Simple Preference Intensity Comparisons

Georgios Gerasimou[*†]
University of St Andrews

December 30, 2019

Abstract

We propose a general and applicable model of simple preference intensity comparisons. The model encompasses those that belong to the utility-difference class, has transparent behavioural underpinnings and features purely ordinal uniqueness properties. Three applications are analysed. First, the model's empirical content is characterized by an easily testable condition on behavioural data that include choices and additional observables with intensity-revealing potential that are often elicited in experimental/empirical work, such as survey ratings, response times or willingness to pay. Second, a special case of the model is argued to facilitate interpersonal comparisons of (strict) intensity relations without requiring interpersonally comparable utilities. Building on this, the novel notion of *intensity efficiency* is introduced in a single-profile social choice setting and is shown to be well-defined and to refine Pareto efficiency by discarding allocations that are dominated on intensity-difference grounds. Finally, the house allocation problem in one-sided matching is revisited when agents have intensity relations of this kind, and a simple algorithm is shown to yield an intensity-efficient allocation in every such market.

Keywords: Preference intensity functions; revealed preference intensity; intensity-efficient allocations; welfare economics; one-sided matching.

JEL: B21, C88, D11, D63, D70, D90

*This article subsumes work previously circulated under the title *"Preference Intensity Representation and Revelation"* (first version: 4th December 2017). Thanks are due to audiences at NASMES 2018 (University of California at Davis), SARP 2018 (University of Oxford), RUD 2019 (Paris School of Economics), EEA/ESEM 2019 (University of Manchester), RES 2019 (University of Warwick), Royal Holloway - University of London, and especially Jean Baccelli, Tore Ellingsen, Marcelo A. Fernandez, Simon Grant, Peter Hammond, Ian Krajbich, Dorothea Kübler, Michael Mandler, Hervé Moulin, François Maniquet, John Quah and Peter Wakker for helpful comments. Matúš Tejiščák provided expert research assistance with coding. Any errors are the author's own.

†Email address: gg26@st-andrews.ac.uk

1 Introduction

The idea that decision makers not only have preferences over the choice alternatives of interest but may also be capable of ranking differences in preference intensities between those alternatives has been present -often implicitly- in many areas of economics, including decision theory and experiments, welfare and public economics, political economy and voting, as well as matching and market design.[1] All existing modelling approaches in this large and diverse body of work assume that an individual's preference intensities are ultimately expressible on a cardinal or pseudo-cardinal utility scale, often under very restrictive or confounding assumptions on the primitives of the analysis, such as quasilinear and/or expected-utility preferences over lotteries. Contrary to these approaches, in this study we model -for the first time- decision makers as capable of making simple and purely ordinal preference intensity comparisons such as *"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"* without assuming that such comparisons can also be quantified in any way. We formalise and give decision-theoretic foundations for this modelling approach; argue that it has many important advantages over all utility-difference approaches; and, finally, provide applications that illustrate its empirical testability and also uncover novel normative implications for welfare economics and matching/allocation problems where monetary transfers are infeasible.

Our analytical approach towards the numerical representation of ordinal preference intensities takes Paul Samuelson's (1938) bivariate formulation of neoclassical cardinal utility representations as a starting point and extends it in a natural and intuitive 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 (rational) preference-intensity modelling that were expressed by Samuelson and his contemporaries and were later 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

