theorem 3.2).[15] The equivalence between the latter two demonstrates that the proposed model nests the utility-difference model whenever the underlying preference intensity order is representable by a preference intensity function s that takes the special *additively separable* form $s(a,b) \equiv u(a) - u(b)$

[14] In Example 1, for instance, $s(a,b) + s(b,c) > s(a,c)$ and $s(b,c) + s(c,d) < s(b,d)$.
[15] A similar characterization –developed independently– also appears in Adams (1965, Theorem 1).

for some function $u : X \to \mathbb{R}$.[16] Although this second equivalence is intimately related to the one provided by Samuelson (1938) in a cardinal framework, it must be emphasized that Samuelson (1938) did not suggest using the bivariate approach as a potentially more general way to represent preference intensity relations. In particular, despite his critical approach towards what we are referring to as triangularly additive preference intensity functions, Samuelson (1938) did not suggest a way of relaxing this property in order to alleviate the concerns that he and some of his contemporaries had raised about the cardinal utility model, or, with hindsight, the utility-difference approach more generally. Ours appears to be the first study in the literature of preference intensity modelling that takes this stand and relaxes triangular additivity with the far less demanding but still sufficiently structured lateral consistency condition.

Table 1: Domain gains of preference-intensity representations relative to utility-difference ones

	A *(utility-difference model)*	*B*	*C* *(preference-intensity model)*	
	Weak Order Reversal Cancellation	Weak Order Reversal Concatenation	**Weak Order Reversal Consistency**	*domain gain in C vs A*
$\lvert X\rvert = 3$	25	25	37	**48%**
$\lvert X\rvert = 4$	723	1,011	3,903	**439%**
$\lvert X\rvert = 5$	63,721	210,361	5,230,801	**8,108%**

Table 1 demonstrates the enormous potential gains in explanatory power that is offered by the proposed preference intensity model relative to the utility-difference one by comparing the total number of distinct relations that satisfy Weak Order, Reversal and Consistency with the number of relations that, in addition, satisfy Cancellation[17]. While computational constraints currently limit these comparisons to the cases where X contains up to five alternatives, this novel output shows that the explanatory gains of the proposed model increase super-exponentially in the cardinality of X within this range. Strikingly, when X has five elements, the proportion of intensity relations that are consistent with Weak Order and Reversal and also satisfy Cancellation (respectively, Concatenation) is a mere 1.22% (respectively, 4%) of the intensity relations that satisfy the former two axioms together with Consistency (or, equivalently, Separability). This fact provides a formal vindication of Samuelson's (1938) criticism of triangular additivity and its Concatenation implication as *"arbitrary"* and *"infinitely improbable"*. This huge gap between the number of complete and internally consistent intensity rankings –however elicited– that can be accommodated by the two models in turn translates into enormous applicability-domain gains of the revealed preference intensity and welfare-theoretic applications of preference intensity functions that are provided in Sections 3 and 4 below.

2.3 Special Case: Random Intensity Relations from Random Binary Choices

The lateral consistency condition (3c) of preference intensity functions is closely related to what is known in the random choice literature as *Strong Stochastic Transitivity* (SST). Before stating this

[16] An additional output of Theorem 1 and Corollary 1 is that they suggest a computational method towards testing whether a relation that is representable by a preference intensity function is also utility-difference representable (hence that it also satisfies the challenging Cancellation axiom). The general idea of the algorithm would be to start with a canonical preference intensity representation of that relation and, if necessary, change its values until no triple of alternatives exists where triangular additivity is violated.

[17] Details on how the entries of Table 1 were computed are available on request.

condition, let us recall that, given the collection $\mathcal{B}_X$ of all binary menus on a set X, a *binary random choice model* on X is a function $p : X \times \mathcal{B}_X \to [0,1]$ such that, for all $a, b \in X$, $p(a, \{a,b\}) + p(b, \{a,b\}) = 1$. Such a model p is said to satisfy SST if $p(a, \{a,b\}) \geqslant \frac{1}{2}$ and $p(b, \{b,c\}) \geqslant \frac{1}{2}$ implies $p(a, \{a,c\}) \geqslant p(a, \{a,b\}), p(b, \{b,c\})$ (we will say that it satisfies SST* if the last inequality is strict whenever one of the first two is also strict). Therefore, interpreting $p(a, \{a,b\}) \geqslant \frac{1}{2}$ as suggesting that the decision maker –whether a single person or, as is sometimes the case in practice, the average in a sample– *probably* prefers a to b, SST requires that a is probably preferred to c more than a is to b and b is to c. This interpretation, in particular, is in line with Debreu's (1958) thesis that stochastic choice data are indicative of differences in preference intensities between alternatives, and hence that one could think of the comparison $p(a, \{a,b\}) \geqslant p(c, \{c,d\})$ as suggesting that a is *probably* preferred to b more than c is to d. More formally, we can define the *random preference intensity relation* $\succsim$ on the set X that is induced by a binary random choice model p on that set *à la* Debreu (1958) by

$$(a,b) \succsim (c,d) \Longleftrightarrow p(a, \{a,b\}) \geqslant p(c, \{c,d\}).$$

As was also noted at the beginning of Section 2.1, such a $\succsim$ is a special case of a preference intensity relation on the set X that is simply defined in terms of differences between real numbers that correspond to choice probabilities.

Moreover, as was also noted previously, Debreu (1958) imposed sufficient structure on the –infinite– set X and the binary random choice model p for the relation $\succsim$ induced by it to admit a cardinal utility difference representation in the sense of (2). The more general class of representations of binary random choice models where the utility function is not necessarily cardinally unique and the set X can be finite without any resulting trivialities[18] is the class of *Fechnerian representations* (Falmagne, 2002). These representations postulate the existence of some $u : X \to \mathbb{R}$ and a strictly increasing $F : \mathbb{R} \to \mathbb{R}$ such that

$$p(a, \{a,b\}) = F\big(u(a) - u(b)\big). \tag{6}$$

A more general class of binary random choice models are known as *scalable* and postulate instead the existence of some $u : X \to \mathbb{R}$ and a function $F : \mathbb{R}^2 \to \mathbb{R}$ that is strictly increasing (decreasing) in its first (second) argument and satisfies

$$p(a, \{a,b\}) = F\big(u(a), u(b)\big). \tag{7}$$

The monotonicity properties of this F together with the fact that the pair (u, F) is no longer required to preserve utility differences as in (6) immediately establishes the greater generality of scalable relative to Fechnerian models.

The class of scalable binary random choice models on a finite set X was axiomatically characterized by Tversky and Russo (1969) by means of SST* and the usually assumed *Positivity* axiom that requires all choice probabilities to be strictly positive. Fechnerian models on the other hand were recently axiomatized by Fudenberg, Iijima, and Strzalecki (2015) by means of Positivity and a

[18]This may be thought of as the stochastic analog to the general model of non-cardinal utility differences that was discussed earlier.

Cancellation-like *Acyclicity* axiom that we will refer to as *FIS-Acyclicity*[19]. Using the random preference intensity relation that is induced by a binary random choice model p, our preceding analysis allows for a connection to be made between preference intensity functions and Fechnerian or scalable representations that invite simple novel interpretations of these models.

Corollary 2
The following are equivalent for a binary random choice model p on a finite set X:

1. p satisfies Positivity and Strong Stochastic Transitivity.*

2. p is scalable.

3. The p-induced $\succsim$ is representable by a preference intensity function.

Corollary 3
The following are equivalent for a binary random choice model p on a finite set X:

1. p satisfies Positivity and FIS-Acyclicity.

2. p is Fechnerian.

3. The p-induced $\succsim$ is representable by a triangularly additive preference intensity function.

As also suggested by the above discussion, the novel part in both Corollaries 2 and 3 is the equivalence between the second and third statements. In the latter case, a formal equivalence is established between FIS-Acyclicity on p and Cancellation on the p-induced relation $\succsim$. In the former case, and recalling also the previous discussion, the equivalence suggests that, by virtue of the SST* property, scalable models can be thought of as having a sufficiently strong structure to ensure transitivity of –and conceptual harmony between– the random intensity relation and the random preference relation that is induced by it and, in particular, that these properties are lost by $\succsim$ if the model p that induces it violates SST*.

We finally note that –under the maintained assumption of Positivity– Tversky and Russo (1969) also established the equivalence between SST* and the random-choice analogs of the Consistency and Separability axioms (called *Substitutability* and *Independence* in that paper, respectively). The first two statements in Theorem 1 extend this equivalence to more general environments where the intensity relation in question is not necessarily defined by differences in choice probabilities.