[1] A non-exhaustive list of relevant references in these areas is provided below. Many of them are discussed in more detail in later sections.
Decision/neuroeconomic theory and experiments: Konovalov and Krajbich (2019), Baldassi, Cerreia-Vioglio, Maccheroni, Marinacci, and Pirazzini (2019), Aguiar and Serrano (2019), Alós-Ferrer, Fehr, and Netzer (2018), Echenique and Saito (2017), Caplin (2014), Butler, Isoni, Loomes, and Navarro-Martinez (2014a), Butler, Isoni, Loomes, and Tsutsui (2014b), Sákovics (2013), Abdellaoui, Barrios, and Wakker (2007), Köbberling (2006), Basu (1982), Shapley (1975), Tversky (1967a), Adams (1965), Scott (1964), Debreu (1958), Scott and Suppes (1958), Suppes and Winet (1955), Mosteller and Nogee (1951), von Neumann and Morgenstern (1947), Samuelson (1938), Alt (1936, 1971).
Utilitarian social choice, welfare economics and public economics: Argenziano and Gilboa (2019), Fleurbaey and Maniquet (2018), Sen (2017/1970), Piacquadio (2017), Fleurbaey and Maniquet (2011), Fleurbaey and Mongin (2005), Moulin (2004), Fleurbaey and Hammond (2004), Bossert and Weymark (2004), Dhillon and Mertens (1999), Ng (1975), Campbell (1973), Harsanyi (1955, 1990).
Political economy and voting: Casella and Palfrey (2019), Casella and Gelman (2008), Casella, Palfrey, and Riezman (2008), Dahl (2006/1956), DeMeyer and Plott (1971).
Matching theory and experiments: Budish and Kessler (2018), Pycia and Ünver (2018), Lee and Yariv (2018), Abdulkadiroğlu, Che, and Yasuda (2011, 2015), Che (2013), Che, Gale, and Kim (2013), Klijn, Pais, and Vorsatz (2013), Budish (2011), Roth and Sotomayor (1990), Hylland and Zeckhauser (1979).

provide a quick preview– is given in Section 3.

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 intensity ordering. In particular, if a is preferred to b more than c is to d, then a higher value is associated with the pair (a, b) than with (c, d). The key novelty of preference intensity representations that formally distinguishes them from utility-difference ones –or bivariate representations that are in fact 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 *at least as much* as a is to b and b is to c. It ensures conceptual harmony between and within the preference intensity and induced preference relations while also establishing transitivity of the latter, and does not force –as is done by utility-difference models– that the value capturing the intensity difference between a and c be the *sum* of the values of the intensity differences between a and b and between b and c. By not imposing this additivity requirement, the model is characterized by standard Weak Order and Reversal together with a very mild Consistency axiom.

The proposed class of functions are the first in the literature that offer a genuinely ordinally unique representation of preference intensity relations that do not portray the agent as being able to somehow quantify her intensity comparisons. This is in stark contrast to cardinal utility models where the decision maker is modelled as if she was able to think of a as being preferred to b *exactly* r many times as c is preferred to b, even though this degree of precision is obviously unrealistic. Other, non-cardinal utility-difference models do not make as sharp predictions but are not invariant with respect to arbitrary strictly increasing transformations either, and ultimately assume that the agent's intensity ordering is captured by some pseudo-cardinal ranking. By contrast, we show that, in their most general formulation, preference intensity functions are simply unique up to arbitrary strictly increasing transformations, hence ordinally unique in the standard sense. This property is important and 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 application of this model is in the study of observable behavioural datasets and testable restrictions thereof that are necessary and sufficient for observable behaviour to be thought of as being generated by a decision maker whose preferences and preference-intensity comparisons belong to the proposed general class. This analysis leads to a significant generalization of the recent related result 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. Generalizing their model, 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; intensity ratings), 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 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. By using preference intensity functions instead of difference-preserving utility functions as its building block, this more general rationalizability notion -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 revealed preference intensity relations.

Our second application 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 baseline numerical representation whereby all agents' (strict) intensity relations 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 cardinal utilitarianism studies where the agents' (usually von Neumann-Morgenstern) utility functions are assumed to have the same range –typically the unit 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 statements such as $s_i(a,b) > s_j(a,b) > 0$ that pertain to the values of agents' i, j intensity functions at the pair of alternatives (a,b) to be interpreted -under an equal-weighting assumption that we discuss- 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 –the analog of 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. Specifically, the underlying concept of *intensity-dominance* encompasses Pareto dominance by ruling out the possibility of mutually beneficial trades between any two individuals, and augments it with the additional novel requirement whereby the assignment of a pair of alternatives to a pair of individuals be made by allocating the jointly preferred alternative to the individual who prefers it *most.*

Our final application puts the concept of intensity efficiency to use in the classic *house allocation problem* of one-sided matching theory (Hylland and Zeckhauser, 1979). In particular, we follow Abdulkadiroğlu and Sönmez (1998) in assuming that agents are initially assigned -possibly at random- some artificial endowments, and show that an extension of David Gale's *top trading cycles* algorithm in the resulting *house market problem* (Shapley and Scarf, 1974) produces an intensity-efficient matching in every such problem whenever the agent's preferences and preference intensities are strict. The algorithm, which we refer to as *intensity efficiency from assigned endowments* (IEAE) assumes an ordering over the agents and an initial house assignment; then proceeds by applying top trading cycles on the house market problem that corresponds to the agents' preferences and initial assignment, hence retrieving a core allocation with respect to those preferences; and finally improves upon this allocation by using the ordering over agents to sequentially identify and -following re-assignment- remove pairs of agents where swapping the alternatives between agents is intensity dominating, *ceteris paribus.*

The remainder of the article is structured as follows. The next section introduces preference intensity relations and the relevant notation. The following section defines the benchmark class of utility-difference representations for such relations and provides a detailed discussion of several conceptual and analytical challenges that are associated with existing manifestations of this modelling approach. Motivated by this discussion, Section 4 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 also made between these models and their special cases -the *Fechnerian* and *simply scalable* models– that emerge when the intensity relations are defined by strictly stochastic binary-choice data. Section 5 analyses the above-mentioned three applications and the last section concludes. Unless otherwise noted, all proofs appear in the Appendix.

2 Preliminaries

Assumed throughout is a set X of general choice alternatives and a primitive binary relation $\succsim$ on $X \times X$ that will be thought of as a *preference intensity relation* on X.[2] The comparison $(a, b) \succsim (c, d)$ will be interpreted as suggesting that a is weakly preferred to b at least as much as c is to d, or that a is weakly inferior to b at most as much as c is to d. Although we will return to this issue in more detail below, it is worth stating from the outset that the comparison

[2]Formally, $\succsim$ can either be thought of as a binary relation on $X \times X$ or as a *quaternary relation* on X. The first economist to formally study properties of preference intensity relations modelled as quaternary relations appears to be Ragnar Frisch (1926; 1971).

$(a,b) \succsim (c,d)$ admits several possible data-driven definitions, of which the following are some examples:

1. *Questionnaires/surveys*: a is stated to be preferred to b more than c is to d;
2. *Willingness to pay*: higher for a at the choice menu $\{a,b\}$ 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*: a is chosen faster at $\{a,b\}$ than c is at $\{c,d\}$;
5. *Physical distance/effort*: more is travelled/exerted in choosing a from $\{a,b\}$ than c from $\{c,d\}$;
6. *Biological indicators*: e.g. higher neural activity is observed when choosing a at $\{a,b\}$ than c at $\{c,d\}$.

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

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

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

3 A Critique of the Utility-Difference Approach

The problem of representing a *complete and transitive* preference intensity relation by a numerical function has been approached in the literature invariably through the class of *utility-difference* models. We state below the formal definition of these representations and also distinguish between *cardinal* and *pseudo-cardinal* ones.[4]

Definition 1
A binary relation $\succsim$ *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* $a,b,c,d \in X$,

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

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

[3] Alternatively, $\gtrapprox$ could be defined by $a \gtrapprox b \Leftrightarrow (a,b) \succsim (c,c)$ for any $c \in X$. The two definitions will be equivalent in our environment.

[4] Importantly, the latter are neither cardinal nor ordinal representations but lie somewhere in between. We return to this issue below.

Cardinal-utility and general, pseudo-cardinal utility-difference representations of various origins (i.e. pertaining to riskless, risk or stochastic-choice domains) 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.[5] 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 study's motivation and focus have been put together and presented below. Although none of the points made here is really new, we have been unable to find an analogously detailed survey in the existing literature that discusses the various utility-difference models both formally and conceptually.