3 Revealed Preference Intensity

Behavioural datasets arising from two-alternative forced-choice experiments are very common in economics, psychology and neuroscience. In many of those experiments, an additional variable is also elicited alongside the decision maker's choice at each menu. Examples of such observables include:

1. *Questionnaires and Likert-scale ratings.* The decision maker can be asked to indicate the degree of preference for the chosen over the non-chosen alternative by selecting a desirability rating on

[19] Fudenberg, Iijima, and Strzalecki (2015) also dealt with the case where X is infinite.

an arbitrarily fine scale that may in turn be divided into broad categories that are suggestive of preference intensity such as "slightly better", "better", "much better", "very much better", as in Butler, Isoni, Loomes, and Tsutsui (2014), for example.[20] Importantly, this is not an absolute rating that conveys information about the desirability of the chosen alternative in some scale[21]; instead it is a relative rating that provides information about the strength of preference of the chosen alternative relative to the rejected one.

2. *Willingness to pay.* The decision maker can be asked to state the amount he would be willing to spend in order to receive the alternative he chose at each menu, as envisaged in Luce and Suppes (1965) for the case where the agent is endowed with b and d and is willing to change them for a and c, for example. Alternatively, as in Butler, Isoni, Loomes, and Navarro-Martinez (2014) for the case of money lotteries, the individual could be asked to indicate how much the rejected alternative needs to be improved in order to become as attractive as the chosen one. In both cases, a menu-specific monetary value is elicited that provides information about the intensity of preference between the alternatives in that menu.

3. *Response times.* Originating in psychology and neuroscience, of increasing interest to economists in recent years is the *drift diffusion* stochastic choice model and its generalizations (Ratcliff, 1978; Ratcliff and McKoon, 2008; Krajbich, Oud, and Fehr, 2014; Clithero, 2018; Konovalov and Krajbich, 2017; Cerreia-Vioglio, Maccheroni, Marinacci, and Rustichini, 2018; Alós-Ferrer, Fehr, and Netzer, 2018). This model postulates the existence of cardinal utility values for all alternatives, which are discovered by the decision maker (possibly with error) via sequential sampling. A key prediction of that model is that there is a negative relationship between the decision maker's response time at a menu and the utility difference between the alternatives at that menu.

In each of these cases, the analyst has access to choice data and to an additional, menu-specific variable or foregone resource that has intensity-revealing potential. The first study that analyzed such extended data in the revealed-preference tradition was Echenique and Saito (2017). Assuming that the additional observable is response times, these authors proposed a notion of rationalizability that built on the non-cardinal utility-difference model in (2), and identified a testable axiom –called *Strong Compensation*– on such data that is necessary and sufficient for them to be *utility-difference rationalizable* in that sense. In particular, the authors' proposed rationalization postulates the existence of a utility function $u : X \to \mathbb{R}$ and a strictly decreasing function $f : \mathbb{R} \to \mathbb{R}$ such that

$$C(\{a,b\}) = a \implies u(a) > u(b) \tag{8a}$$

$$u(a) - u(b) = f\big(t(a,b)\big), \tag{8b}$$

where $C(\{a,b\}) = a$ means that a is chosen over b at menu $\{a,b\}$ and $t(a,b)$ is the observed response time for this choice. Despite this novel application of the utility-difference approach to intensity modelling, however, as the analysis in the previous section –and Table 1 in particular– suggests,

[20] In non-binary choice environments, Abdellaoui, Barrios, and Wakker (2007) also elicited riskless utility over money through strength-of-preference statements.

[21] Such absolute ratings are often elicited in empirical and experimental studies, but they effectively force the decision make to form a pseudo-cardinal ranking of the alternatives directly.

the existence of a utility-difference representation that lies at the heart of the Echenique and Saito (2017) rationalization is a particularly restrictive requirement and rules out a wide range of perfectly consistent behaviour, both in terms of the choices alone and in terms of the choices and response times taken together.

In response to this limitation, and also to highlight the greater generality and explanatory power of the purely ordinal model of preference intensities that was laid out in (3), we will use this model to generalize the analysis in Echenique and Saito (2017) by considerably relaxing the restrictions imposed on the data by means of a simple *Congruence*-like Richter (1966) axiom (which, in particular, motivates a purely constructive approach), and also by introducing a more general –purely on interpretational grounds– definition of a dataset than Echenique and Saito (2017) that allows for the additional observable to be any relevant menu-specific variable.

Definition 6
A binary behavioural dataset $\mathcal{D} = \{\{a_i, b_i\}, C(\{a_i, b_i\}), r_{\{a_i,b_i\}}\}_{i=1}^{k}$ *on a finite set* X *is a collection of triples consisting of a binary menu* $\{a_i, b_i\} \in \mathcal{B}_X$*, the observed choice* $C(\{a_i, b_i\})$ *at that menu, and the value* $r_{\{a_i,b_i\}}$ *of an observable and menu-specific intensity-revealing variable/resource.*

We will now use preference intensity functions to introduce the following concept of rationalizability that disciplines this class of datasets while at the same time encompassing Echenique and Saito (2017) utility-difference rationalizability as a special case. From now on, we will use the notation $a \ggcurly^B b$ whenever $C(\{a, b\}) = \{a\}$ for some menu $\{a, b\}$ in $\mathcal{D}$. That is, $\ggcurly^B$ denotes the (binary) revealed preference relation.

Definition 7
A binary behavioural dataset $\mathcal{D}$ *on a finite set* X *is* preference-intensity rationalizable *if there exist functions* $s : X \times X \to \mathbb{R}$ *and* $f : \mathbb{R} \to \mathbb{R}$*, with* f *strictly monotonic, such that, for all* $a, b, c \in X$*,*

$$a \ggcurly^B b \quad \Longrightarrow \quad s(a, b) > 0 \tag{9a}$$
$$s(a, b) \quad = f\left(r_{\{a,b\}}\right) = \quad -s(b, a) \tag{9b}$$
$$s(a, b), s(b, c) > 0 \quad \Longrightarrow \quad s(a, c) \geqslant s(a, b), s(b, c) \tag{9c}$$

In words, $\mathcal{D}$ is rationalizable in the above sense if: (i) a preference intensity function s can be constructed so that each pair of alternatives (a, b) where a is revealed preferred to b is associated with a strictly positive value; (ii) this value is itself a strictly monotonic function of the intensity-revealing resource at $\{a, b\}$; and (iii) s also satisfies a (slightly weaker) version of the lateral consistency requirement that still ensures conceptual harmony between –and within– revealed preferences and revealed preference intensities by requiring both to be acyclic.

Turning to the model's empirical content, the following axiom –laid out in two versions, and with the appropriate one for the given task being left to the analyst to determine– combines Richter's (1966) Congruence axiom in the present binary-choice setting with a weak and intuitive monotonicity requirement.

Congruent Monotonicity (Negative)
If $a_1 \ggg^B a_2 \cdots \ggg^B a_n$ *and* $a_i \ggg^B a_{i+h}$ *for some* $i, i+h \leqslant n$ *with* $h > 1$*, then* $a_n \not\ggg^B a_1$ *and* $r_{\{a_i,a_{i+h}\}} \leqslant r_{\{a_i,a_{i+1}\}}$*.*

Congruent Monotonicity (Positive)
If $a_1 \ggg^B a_2 \cdots \ggg^B a_n$ *and* $a_i \ggg^B a_{i+h}$ *for some* $i, i+h \leqslant n$ *with* $h > 1$*, then* $a_n \not\ggg^B a_1$ *and* $r_{\{a_i,a_{i+h}\}} \geqslant r_{\{a_i,a_{i+1}\}}$*.*

The monotonicity part of the axiom predicts that the value of the intensity-revealing resource is monotonically increasing or decreasing for every sequence of alternatives where any two consecutive elements in the sequence are related by revealed preference. Specifically, if there is a sequence of alternatives with the revealed-preference structure postulated in the hypothesis part of the axiom, then $a_i \ggg^B a_{i+h}$ with $h > 1$ suggests that, with a_i and a_{i+h} being further apart from each other in the decision maker's preference ranking than a_i and a_{i+1} are, a_i is preferred to a_{i+h} more than a_i is to a_{i+1}. This is reflected in the way in which the resource values at these menus are required to be ordered. A dataset $\mathcal{D}$ will be said to satisfy Congruent Monotonicity if it satisfies either of the above two versions of the axiom, for all observations contained in it. It is worth noting that if $\mathcal{D}$ is complete in that it contains observations from all binary menus that can be derived from the underlying finite choice set, then the axiom is the conjunction of the Transitivity and Monotonicity axioms in Echenique and Saito (2017). While both these axioms are intuitive implications of the Strong Compensation axiom that characterizes rationalizability in the authors' utility-difference framework, there are many additional classes of restrictions that Strong Compensation inevitably needs to impose on the data in order to ensure a utility-difference representation that go beyond these two kinds of restrictions.[22]