Riskless neoclassical cardinal utility

At the analytical level, an important limitation of the neoclassical cardinal utility model that has been appreciated thanks to its various axiomatizations (Alt, 1936, 1971; Suppes and Winet, 1955; Shapley, 1975; Köbberling, 2006) is that a particularly restrictive solvability condition must be imposed on the preference intensity relation in order for the model to be well-defined:

Solvability

If $(a,c) \succ (d,e) \succ (b,c)$*, then there is* $f \in X$ *such that* $(f,c) \sim (d,e)$*;*

This technical condition essentially forces the underlying set of alternatives to be infinite, for when the set of alternatives is specifically assumed finite, then the axiom (and the model) are known to be well-defined only when an "equal-spacing" condition is satisfied that requires preference intensity differences between any two pairs of consecutive -in the induced preference ranking- alternatives to be equivalent.[6] This equal-spacing assumption, however, trivializes the model and effectively makes it non-operational in finite domains, despite the significance of these domains for practical and welfare-relevant allocation problems.

At the conceptual level, moreover, this 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 $\succsim$ is represented by a utility function u of this kind, then cardinality of u implies that the statement $(a,b) \succsim (c,d)$ is equivalent 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

[5] For detailed and complementary accounts of these models, the debates they have been associated with and their implications for economic analysis, the reader is also 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).

[6] Krantz et al. (1971; Theorem 5, p. 168).

into the claim that a is preferred to b *exactly* r times as much as c is preferred to d, as if human perception of intensity differences were as precise as temperature differences measured by professional thermometers, for example.

We finally note another conceptually unappealing feature of this model that was first pointed out in Basu (1982): whenever a neoclassical cardinal utility index u exists and is well-behaved in the sense that its range $u(X)$ is an interval of real numbers, then the model implies that if the decision maker is able to compare first differences in utilities (equivalently, first differences in preference intensities), then she is also able to compare utility differences of the n-*th* order (equivalently, n-*th*-order differences in preference intensities), for any order n. Even for $n = 2$, however, the model's prediction that an individual is able to say that the difference in her preference intensity difference between a, b and c, d is r times higher than that between e, f and g, h seems implausibly demanding. Yet, such predictions come as necessary "extra baggage" of that model.

von Neumann-Morgenstern cardinal utility

It is by no means 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) (vNM) expected-utility model to be interpreted as indicating differences in preference intensities between those alternatives. Yet, it has been well-known among decision and welfare theorists for a long time that such utility differences confound the individuals' preference intensities over the underlying set of riskless alternatives with their attitudes toward risk (Luce and Raiffa, 1957; Schoemaker, 1982; Hammond, 1991; Ellingsen, 1994; Abdellaoui, Barrios, and Wakker, 2007; Blume, 2010; Baccelli and Mongin, 2016; Fleurbaey and Maniquet, 2018), or, as Arrow (1951, p. 10) noted when cautioning against interpersonal comparisons of vNM utilities in social choice, with *"the tastes of individuals for gambling"*. Long after the conclusion of that debate, for example, Schoemaker (1982, p. 533), articulated this insight by noting that *"preferences among lotteries are determined by at least two separate factors; namely (1) strength of preference for the consequences under certainty, and (2) attitude to risk. The [vNM] utility function is a compound mixture of these two."*. Despite its many merits as a normative model of choice under risk, therefore, on these grounds one can argue that expected utility provides an analytically convenient but conceptually unsatisfactory approach to preference-intensity modelling.

Cardinal utility from stochastic choices

Debreu's (1958) axiomatization of cardinal utility was based on comparisons of binary choice probabilities, where $p(a, \{a, b\})$ is the choice probability of a at menu $\{a, b\}$ and $p(a, \{a, b\}) +$

$p(b, \{a, b\}) = 1$. His axiomatization featured the following two axioms, which are labelled here after Suppes, Krantz, Luce, and Tversky (1989):

(Stochastic) Solvability

For any $t \in (0, 1)$ *satisfying* $p(a, \{a, b\}) \geqslant t \geqslant p(b, \{a, d\})$ *there is* $c \in X$ *such that* $p(a, \{a, c\}) = t$.