Theorem 2
The following are equivalent for a binary behavioural dataset $\mathcal{D}$ *on a finite set* X*:*

1. $\mathcal{D}$ *is preference-intensity rationalizable.*

2. $\mathcal{D}$ *satisfies Congruent Monotonicity.*

Example 2

Suppose that $X = \{a, b, c, d\}$ and consider two behavioural datasets $\mathcal{D}_1$ and $\mathcal{D}_2$ with common binary choices and generally distinct resource values $r^1_{\{\cdot\}}$, $r^2_{\{\cdot\}}$ where, without loss of generality, lower values point to higher intensities (e.g. as in response times):

$$\begin{array}{ccccccccccc} a \ggg^B b, & & b \ggg^B c, & & c \ggg^B d, & & a \ggg^B c, & & a \ggg^B d, & & b \ggg^B d \\ r^1_{\{a,d\}} & < & r^1_{\{b,d\}} & < & r^1_{\{a,c\}} & < & r^1_{\{a,b\}} & < & r^1_{\{b,c\}} & < & r^1_{\{c,d\}} \\ r^2_{\{a,d\}} & < & r^2_{\{b,d\}} & < & r^2_{\{a,c\}} & < & r^2_{\{b,c\}} & < & r^2_{\{a,b\}} & < & r^2_{\{c,d\}}. \end{array}$$

Notice now that the orderings of the r^1s and r^2s are as in the preference intensity relations $\succsim_1$ and $\succsim_2$

[22]For, example, Concatenation – called Time-Transitivity in Echenique and Saito (2017)– as well as other distinct implications of Cancellation such as those in the second parts of Examples 1 and 2.

of Example 1. Both these relations induce the same preference ordering on X that coincides with the revealed preference relation $\ggcurly^B$. It is immediate, therefore, that both $\mathcal{D}_1$ and $\mathcal{D}_2$ satisfy Congruent Monotonicity and are preference-intensity rationalizable, e.g. by the pair (s_i, f_i) for $i = 1, 2$ such that $f_i\big(r^i_{\{a',b'\}}\big) = \frac{1}{r^i_{\{a',b'\}}}$ and $s_i(a', b') = \big(f_i(r^i_{\{a',b'\}})\big)^{-1}$ if $a' \ggcurly^B b'$ and $s_i(a', b') = -s_i(b', a')$ if $b' \ggcurly^B a'$. However, since the intensity orderings that are induced by the additional data $r^1_{\{\cdot\}}$ and $r^2_{\{\cdot\}}$ both violate Cancellation (and, in the first case, Concatenation too), it follows that they do not admit an Echenique-Saito rationalization, and hence that their Strong Compensation axiom fails. ◊

Corollary 4

The following are equivalent for a binary behavioural dataset $\mathcal{D}$ on a finite set X:

1. *$\mathcal{D}$ satisfies Strong Compensation.*
2. *$\mathcal{D}$ is utility-difference rationalizable.*
3. *$\mathcal{D}$ is triangularly additively preference-intensity rationalizable.*

Using the special triangular additivity property of preference intensity functions once again, the equivalence between the last two statements clarifies the sense in which preference-intensity rationalizability encompasses utility-difference rationalizability in the sense of Echenique and Saito (2017) as a special case.

4 Welfare-Theoretic Application: Intensity-Efficient Allocations

In this section we demonstrate the generality and applicability of the proposed model of preference intensity measurement from a welfare economics perspective. Consider a society with n agents and assume that there is a finite set X of general choice alternatives with n elements that must be allocated to these agents. A *preference intensity profile* $\succsim = (\succsim_1, \ldots, \succsim_n)$ on X is an ordered n-tuple of binary relations on $X \times X$, where each $\succsim_i$ is representable by a preference intensity function, hence satisfies Weak Order, Reversal and Consistency. The ordinary preference profile induced by $\succsim$ is denoted by $\gtrapprox$. An *allocation* is an n-tuple $x \in X^n$, where $x_i \in X$ corresponds to the alternative allocated to agent i by x. A preference intensity profile $\succsim$ is *strict* if $(a, b) \sim_i (c, d)$ implies $a = b$ and $c = d$ for every agent i and alternatives a, b, c and d. By Theorem 1, each $\succsim_i$ in $\succsim$ is representable by a canonical preference intensity function s_i. Moreover, if $\succsim$ is strict, then

$$s_1(X \times X) = \cdots = s_j(X \times X) = \{-k, -k+1, \ldots, 1, 0, 1, \ldots, k-1, k\},$$

where k here is the number of distinct pairs of distinct alternatives in X.

As previously noted, putting such agents' preference intensity functions in canonical form allows for the interpersonal comparability of their preference *intensities* without any requirement for interpersonal comparability of their *utilities*. In particular, this common normalization across agents allows for $s_i(a, b) > s_j(a, b) > 0$ to reveal that i prefers a to b more than j does because the pair (a, b) is now know to lie higher in agent i's intensity ordering than it does in agent j's. Importantly, even in the very special case where all agents' intensity relations are utility-difference representable, an analogous normalization of the agents' utility differences is generally impossible even if all agents' utility functions are normalized to have the same minimum and maximum value. Therefore, compared to utility-difference representations, (canonical) preference intensity ones offer a far greater domain of application as well as a simple and general way of achieving interpersonal comparisons of preference intensities that is impossible under the former approach.

Building on this intuitive input of canonical normalizations, we can now introduce the following novel notions of dominance and efficiency.

Definition 8
An allocation x intensity-dominates *another allocation y with respect to a strict intensity profile $\gtrsim = (\gtrsim_1, \ldots, \gtrsim_n)$ with canonical preference-intensity representation $s = (s_1, \ldots, s_n)$ if:*

(i) $s_i(x_i, x_j) \geqslant s_j(y_j, y_i)$ whenever $(x_i, x_j) = (y_j, y_i)$ for some pair of agents (i, j);

(ii) $s_l(x_l, y_l) \geqslant 0$ for every agent l not in such a pair;

(iii) at least one inequality in (i) or (ii) is strict.

An allocation x is intensity-efficient *with respect to $\gtrsim$ if it is not intensity-dominated.*

Allocation x intensity-dominates y if: (i) in every pair of agents that is "flipped" by x and y in the sense that both allocations assign the same two alternatives a and b to the two agents in that pair but do so in opposite ways, the agent receiving a under x (dis-)prefers[23] it to b more (less) than the agent who receives a under y; (ii) all agents who do not belong to such a pair are weakly better off at x than at y; and (iii) at least one interpersonal comparison in (i) or intrapersonal comparison in (ii) is strict. For two allocations x and y that are related in this way it holds that whenever ordinary Pareto dominance (i.e. $s_i(x_i, y_i) \geqslant 0$ for all i and strictly for some) does not apply, the interpersonal preference trade-offs in all pairs of agents that receive the same two alternatives under x and y but in reverse order are resolved in favour of the agent who prefers the alternative *more* than the other. Compactly, if $(x_i, x_j) = (y_j, y_i) = (a, b)$ and $s_i(a, b) > s_j(a, b)$, then, assuming (ii) is satisfied for all other agents, x intensity-dominates y.

Proposition 3
The following are true for a strict preference-intensity profile $\gtrsim = (\gtrsim_1, \ldots, \gtrsim_n)$ on a finite set X:

1. *An intensity-efficient allocation with respect to $\gtrsim$ exists.*
2. *If x is intensity-efficient with respect to $\gtrsim$, then x is Pareto efficient with respect to $\gtrapprox$.*

Unlike standard utilitarian notions of efficiency that assume cardinal utility-difference representations (typically –but not always– of the von Neumann-Morgenstern kind), the proposed notion of intensity efficiency is based on interpersonal comparisons of the agents' ordinal preference intensity rankings that are defined on *pairs* of alternatives, in line with the general approach of this paper. Proposition 3 further shows that this concept is well-defined and refines Pareto efficiency. Like the latter notion, however, and unlike utilitarian efficiency, it generally specifies an incomplete social ranking over the set of allocations for a given preference intensity profile. Yet, as the following

[23] By "a is *dis-preferred* to b" we mean that a is considered inferior to b.

example illustrates, the refinement that intensity efficiency offers over the set of Pareto efficient allocations by discarding ones that are intensity dominated can be quite substantial.