(Stochastic) Monotonicity

If $p(a, \{a, b\}) \geqslant p(a', \{a', b'\})$ *and* $p(b, \{b, c\}) \geqslant p(b', \{b', c'\})$*, then* $p(a, \{a, c\}) \geqslant p(a', \{a', c'\})$*; and if either antecedent inequality is strict, so is the conclusion.*

Similar to the case of the riskless neoclassical model, the above version of Solvability also forces the set of alternatives to be infinite in order to avoid the equal-spacing triviality mentioned earlier. Monotonicity is a special case of the *Concatenation* axiom that we discuss in its general version below.

In addition to the limitations already discussed in the context of the riskless neoclassical model, important conceptual concerns surround the stochastic-choice approach to modelling intensities. First, as pointed out in Davidson and Marschak (1959), such a cardinal utility model effectively rules out choice probabilities that are either 0 or 1. For if that was not the case, then 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, for example, then $u(\$5) = u(\$5000)$.[7] This difficulty is circumvented by simply assuming it away through the *Positivity* axiom whereby all choice probabilities are strictly positive. But one can then refer to an argument that was put forward in Baccelli and Mongin (2016) and respond that even though such genuine choice stochasticity precludes the possibility of the decision maker behaving like an ordinal utility maximizer, the model in fact portrays her as one whose behaviour is describable by maximization of a utility function that has even more refined uniqueness properties, which is somewhat paradoxical.

While these arguments do in our view challenge the capacity of choice probabilities alone to reliably convey information about an individual agent's preference intensities, we also note that current -and independent from our study- work in this area that aims to combine choice probabilities with additional information such as response times or neurophysiological variables[8] provides a more holistic and promising approach towards extracting preference-intensity infor-

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

[8]Relevant recent works include Baldassi, Cerreia-Vioglio, Maccheroni, Marinacci, and Pirazzini (2019); Webb, Levy, Lazzaro, and Glimcher (2019); Alós-Ferrer, Fehr, and Netzer (2018); Rustichini (2018).

mation from potentially "noisy"/bounded-rational decisions.

Non-cardinally-unique utility differences

In contrast to riskless neoclassical and cardinal utility from stochastic-choice data, basic (pseudo-cardinal) 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 were independently given in Scott (1964), Adams (1965) and Tversky (1967b), all of these involve a complicated Cancellation axiom. Scott's version of the axiom reads as follows:

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

While Cancellation summarizes in a remarkably compact way the need for there to be a solution to a system of linear inequalities in order for the intended representation to exist, it is impossible to understand and interpret all the behavioural restrictions that it imposes. A case in point is the way Luce and Suppes (1965, p. 277) commented on it: *"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."*

Further to that, in order to cover finite sets of arbitrary cardinalities, the number n capturing the width of the relevant sequences in the statement of Cancellation is unbounded. As such, the utility-difference model is not "finitely axiomatizable" (Scott and Suppes, 1958), even though the axiom is indeed testable in principle. As was shown in Fishburn (2001), for example, once a finite set X is fixed, the number n in the axiom's statement is bounded above by the cardinality of X by $2 \cdot |X| - 1$.

In addition to the "black-box" nature of the model, Samuelson (1938) observed[9] that a utility-difference representation necessitates the following condition on the agent's preference intensity relation:

[9]Samuelson did so in the context of the riskless neoclassical cardinal utility model, but his points apply more generally.

Concatenation

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

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 in Example 1 below, 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 this model too comes with more "baggage" than perhaps is necessary for the problem at hand.

Finally, if a preference intensity relation $\succsim$ 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.[10] This uniqueness property is between cardinal and ordinal,[11] 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?

Non-cardinally-unique utility differences from semi-ordered preferences

We conclude by noting that an alternative to rational utility-difference representations of preference intensity relations is the utility-difference-like approach that takes as primitive a *semi-order* P on X that is representable (Luce, 1956; Scott and Suppes, 1958; Scott, 1964;

[10] This is formally pinned down in Krantz et al. (1971; Theorem 2, p. 431).

[11] Specifically, all positive affine transformations of u also represent $\succsim$, but not all strictly increasing transformations do so. The permissible transformations of u therefore lie between these two extreme polar cases in a $\succsim$-dependent way. There is a conceptual analogy between these and transformations that preserve decreasing marginal utility, which also lie between the cardinal and ordinal extreme cases. Mandler (2006) provides a formal investigation of this problem.