Example 3

Suppose that $X = \{a, b, c, d\}$ and consider the strict intensity profile $\succsim = (\succsim_1, \succsim_2, \succsim_3, \succsim_4)$ that is represented canonically by:

$$
\begin{array}{llll}
s_1(a,d) = 6 & s_2(d,a) = 6 & s_3(a,d) = 6 & s_4(d,a) = 6 \\
s_1(b,d) = 5 & s_2(d,c) = 5 & s_3(a,c) = 5 & s_4(c,a) = 5 \\
s_1(a,c) = 4 & s_2(d,b) = 4 & s_3(a,b) = 4 & s_4(d,b) = 4 \\
s_1(b,c) = 3 & s_2(c,a) = 3 & s_3(b,d) = 3 & s_4(c,b) = 3 \\
s_1(a,b) = 2 & s_2(c,b) = 2 & s_3(c,d) = 2 & s_4(b,a) = 2 \\
s_1(c,d) = 1 & s_2(b,a) = 1 & s_3(b,c) = 1 & s_4(d,c) = 1
\end{array}
$$

The preference profile $\gtrapprox$ that is induced by $\succsim$ is

$$
\begin{array}{ccccccc}
a & \ggg_1 & b & \ggg_1 & c & \ggg_1 & d \\
d & \ggg_2 & b & \ggg_2 & c & \ggg_2 & a \\
a & \ggg_3 & b & \ggg_3 & c & \ggg_3 & d \\
d & \ggg_4 & b & \ggg_4 & c & \ggg_4 & a.
\end{array}
$$

The Pareto efficient allocations with respect to $\gtrapprox$ are

$$w = (a, c, b, d), \qquad x = (a, d, b, c), \qquad y = (b, c, a, d), \qquad z = (b, d, a, c).$$

Finally, the unique[24] intensity-efficient allocation with respect to $\succsim$ is z. $\Diamond$

5 Concluding Remarks

The existing decision-theoretic apparatus for preference-intensity modelling –primarily developed in the 1940's-60s and invariably revolving around utility-difference representations of various types– has been associated with several conceptual and analytical challenges. This fact has deprived economists from a model of preference intensities that combines simplicity, generality, ordinality and tractability with transparency in its behavioural foundations. This paper aims to contribute towards filling this gap by proposing and axiomatically characterizing the novel model of preference intensity functions that was claimed to have these desirable features, and by also identifying its observationally testable beavioural content in the spirit of the revealed-preference analysis tradition. The model can be thought of as the simplest analogue of ordinal utility functions that allows for a decision maker's preferences and preference intensity comparisons to be represented in a genuinely ordinal way, without introducing any deviations from currently conventional notions of rationality on either of these relations, and at the same time without assuming that preference intensity comparisons are as if they were made by precision instruments. Moreover, the canonical normalization afforded by the model motivates a simple way to make ordinal interpersonal comparisons of preference intensities that do

[24] The reader can verify that z intensity-dominates all other allocations; x and z dominate w; and x, y are incomparable in this sense.

not rely on interpersonal comparisons of utilities. This in turn motivates the novel notion of intensity efficiency that was shown to be well-defined and also to refine Pareto efficiency by discarding allocations that are dominated on intensity-difference grounds.

References

Abdellaoui, M., C. Barrios, and P. P. Wakker (2007): "Reconciling Introspective Utility with Revealed Preference: Experimental Arguments Based on Prospect Theory," *Journal of Econometrics*, 138, 356–378.

Aczél, J. (1966): *Lectures on Functional Equations and their Applications*. New York: Academic Press.

Adams, E. W. (1965): "Elements of a Theory of Inexact Measurement," *Philosophy of Science*, 32, 205–228.

Alós-Ferrer, C., E. Fehr, and N. Netzer (2018): "Time Will Tell: Recovering Preferences when Choices Are Noisy," *working paper.*

Alt, F. (1936): "Über die Messbarkeit des Nutzens," *Zeitschrift für Nationalökonomie*, 7, 161–169.

——— (1971): "On the Measurability of Utility," in *Preferences, Utility and Demand*, ed. by J. S. Chipman, L. Hurwicz, M. K. Richter, and H. F. Sonnenschein, chap. 20. New York: Harcourt Brace Jovanovich, (English translation of Alt (1936) by S. Schach).

Argenziano, R., and I. Gilboa (2019): "Perception-Theoretic Foundations of Weighted Utilitarianism," *Economic Journal*, forthcoming.

Baccelli, J., and P. Mongin (2016): "Choice-Based Cardinal Utility: A Tribute to Patrick Suppes," *Journal of Economic Methodology*, 23, 268–288.

Bossert, W., and J. A. Weymark (2004): "Utility in Social Choice," in *Handbook of Utility Theory, Volume 2: Extensions*, ed. by S. Barbera, P. Hammond, and C. Seidl, pp. 1099–1177. Dordrecht: Kluwer.

Butler, D., A. Isoni, G. Loomes, and D. Navarro-Martinez (2014): "On the Measurement of Strength of Preference in Units of Money," *Economic Record*, 90, 1–15.

Butler, D., A. Isoni, G. Loomes, and K. Tsutsui (2014): "Beyond Choice: Investigating the Sensitivity and Validity of Measures of Strength of Preference," *Experimental Economics*, 17, 537–563.

Cerreia-Vioglio, S., F. Maccheroni, M. Marinacci, and A. Rustichini (2018): "Multinomial Logit Processes and Preference Discovery: Inside and Outside the Black Box," *working paper.*

Clithero, J. A. (2018): "Improving Out-of-Sample Predictions Using Response Times and a Model of the Decision Process," *Journal of Economic Behavior & Organization*, 148, 344–375.

Davidson, D., and J. Marschak (1959): "Experimental Tests of a Stochastic Decision Theory," in *Measurement: Definitions and Theories*, ed. by C. W. Churchman, and P. Ratoosh, pp. 233–269. New York: Wiley.

Debreu, G. (1958): "Stochastic Choice and Cardinal Utility," *Econometrica*, 26, 440–444.

Dhillon, A., and J.-F. Mertens (1999): "Relative Utilitarianism," *Econometrica*, 67, 471–498.

Echenique, F., and K. Saito (2017): "Response Time and Utility," *Journal of Economic Behavior & Organization*, 139, 49–59.

Ellingsen, T. (1994): "Cardinal Utility: A History of Hedonimetry," in *Cardinalism*, ed. by M. Allais, and O. Hagen, pp. 105–165. Dordrecht: Kluwer.

Falmagne, J. C. (2002): *Elements of Psychophysical Theory*. Oxford: Oxford University Press.

Fishburn, P. C. (1970a): *Utility Theory for Decision Making*. New York: Wiley.

——— (1970b): "Utility Theory with Inexact Preferences and Degrees of Preference," *Synthese*, 21, 204–221.

——— (2001): "Cancellation Conditions for Finite Two Dimensional Additive Measurement," *Journal of Mathematical Psychology*, 45, 2.

Fleurbaey, M., and P. Hammond (2004): "Interpersonally Comparable Utility," in *Handbook of Utility Theory, Volume 2: Extensions*, ed. by S. Barbera, P. Hammond, and C. Seidl, pp. 1179–1285. Dordrecht: Kluwer.

Fudenberg, D., R. Iijima, and T. Strzalecki (2015): "Stochastic Choice and Revealed Perturbed Utility," *Econometrica*, 83, 2371–2409.

Hammond, P. (1991): "Interpersonal Comparisons of Utility: Why and How They Are and Should Be Made," in *Interpersonal Comparisons of Well-Being*, ed. by J. Elster, and J. E. Roemer, pp. 200–254. Cambridge: Cambridge University Press.

Isbell, J. R. (1959): "Absolute Games," in *Contributions to the Theory of Games, Volume IV*, ed. by A. W. Tucker, and R. D. Luce, pp. 357–396. Princeton: Princeton University Press.

Köbberling, V. (2006): "Strength of Preference and Cardinal Utility," *Economic Theory*, 27, 375–391.

Konovalov, A., and I. Krajbich (2017): "Revealed Indifference: Using Response Times to Infer Preferences," *working paper*.

Krajbich, I., B. Oud, and E. Fehr (2014): "Benefits of Neuroeconomic Modelling: New Policy Interventions and Predictors of Preference," *American Economic Review: Papers & Proceedings*, 104, 501–506.

Krantz, D. H., R. D. Luce, P. Suppes, and A. Tversky (1971): *Foundations of Measurement, Volume I*. New York: Wiley.