Fishburn, 1970) by a pair (u, δ) where $u : X \to \mathbb{R}$ and $\delta > 0$ are such that

$$aPb \Longleftrightarrow u(a) - u(b) > \delta. \tag{3}$$

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. As with the pseudo-cardinal rational utility-difference model discussed above, the uniqueness properties of such representations are also between ordinal and cardinal (Roberts, 1979). 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 have often been used to define a binary relation on $X \times X$ by

$$(a, b)\widehat{P}(c, d) \Longleftrightarrow u(a) - u(b) > u(c) - u(d) > \delta. \tag{4}$$

Unlike the class of rational utility-difference models captured in (2) or the more general ones that we propose below, the decision maker here is portrayed as not always being able to rank alternatives or intensity differences between them. While, as is well-known, this implication of (3) is intuitive from a descriptive point of view when the underlying choice set X contains very similar or hard-to-compare alternatives, it has also proved fruitful in welfare-theoretic applications pertaining to axiomatisations of weighted utilitarianism (Ng, 1975; Argenziano and Gilboa, 2019) where its preference-intensity interpretation has been used explicitly as a motivation. Yet, (3) and (4) predict that the decision maker's preference and preference-intensity comparisons are incomplete under one interpretation or complete and intransitive under another, and hence not representable by a utility function in the standard sense. Along these lines, Hammond (1991, pp. 217-218) and Sen (2017/1970, pp. 146-148), for example, have expressed reservations on the conceptual appeal of such a foundation to utilitarianism. It is worth reminding the reader, finally, that in the special case where the decision maker's strict preference relation P in this model is complete/total (and hence $\delta = 0$), the pseudo-cardinal uniqueness property of the utility function u that has motivated the preference-intensity interpretations of the relevant intra- and inter-personal comparisons of its value differences disappears.

4 Preference Intensity Functions

The alternative modelling approach that we propose and analyse in this section is informed by – and reflects– the preceding critical discussion of all utility-difference approaches and circumvents the analytical and/or conceptual issues raised there. In summary, the proposed model:

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

(ii) is axiomatised 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 class of riskless pseudo-cardinal utility-difference models on finite sets as a special case in a way that is made precise;

(iv) is well-defined on infinite domains too;

(v) provides an applicable decision-theoretic foundation for equally-weighted interpersonal preference intensity comparisons that do not require interpersonal comparisons of utilities.

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* $a, b, c, d \in X$*,*

$$(a, b) \succsim (c, d) \iff s(a, b) \geqslant s(c, d) \tag{5a}$$

$$s(a, b) = -s(b, a) \tag{5b}$$

$$s(a, b), s(b, c) \geqslant 0 \implies s(a, c) \geqslant s(a, b), s(b, c), \tag{5c}$$

where s *is unique up to an* odd[12] and strictly increasing transformation *in the sense that* $t : X \times X \to \mathbb{R}$ *also represents* $\succsim$ *as in* (5) *if and only if* $t = f \circ s$ *for some* $f : \mathbb{R} \to \mathbb{R}$ *that is odd and strictly increasing in* $s(X \times X)$*.*

As captured by the order-preservation requirement (5a), 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 (5a), the *skew-symmetry* condition (5b) 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 she 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

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

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, (5c), will be referred to as *lateral consistency*. Under the maintained assumption that (5a) and (5b) 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 her preference ranking from a to b and from b to c, her 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 (5a) and (5b) alone.[13]

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.

Operational in the special case of a finite domain, 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. It will also prove particularly fruitful in other ways later on.

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* (5)*, its range* $s(X \times X)$ *is a symmetric set of consecutive integers* $\{-k, \ldots, -1, 0, 1, \ldots, k\}$.

The integer k in such a representation corresponds to the number of distinct intensity-equivalence classes of pairs of alternatives where the first element in the pair is strictly preferred to the second (or vice versa). We will return to canonical preference intensity representations in the next section when we analyse some welfare-theoretic and allocation implications of the model.