Luce, R. D. (1956): "Semiorders and A Theory of Utility Discrimination," *Econometrica*, 24, 178–191.

Luce, R. D., D. H. Krantz, P. Suppes, and A. Tversky (1990): *Foundations of Measurement, Volume III.* San Diego and London: Academic Press.

Luce, R. D., and H. Raiffa (1957): *Games and Decisions.* New York: Wiley.

Luce, R. D., and P. Suppes (1965): "Preference, Utility and Subjective Probability," in *Handbook of Mathematical Psychology, Volume 3*, ed. by R. D. Luce, R. R. Bush, and E. H. Galanter, pp. 249–410. New York: Wiley.

Mandler, M. (1999): *Dilemmas in Economic Theory: Persisting Foundational Problems of Microeconomics.* New York: Oxford University Press.

——— (2006): "Cardinality versus Ordinality: A Suggested Compromise," *American Economic Review*, 96, 1114–1136.

Moscati, I. (2018): *Measuring Utility. From the Marginal Revolution to Behavioral Economics.* New York: Oxford University Press.

Ng, Y.-K. (1975): "Bentham or Bergson? Finite Sensibility, Utility Functions and Social Welfare Functions," *Review of Economic Studies*, 42, 545–569.

Ratcliff, R. (1978): "A Theory of Memory Retrieval," *Psychological Review*, 85, 59–108.

Ratcliff, R., and G. McKoon (2008): "The Diffusion Decision Model: Theory and Data for Two-Choice Decision Tasks," *Neural Computation*, 20, 873–922.

Richter, M. K. (1966): "Revealed Preference Theory," *Econometrica*, 34(3), 635–645.

Samuelson, P. A. (1938): "The Numerical Representation of Ordered Classifications and the Concept of Utility," *Review of Economic Studies*, 6, 65–70.

Scott, D. (1964): "Measurement Structures and Linear Inequalities," *Journal of Mathematical Psychology*, 1, 233–247.

Scott, D., and P. Suppes (1958): "Foundational Aspects of Theories of Measurement," *Journal of Symbolic Logic*, 23, 113–128.

Suppes, P., D. H. Krantz, R. D. Luce, and A. Tversky (1989): *Foundations of Measurement, Volume II.* San Diego and London: Academic Press.

Tversky, A., and J. E. Russo (1969): "Similarity and Substitutability in Binary Choices," *Journal of Mathematical Psychology*, 6, 1–12.

von Neumann, J., and O. Morgenstern (1947): *Theory of Games and Economic Behavior.* Princeton: Princeton University Press, 2nd edn.

Appendix: Proofs

Proof of Theorem 1.

We first establish the following auxiliary results.

Claim 1. *If $\succsim$ is a Weak Order and satisfies Reversal, then $(a,b) \succsim (b,a)$ and $(c,d) \succsim (d,c)$ implies $(a,b) \succsim (d,c)$. Moreover, if either $(a,b) \succ (b,a)$ or $(b,c) \succ (c,b)$ is also true, then $(a,b) \succ (d,c)$.*

For the first part, let $(a,b) \succsim (b,a)$, $(c,d) \succsim (d,c)$ and suppose to the contrary that $(b,a) \succ (c,d)$. Transitivity and $(a,b) \succsim (b,a) \succ (c,d) \succsim (d,c)$ implies $(a,b) \succ (d,c)$. In view of Reversal, this is a contradiction. Moreover, in view of Completeness, $(b,a) \not\succ (d,c)$ implies $(d,c) \succsim (b,a)$, which, under Reversal, further implies $(a,b) \succsim (c,d)$, as required. For the second part, let $(a,b) \succ (b,a)$, $(c,d) \succsim (d,c)$ and suppose to the contrary that $(a,b) \sim (d,c)$. Reversal implies $(c,d) \sim (b,a)$. Now noting that $(a,b) \succ (b,a)$ holds, Transitivity and $(d,c) \sim (a,b) \succ (b,a) \sim (c,d)$ together imply $(d,c) \succ (c,d)$, which contradicts $(c,d) \succsim (d,c)$. The argument in the case where $(a,b) \succsim (b,a)$ and $(c,d) \succsim (d,c)$ is symmetric. $\Diamond$

Claim 2. *If $\succsim$ satisfies Weak Order, Reversal and Consistency, then $\gtrapprox$ is a weak order on X.*

Completeness of $\gtrapprox$ immediately follows from its definition and the assumed Completeness of $\succsim$. For Transitivity, suppose $a \gtrapprox b$ and $b \gtrapprox c$, and assume to the contrary that $c \ggg a$. We have $(a,b) \succsim (b,a)$, $(b,c) \succsim (c,b)$ and $(c,a) \succ (a,c)$. In view of Claim 1, $(a,b) \succsim (b,a)$ and $(c,a) \succ (a,c)$ implies $(a,b) \succ (a,c)$. By Reversal, $(c,a) \succ (b,a)$. By Consistency, $(c,b) \succ (b,c)$. This is a contradiction. $\Diamond$

Claim 3. *If $\succsim$ satisfies Weak Order, Reversal and Consistency, then $(a,b) \succsim (b,a)$ and $(b,c) \succsim (c,b)$ implies $(a,c) \succsim (a,b)$ and $(a,c) \succsim (b,c)$.*

Suppose $(a,b) \succsim (b,a)$ and $(b,c) \succsim (c,b)$ and assume to the contrary that $(a,b) \succ (a,c)$. By Claim 1, $(a,b) \succsim (b,a)$ and $(a,b) \succ (a,c)$ together imply $(a,b) \succ (b,a)$. Similarly, $(b,c) \succsim (c,b)$ and $(a,b) \succ (a,c)$ together imply $(b,c) \succ (a,c)$. Moreover, $(a,b) \succ (b,a)$ and $(b,c) \succ (a,c)$ implies $(a,b) \succ (a,c)$. By Reversal, $(c,a) \succ (b,a)$. By Consistency, $(c,b) \succ (b,c)$. This is a contradiction. The implication $(a,c) \succsim (b,c)$ is established symmetrically. $\Diamond$

$1 \Rightarrow 2$. Suppose $(a,c) \succsim (b,c)$ and assume to the contrary that $(a,d) \not\succsim (b,d)$. By Completeness, this implies $(b,d) \succ (a,d)$. Consistency and $(a,c) \succsim (b,c)$ implies $(a,b) \succsim (b,a)$. Consistency and $(b,d) \succ (a,d)$ also implies $(b,a) \succsim (a,b)$. Suppose to the contrary that $(a,b) \sim (a,b)$. Since $(b,d) \succ (a,d)$ is also true, it follows from Claim 1 that $(a,b) \sim (b,a) \succ (a,d)$ and, by Transitivity and Reversal, $(a,b) \succ (a,d)$ and $(d,a) \succ (b,a)$, respectively. The latter and Consistency together imply $(d,b) \succsim (b,d)$. This, together with Transitivity and $(b,d) \succ (a,d)$, implies $(d,b) \succ (a,d)$ which, by Reversal, is equivalent to $(d,a) \succ (b,d)$. This contradicts the postulate $(b,d) \succ (a,d)$.

$2 \Rightarrow 1$. Suppose $(a,c) \succsim (b,c)$ and assume to the contrary that $(a,b) \not\succsim (b,a)$. By Completeness, $(b,a) \succ (a,b)$. By Separability and Reversal, $(a,b) \succsim (b,b) \succsim (b,a)$. This is a contradiction.

$1 \Rightarrow 3$. Note first that, since X is finite and $\gtrapprox$ is a weak order on X (Claim 2), there exist k

$\approx$-equivalence classes $[a_i]$ which, with a slight abuse of notation, can be strictly ordered as

$$[a_1] \ggcurly \ldots \ggcurly [a_k].$$

The above ordering will be held fixed throughout the proof. In particular, it is understood that, for any $i \leqslant k$, $a, b \in [a_i] \Leftrightarrow a \approx b$ and also that, for any $i < j$, $a \in [x_i]$ and $b \in [x_j] \Leftrightarrow a \ggcurly b$.

Let the $\approx$-quotient set of X be defined by $X_{\approx} \equiv \mathcal{X} := \{[a_1], \ldots, [a_k]\}$. Let also

$$A \;\; := \;\; \{[a_i] \times [a_j] \in \mathcal{X} \times \mathcal{X} : i < j\}$$

and

$$Q_{\succ}(a_i, a_j) \;\; := \;\; \{[a_h] \times [a_s] \in A : (a_i, a_j) \succ (a_h, a_s)\}$$

That is, $[a_h] \times [a_s] \in Q_{\succ}(a_i, a_j)$ if and only if, for all $(a_h, a_s) \in [a_h] \times [a_s]$, $a_h \ggcurly a_s$ and $(a_i, a_j) \succ (a_h, a_s)$. Notice that $Q_{\succ}(a_i, a_j) \neq \varnothing$ implies $i < j$ but the converse is not true in general.

Now define the function $s : X \times X \to \mathbb{R}$ by

$$s(a_i, a_j) := \begin{cases} 1 + |Q_{\succ}(a_i, a_j)|, & \text{if } i < j \\ \\ 0, & \text{if } i = j \\ \\ -s(a_j, a_i), & \text{if } i > j \end{cases}$$

Note that this s is well-defined in $X \times X$ since $(a_i, a_j) \in X \times X$ if and only if $(a_i, a_j) \in [a_i] \times [a_j]$ for some $[a_i], [a_j] \in \mathcal{X}$ where, clearly, exactly one of $i < j$, $i = j$ and $i > j$ is true. Moreover, s satisfies (3b) by construction. We will show that s also satisfies (3a), and it will then follow from Claim 3 that s obeys (3c) as well.

Notice first that it follows from the definitions of s and $Q_{\succ}(\cdot)$, and also from $[a_1] \ggcurly \ldots \ggcurly [a_k]$, that $s(a_i, a_j) > 0 \Leftrightarrow a_i \ggcurly a_j$ and $s(a_i, a_j) = 0 \Leftrightarrow a_i \approx a_j$. Now suppose $(a_j, a_l) \succsim (a_m, a_n)$ and assume $j \leqslant l$. It holds that $[a_m] \times [a_n] \in Q_{\succ}(a_j, a_l)$ or $Q_{\succ}(a_j, a_l) = \varnothing$. Given the definitions of s and $Q_{\succ}(\cdot)$, the first case implies $s(a_j, a_l) > s(a_m, a_n)$ because $(a_j, a_l) \succ (a_m, a_n)$ and therefore $Q_{\succ}(a_j, a_l) \supset Q_{\succ}(a_m, a_n)$ since $\succsim$ is a weak order on $X \times X$. The second case, $Q_{\succ}(a_j, a_l) = \varnothing$, implies $s(a_j, a_l) = 1$. Moreover, if $Q_{\succ}(a_j, a_l) = \varnothing$ and $m \leqslant n$, then $(a_j, a_l) \succsim (a_m, a_n)$ implies $(a_j, a_l) \sim (a_m, a_n)$, which further implies $s(a_m, a_n) = 1 = s(a_j, a_l)$. On the other hand, $Q_{\succ}(a_j, a_l) = \varnothing$ and $m > n$ implies $s(a_m, a_n) < 0 < s(a_j, a_l) = 1$. Assume now that $j > l$. In view of Claim 1, this implies $m > n$. Reversal now implies $(a_n, a_m) \succsim (a_l, a_j)$. Applying the above argument to this case establishes that $s(a_n, a_m) \geqslant s(a_l, a_j)$ and, given that $s(a, b) = -s(b, a)$ for all $a, b \in X$ is true by construction, $s(a_j, a_l) \geqslant s(a_m, a_n)$. Thus, for all $a_j, a_l, a_m, a_n \in X$, $(a_j, a_l) \succsim (a_m, a_n)$ implies $s(a_j, a_l) \geqslant s(a_m, a_n)$. Conversely, suppose $s(a_j, a_l) \geqslant s(a_m, a_n)$. Assume to the contrary that $(a_j, a_l) \not\succsim (a_m, a_n)$. Since $\succsim$ is complete, this implies $(a_m, a_n) \succ (a_j, a_l)$. It now follows from the above arguments that $s(a_m, a_n) > s(a_j, a_l)$, a contradiction. Therefore, s represents $\succsim$ as in (3). Moreover, by construction, $s(X \times X)$ is a symmetric set of consecutive integers. Hence, s constitutes a canonical preference intensity representation of $\succsim$.

3 $\Rightarrow$ 4. Since a canonical preference intensity function is a special case of a preference intensity function, the existence claim is obviously true. To establish the uniqueness property, let s be a preference intensity function that represents $\succsim$ and let t be an odd and strictly increasing transformation of s. Since $s(a,b) = -s(b,a)$ and $t(a,b) = f(s(a,b))$ for some function $f : \mathbb{R} \to \mathbb{R}$ that is odd in $s(X \times X)$, we have

$$t(a,b) = f(s(a,b)) = -f(-s(a,b)) = -f(s(b,a)) = -t(b,a).$$

Now suppose $(a,b) \succsim (c,d)$. This is equivalent to $s(a,b) \geqslant s(c,d)$. Since t is a strictly increasing transformation of s, it follows that $t(a,b) \geqslant t(c,d)$ too. Conversely, suppose $\succsim$ is represented by two distinct preference intensity functions s and t. Let $t := f \circ s$ for some function $f : \mathbb{R} \to \mathbb{R}$. Suppose $f(-z) \neq -f(z)$ for some $z \in s(X \times X)$. Let $z = s(a,b)$. Since $s(a,b) = -s(b,a)$ and $t(a,b) = f(s(a,b))$, by assumption, it follows that $t(b,a) = f(s(b,a)) = f(-s(a,b)) \neq -f(s(a,b)) = -t(a,b)$, which contradicts the assumption that t represents $\succsim$. Therefore, f is odd in $s(X \times X)$. Now suppose $f(z) \leqslant f(z')$ for some $z, z' \in s(X \times X)$ such that $z > z'$. Suppose $z = s(a,b)$ and $z' = s(c,d)$. By assumption, $(a,b) \succ (c,d)$. Since $f(z) = f(s(a,b)) = t(a,b) \leqslant t(c,d) = f(s(c,d)) = f(z')$, this again contradicts the assumption that t represents $\succsim$. Therefore, f is strictly increasing in $s(X \times X)$.

4 $\Rightarrow$ 5. It is obvious that if there exists $s : X \times X \to \mathbb{R}$ that represents $\succsim$ as in (3), then s also represents $\succsim$ as in (4). The relevant part in the proof that 3 $\Rightarrow$ 4 can be invoked to also establish uniqueness up to a strictly increasing transformation of an arbitrary $g : X \times X \to \mathbb{R}$ that represents $\succsim$ as in (4).

5 $\Rightarrow$ 1. Weak Order is implied by (4a) and Reversal is implied by (4b). To show that Consistency is also implied by (4), suppose to the contrary that $(a,c) \succsim (b,c)$ and $(b,a) \succ (a,b)$. We have $g(a,c) \geqslant g(b,c)$ and $g(b,a) > g(a,b)$. Notice first that, from (4a) and (4b), $g(b,a) > g(a,b)$ implies $g(b,a) > g(a,a) > g(a,b)$.

Suppose first that $g(b,c) \geqslant g(b,a)$. We then have $g(a,c) \geqslant g(b,c) \geqslant g(b,a) > g(a,a) > g(a,b)$. Hence, $g(a,c) > g(a,a) > g(a,b)$ and, by (4b), $g(b,a) > g(a,a) > g(c,a)$. It follows then that $g(b,a), g(a,c) > g(a,a)$ and, by (4c), $g(b,c) > g(a,c)$. This is a contradiction. Now suppose $g(b,a) > g(b,c)$ instead. Then, either $g(b,a) > g(a,c)$ or $g(a,c) \geqslant g(b,a)$ is also true. Consider the former case first. We have $g(b,a) > g(a,a) > g(a,b)$ and $g(b,a) > g(a,c) \geqslant g(b,c)$. Suppose $g(a,c) \geqslant g(a,a)$. Then, by (4c), $g(b,a) > g(a,a)$ and $g(a,c) \geqslant g(a,a)$ implies $g(b,c) \geqslant g(b,a)$, which contradicts the above postulate. Now suppose $g(a,a) > g(a,c)$ instead. We have $g(a,a) > g(a,c) \geqslant g(b,c)$. By (4b), this implies $g(c,b) > g(a,a)$. By (4c), moreover, this and $g(b,a) > g(a,a)$ together imply $g(c,a) > g(c,b)$. By (4b) again, this is equivalent to $g(b,c) > g(a,c)$ which contradicts the above postulate. Consider, finally, the case where $g(a,c) \geqslant g(b,a)$. We have $g(a,c) \geqslant g(b,a) > g(a,a) > g(a,b)$ and $g(b,a) > g(b,c)$. Therefore, by (4c), $g(b,a) > g(a,a)$ and $g(a,c) > g(a,a)$ implies $g(b,c) > g(b,a)$, a contradiction. ■

Proof of Corollary 1.