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

[13]Notice that (5a) and (5b) together imply that, for all $a, b \in X$, $a \overset{\gg}{\approx} b \Leftrightarrow (a,b) \succsim (b,a) \Leftrightarrow s(a,b) \geqslant 0 \geqslant s(b,a)$. It then follows from this and (5c) that $a \overset{\gg}{\approx} b \overset{\gg}{\approx} c \Rightarrow s(a,c) \geqslant s(c,a) \Leftrightarrow (a,c) \succsim (c,a) \Leftrightarrow a \overset{\gg}{\approx} c$.

Definition 4

A binary relation $\gtrsim$ *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* $a, b, c, d \in X$,

$$(a, b) \gtrsim (c, d) \iff g(a, b) \geqslant g(c, d) \tag{6a}$$

$$g(a, b) \geqslant g(c, d) \implies g(d, c) \geqslant g(b, a) \tag{6b}$$

$$g(a, b), g(b, c) \geqslant g(a, a) \implies g(a, c) \geqslant g(a, b), g(b, c). \tag{6c}$$

For a function g that satisfies (6) but not (5), 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 (5) becomes $a \gtrapprox b \Leftrightarrow s(a, b) \geqslant s(b, a)$ in (6). Consistent with this observation, the comparison of (5c) and (6c) 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 (6), the uniqueness property in the class of general preference intensity functions coincides with that of standard ordinal transformations.

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

Weak Order

For all $a, b, c, d \in X$, $(a, b) \gtrsim (c, d)$ *or* $(c, d) \gtrsim (a, b)$.
For all $a, b, c, d, e, f \in X$, $(a, b) \gtrsim (c, d) \gtrsim (e, f)$ *implies* $(a, b) \gtrsim (e, f)$.

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 also imposes these axioms as well as many additional ones. It is envisaged that both parts of the Weak Order axiom will be relaxed in future work.

Reversal

For all $a, b, c, d \in X$, $(a, b) \gtrsim (c, d)$ *implies* $(d, c) \gtrsim (b, a)$.

Reversal too is a standard condition and allows for the relation $\gtrsim$ 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, $(d, c) \gtrsim (b, a)$ should hold.

Consistency

For all $a, b, c \in X$, $(a, c) \gtrsim (b, c)$ *implies* $(a, b) \gtrsim (b, a)$.

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 $a, b, c, d \in X$, $(a, c) \gtrsim (b, c)$ *implies* $(a, d) \gtrsim (b, d)$.

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 $\gtrsim$ *on a finite set* $X \times X$*:*

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

The example below illustrates Theorem 1 and highlights the explanatory relevance of preference intensity functions by presenting two relations that are preference-intensity but not utility-difference representable. In the case of the first relation this is due to failure of Concatenation (hence also of Cancellation), whereas the second relation conforms with Concatenation and the axioms of Theorem 1 but still violates Cancellation in a different way. The point of these examples is to illustrate the severe restrictiveness of the utility-difference approach to intensity modelling by demonstrating that it fails to account for perfectly reasonable and internally consistent intensity orderings. As we show in Table 1 below, orderings of this kind that are left out of the utility-difference model as unexplainable are in fact the rule rather than the exception.

Example 1

Suppose that $X = \{a, b, c, d\}$ and consider two intensity relations $\succsim_1$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 $i = 1, 2$ and $a', b', c', d' \in X$. 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.

Now 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 case of relation $\succsim_1$ we have

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

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

In the case of relation $\succsim_2$ 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 $u_2(c) - u_2(b) > u_2(c) - u_2(b)$. This contradiction stems from a different violation of the Cancellation axiom that, unlike the one above, does not seem to admit a straightforward interpretation.[14]

Notice, finally, that despite the failure of the utility-difference model to account for these perfectly 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 $i = 1,2$, where the only difference between s_1 and s_2 is their values at the pairs (a,b), (b,c) and (b,a), (c,b), respectively. $\diamondsuit$

Theorem 1 characterizes 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 the utility-difference model necessitates. 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- and matching-theoretic ones that we pursue in Section 5.