1 $\Leftrightarrow$ 2. See Theorem 3.2 in Scott (1964).

2 $\Rightarrow$ 3. Defining $s : X \times X \to \mathbb{R}$ by $s(x,y) := u(x) - u(y)$ trivially establishes the claim.

3 ⇒ 2. It is well-known (see, for example, Theorem 2, p. 356 in Aczél, 1966 or pp. 97-98 in Falmagne, 2002) that, under very general conditions which encompass those of Corollary 1, the solution to a so-called *Sincov* functional equation $f(x,y) = f(x,z) + f(z,y)$ is given by $f(x,y) = g(x) - g(y)$ for a unique function g, thereby establishing the claim. For completeness, a simple direct proof is also provided below.

Let z be an arbitrary element of X. Suppose $\gtrsim$ is represented by the triangularly additive preference intensity function $s : X \times X \to \mathbb{R}$. We have

$$\begin{aligned} a \gtrapprox b &\iff s(a,b) \geqslant s(b,a) \\ &\iff s(a,b) \geqslant 0 \\ &\iff s(a,b) + s(b,z) \geqslant s(b,z) \\ &\iff s(a,z) \geqslant s(b,z) \end{aligned} \tag{10}$$

where the last step makes use of the fact that s is triangularly additive. Now define the function $u : X \to \mathbb{R}$ by

$$u(a) := s(a,z). \tag{11}$$

It follows from (10) and (11) that

$$\begin{aligned} a \gtrapprox b &\iff s(a,z) \geqslant s(b,z) \\ &\iff u(a) \geqslant u(b). \end{aligned}$$

Thus,

$$\begin{aligned} (a,b) \gtrsim (c,d) &\iff s(a,b) \geqslant s(c,d) \\ &\iff s(a,z) + s(z,b) \geqslant s(c,z) + s(z,d) \\ &\iff s(a,z) - s(b,z) \geqslant s(c,z) - s(d,z) \\ &\iff u(a) - u(b) \geqslant u(c) - u(d). \end{aligned}$$

■

Proof of Corollary 2.

1 ⇔ 2: See Tversky and Russo (1969).

2 ⇔ 3: Omitted (analogous to the proof of 2 ⇔ 3 in Corollary 1). ■

Proof of Corollary 3.

1 ⇔ 2: See Proposition 1 in Fudenberg, Iijima, and Strzalecki (2015).

2 ⇔ 3: Omitted (analogous to the proof of 2 ⇔ 3 in Corollary 1). ■

Proof of Theorem 2.

With a slight abuse of notation, $\{a,b\} \in \mathcal{D}$ and $r_{\{a,b\}} \in \mathcal{D}$ is written when it is understood that choice from menu $\{a,b\}$ and the resource value $r_{\{a,b\}}$ at that menu are observable in $\mathcal{D}$. Moreover, without loss of generality, the proof assumes that Congruent Monotonicity *(negative)* is satisfied. The case where Congruent Monotonicity *(positive)* is satisfied instead can be dealt with in a symmetric way.

The argument for the "only if" part of the claim is easy and omitted. Now, for $a, b \in X$ such that $a \not\ggcurly^B b$ and $b \not\ggcurly^B a$ (to be written $a \parallel^B b$), denote by $[a,b]$ the –possibly empty– collection of all maximal sequences $\{a_1, \ldots, a_k\}$ such that $a = a_1$, $b = a_k$ and either $a_i \ggcurly^B a_{i+1}$ for all $i = 1, \ldots, k-1$ or $a_{i+1} \ggcurly^B a_i$ for all $i = 1, \ldots, k-1$ (the latter two situations will be denoted by $a \ggcurly^{\hat{B}} b$ and $b \ggcurly^{\hat{B}} a$, respectively). Moreover, if $a \ggcurly^B b$ or $b \ggcurly^B a$, write $[a,b] := \{a,b\}$. Finally, if $a \not\ggcurly^{\hat{B}} b$ and $b \not\ggcurly^{\hat{B}} a$, write $a \parallel^{\hat{B}} b$. In light of these definitions, we have

$$\begin{aligned}
&[a,b] = \{a,b\}, && \text{if } a \ggcurly^B b \text{ or } b \ggcurly^B a \\
&[a,b] \ni \{x_i\}_{i=1}^{k \geqslant 3} \supset \{a,b\}, && \text{if } a \not\ggcurly^B b,\ b \not\ggcurly^B a \text{ and } a \ggcurly^{\hat{B}} b \text{ or } b \ggcurly^{\hat{B}} a \\
&[a,b] = \varnothing, && \text{if } a \parallel^{\hat{B}} b
\end{aligned}$$

Now observe that, by the first requirement of Congruent Monotonicity, $x \ggcurly^{\hat{B}} y$ implies $y \not\ggcurly^{\hat{B}} x$. Hence, we have the following:

Observation. *For all $a, b \in X$, exactly one of the following is true: $a \ggcurly^{\hat{B}} b$; $b \ggcurly^{\hat{B}} a$; $a \parallel^{\hat{B}} b$.*

Next, define the function $s : X \times X \to \mathbb{R}$ by

$$s(a,b) := \begin{cases} \left(\min\limits_{\{a_i\}_{i=1}^k \in [a,b]} \min\limits_{\substack{\{a_i, a_{i+l}\} \in \mathcal{D}, \\ i+l \leqslant k}} r_{\{a_i, a_{i+l}\}} \right)^{-1} - \epsilon, & \text{if } a \ggcurly^{\hat{B}} b \\ 0, & \text{if } a \parallel^{\hat{B}} b \\ \epsilon - \left(\min\limits_{\{a_i\}_{i=1}^k \in [a,b]} \min\limits_{\substack{\{a_i, a_{i+l}\} \in \mathcal{D}, \\ i+l \leqslant k}} r_{\{a_i, a_{i+l}\}} \right)^{-1}, & \text{if } b \ggcurly^{\hat{B}} a \end{cases} \tag{12}$$

where

$$\epsilon := \frac{1}{2} \min_{\{a,b\} \in \mathcal{D}} r_{\{a,b\}}.$$

In view of the Observation and the preceding remarks, s is well-defined in $X \times X$.

Note next that, for any function $f : \mathbb{R} \to \mathbb{R}$ with the property that

$$f(r_{\{a,b\}}) := \begin{cases} \frac{1}{r_{\{a,b\}}} - \epsilon, & \text{if } a \gg^B b \\ \\ \epsilon - \frac{1}{r_{\{a,b\}}}, & \text{if } b \gg^B a \end{cases}$$

for all $r_{\{a,b\}} \in \mathcal{D}$, it holds that $s(a,b) = f\left(r_{\{a,b\}}\right)$ and f is strictly decreasing in $r_{\{a,b\}}$. Moreover, $s(a,b) = -s(b,a)$ and $s(a,b) > 0 \Leftrightarrow a \gg^{\hat{B}} b$ also hold by definition (for the latter claim, recall also the definition of ϵ). The above implies, in particular, that $a \gg^B b \Rightarrow s(a,b) > 0$. Therefore, (9a) and (9b) hold. It remains to be verified that $s(a,b), s(b,c) > 0$ implies $s(a,c) \geqslant s(a,b), s(b,c)$.

To this end, note first that $s(a,b), s(b,c) > 0$ implies $a \gg^{\hat{B}} b$ and $b \gg^{\hat{B}} c$, which in turn implies $a \gg^{\hat{B}} c$. The claim that $s(a,c) \geqslant s(a,b), s(b,c)$ will be established by showing that the second part of Congruent Monotonicity implies

$$\min_{\{x_i\}_{i=1}^k \in [a,c]} \min_{\substack{\{x_i, x_{i+l}\} \in \mathcal{D} \\ i+l \leqslant k}} r_{\{x_i, x_{i+l}\}} \leqslant \min_{\{y_i\}_{i=1}^m \in [a,b]} \min_{\substack{\{y_i, y_{i+l}\} \in \mathcal{D} \\ i+l \leqslant m}} r_{\{y_i, y_{i+l}\}}, \min_{\{z_i\}_{i=1}^n \in [b,c]} \min_{\substack{\{z_i, z_{i+l}\} \in \mathcal{D} \\ i+l \leqslant n}} r_{\{z_i, z_{i+l}\}}. \tag{13}$$

It is immediate that the minimum operators in (13) are redundant and the inequality holds strictly if $[a,b] = \{a,b\}$, $[b,c] = \{b,c\}$ and $[a,c] = \{a,c\}$, in which case $a \gg^B b \gg^B c$ and $a \gg^B c$ apply. Similarly, the claim that (13) holds is also immediate if $[a,b] \neq \{a,b\}$, $[b,c] \neq \{b,c\}$ and there are unique maximal sequences $\{b_i\}_{i=1}^m$ and $\{c_i\}_{i=1}^n$ such that $[a,b] = \{b_i\}_{i=1}^m$, $[b,c] = \{c_i\}_{i=1}^n$, in which case $[a,c] = \{b_i\}_{i=1}^m \cup \{c_i\}_{i=1}^n$ is also true. Now suppose $[a,b]$ or $[b,c]$ comprises more than one maximal sequence with the above properties. Then, each $\{a_i\}_{i=1}^k \in [a,c]$ can be written as

$$\{a_i\}_{i=1}^k = \{b_i\}_{i=1}^m \cup \{c_i\}_{i=1}^n \qquad \text{for some } \{b_i\}_{i=1}^m \in [a,b],\ \{c_i\}_{i=1}^n \in [b,c]. \tag{14}$$

Suppose to the contrary that

$$\min_{\{a_i\}_{i=1}^k \in [a,c]} \min_{\substack{\{a_i, a_{i+l}\} \in \mathcal{D} \\ i+l \leqslant k}} r_{\{a_i, a_{i+l}\}} > \min_{\{b_i\}_{i=1}^m \in [a,b]} \min_{\substack{\{b_i, b_{i+l}\} \in \mathcal{D} \\ i+l \leqslant m}} r_{\{b_i, b_{i+l}\}}. \tag{15}$$

It holds that $\min_{\{b_i\}_{i=1}^m \in [a,b]} \min_{\substack{\{b_i, b_{i+l}\} \in \mathcal{D} \\ i+l \leqslant m}} r_{\{b_i, b_{i+l}\}} = r_{\{a',b'\}}$ for some $a', b' \in X$ such that $a' \gg^B b'$. By (14) and Congruent Monotonicity,

$$r_{\{a',b'\}} \geqslant \min_{\{a_i\}_{i=1}^k \in [a,c]} \min_{\substack{\{a_i, a_{i+l}\} \in \mathcal{D} \\ i+l \leqslant k}} r_{\{a_i, a_{i+l}\}}.$$

This contradicts (15). The case where $[a,b]$ is replaced by $[b,c]$ on the right hand side of (15) is similarly ruled out. Therefore, (13) holds. ■

Proof of Corollary 4.

$1 \Leftrightarrow 2$. See Theorem 3 in Echenique and Saito (2017).

$2 \Leftrightarrow 3$. Omitted (analogous to the proof of $2 \Leftrightarrow 3$ in Corollary 1). ■

Proof of Proposition 3.

1. It suffices to show that the intensity-dominance relation on X^n is transitive. Suppose that x intensity-dominates y, and y intensity-dominates z. We will show that x intensity-dominates z. We will proceed by partitioning the set of agents as follows:

Class 1. Agents in some pair (i, j) such that the first criterion of Definition 8 is satisfied for that pair both with respect to allocations x, y and with respect to y, z.

Class 2. Agents in some pair (i, j) such that the first criterion of Definition 8 is satisfied for that pair with respect to either x, y or y, z but not both.

Class 3. Agents that do not belong to one of the above two classes.

We will say that the pair (i, j) is *flipped* by x and y if $(x_i, x_j) = (y_j, y_i) = (a, b)$ for some $a, b \in X$.

Case 1. Suppose that (i, j) is a pair in Class 1, and suppose to the contrary that $s_i(x_i, x_j) > s_j(y_j, y_i)$, $s_i(y_i, y_j) \geqslant s_j(z_j, z_i)$ or $s_i(x_i, x_j) \geqslant s_j(y_j, y_i)$, $s_i(y_i, y_j) > s_j(z_j, z_i)$. Since (i, j) is flipped by both x, y and y, z by assumption, it follows that, for some $a, b \in X$, $(x_i, x_j) = (y_j, y_i) = (a, b)$ and $(y_i, y_j) = (z_j, z_i) = (b, a)$. Without loss of generality, consider the first of the above two cases. It holds that $s_i(x_i, x_j) = s_i(a, b) > s_j(a, b) = s_j(y_j, y_i)$ and $s_i(y_i, y_j) = s_i(b, a) = -s_i(a, b) \geqslant -s_j(a, b) = s_j(b, a) = s_j(z_j, z_i)$. This is a contradiction. Therefore, $s_i(x_i, x_j) = s_j(y_j, y_i)$ and $s_i(y_i, y_j) = s_j(z_j, z_i)$ must be true. This in turn implies that allocations x and z are identical with respect to every such pair (i, j), i.e. $(x_i, x_j) = (z_i, z_j)$.

Case 2. Suppose that the pairs (i, j) and (k, l) are flipped by x, y and y, z, respectively, and $s_i(x_i, x_j) \geqslant s_j(y_j, y_i)$, $s_k(y_k, y_l) \geqslant s_l(z_l, z_k)$. By assumption, $(i, j) \neq (k, l)$, and $s_i(y_i, z_i), s_j(y_z, z_j) \geqslant 0$, $s_k(x_k, y_k), s_l(x_l, y_l) \geqslant 0$. Consider first the possibility that (i, j) or (k, l) is flipped by x and z. Without loss of generality, suppose this is true for (i, j). This implies $(x_i, x_j) = (z_j, z_i)$. It now follows from the above that $(y_j, y_i) = (z_j, z_i)$, and therefore $s_i(x_i, x_j) \geqslant s_j(z_j, z_i)$. Thus, the first condition of Definition 8 for x intensity-dominating z is satisfied in this case for agents i, j, k, l. Now consider the possibility where neither (i, j) nor (k, l) is flipped by x and z. Notice that we can assume without loss of generality that $s_i(x_i, y_i) \geqslant 0$ (indeed, if $s_i(x_i, y_i) \equiv s_i(x_i, x_j) \leqslant 0$, then the above implies $s_j(y_j, x_j) \equiv s_j(y_j, x_j) \leqslant 0$, and hence $s_j(x_j, x_i) \geqslant s_i(y_i, y_j)$ so that x intensity-dominates y at the pair (j, i) instead of the pair (i, j) with respect the first part of Definition 8). Given this and $s_i(y_i, z_i) \geqslant 0$ also being true by the assumption that y intensity-dominates z, it now follows from (3c) that $s_i(x_i, z_i) \geqslant 0$. Thus, the second condition in Definition 3c for the intensity-dominance of x over z is satisfied in this case. Finally, by the dominance assumption, at least one inequality is strict in at least one of the above two cases. This implies that the third part of Definition 8 for the intensity-dominance of x over z is also satisfied.

Case 3. Consider some agent i that belongs to Class 3. The dominance assumption for x over y and y over z implies $s_i(x_i, y_i), s_i(y_i, z_i) \geqslant 0$, which, by (3c), implies $s_i(x_i, z_i) \geqslant 0$. Therefore, the second part of Definition 8 is satisfied for every such agent.

It follows from the above that if x intensity-dominates y and y intensity-dominates z, then the dominance with respect to the first two parts of Definition 8 is carried over for x over z for all agents in each of the above three equivalence classes, with at least one strict inequality holding with respect to either the first or the second part of Definition 8. This establishes that x intensity-dominates z.

2. It suffices to show that whenever x Pareto-dominates y, x also intensity-dominates y. Suppose then that x Pareto-dominates y. This implies that $s_i(x_i, y_i) \geqslant 0$ holds for all $i \leqslant n$, with strict inequality for some i. Suppose first that there is no pair of agents (i, j) that is flipped by x and y. This implies that the first part of Definition 8 is trivially satisfied, while, from the above, the second and third parts are satisfied too. Therefore, x intensity-dominates y. Now suppose that there is some pair (i, j) that is flipped by x and y, and assume to the contrary that $s_i(x_i, x_j) < s_j(y_j, y_i)$. By assumption, $(x_i, x_j) = (y_j, y_i) = (a, b)$ for some $a, b \in X$. Since x Pareto-dominates y, we have $s_i(x_i, x_j) \equiv s_i(x_i, y_i) = s_i(a, b) \geqslant 0$ and $s_j(x_j, y_j) \equiv s_j(y_i, y_j) = s_j(b, a) \geqslant 0 \geqslant s_j(a, b) = s_j(y_j, y_i)$. It follows that $0 \geqslant s_j(y_j, y_i) > s_i(x_i, x_j) \geqslant 0$, which is a contradiction. Thus, all three parts of Definition 8 are satisfied in this case as well, thereby establishing that x intensity-dominates y. ■