4.1 Special Case 1: Utility-Difference Representations

The concept introduced next will be helpful in clarifying the formal 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 $a, b, c \in X$,*

$$s(a,c) = s(a,b) + s(b,c). \tag{7}$$

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.When this condition *is* satisfied, however, it implies both (5b) and (5c). Trian-

[14]Indeed, Cancellation and the following comparisons in $\succsim_2$ (rephrased in the permutation language of the Cancellation axiom)

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

together imply

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

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

gularly additive preference intensity functions are therefore characterized by (5a) and (7) only. This concept –unnamed, and accompanied by a severely 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 $\gtrsim$ on a finite set $X \times X$:

1. *$\gtrsim$ satisfies Completeness, Reversal and Cancellation.*
2. *$\gtrsim$ is utility-difference representable.*
3. *$\gtrsim$ 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 $s(a, b) \equiv u(a) - u(b)$ for some function $u : X \to \mathbb{R}$.[16] Although this second equivalence essentially goes back to Samuelson (1938) who established it in a cardinal-utility framework, we emphasize 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 (e.g. Phelps-Brown, 1934) had raised about the cardinal utility model, or, from a contemporary perspective, the utility-difference approach more generally. To our knowledge, the present study is the first in the literature of preference intensity modelling that takes this stand and relaxes triangular additivity with the far less demanding but still non-trivially structured lateral consistency condition.

A special class of preference intensity relations that are likely to be of interest in applications are those that, in addition to satisfying Weak Order, Reversal and Consistency, are *strict* in the following sense:

Strictness

For all $a, b, c, d \in X$, $(a, b) \sim (c, d)$ implies $(a, b) = (c, d)$ or $a = b$ and $c = d$.

[15] A similar characterization –developed independently– also appears in Adams (1965, Theorem 1).

[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.

By analogy to strict preferences, intensity relations are strict if the agent never perceives intensity differences between distinct pairs of distinct alternatives to be exactly equivalent. Considering the increased cognitive demands associated with making preference intensity comparisons, expecting the agent to be able to delineate very subtle differences so that a transitive and non-degenerate intensity-equivalence relation emerges from these comparisons may be too much to ask for. In such situations, if strict intensities across the various pairs of alternatives are sufficiently discernible (for example, when the alternatives are the possible school's that the decision maker's child might be enrolled in), Strictness may convey not only analytical convenience but also descriptive relevance.

Table 1: Exact enumeration of the distinct intensity orderings (strict and non-strict) that are compatible with utility-difference representations (second column) and general preference intensity representations (fourth column) in small choice domains.

	Weak Order Reversal Cancellation (& Strictness)	**Weak Order Reversal Concatenation (& Strictness)**	**Weak Order Reversal Consistency (& Strictness)**	***explanatory gain***
$\lvert X\rvert = 3$	25 (12)	25 (12)	37 (12)	48% (0%)
$\lvert X\rvert = 4$	723 (240)	1,011 (336)	3,903 (384)	439% (60%)
$\lvert X\rvert = 5$	63,721 (13,680)	210,361 (39,120)	5,230,801 (92,160)	8,108% (574%)

With all requisite concepts in place, we can now proceed to a formal comparison of the explanatory power between the pseudo-cardinal utility-difference model and the one proposed above. In particular, Table 1 shows how the explanatory gains of the latter model increase in the size of the choice set by comparing the total number of distinct intensity relations that satisfy Weak Order, Reversal and Consistency with the number of such relations that satisfy Weak Order, Reversal and Cancellation (in both cases, with and without Strictness).[17] While computational constraints currently limit similarly exact comparisons when there are more than five alternatives, this novel output shows that the explanatory gains of the proposed model increase super-exponentially in the cardinality of the set X when the latter lies in this small range. When there are five elements in the choice set, in particular, the proportion of intensity relations that are consistent with Weak Order and Reversal and also satisfy Cancellation is just about a mere 15% and 1% -with and without Strictness, respectively- of those where Consistency replaces Cancellation. The corresponding enumerations for the class of intensity relations that is defined by Weak Order, Reversal and *Concatenation* (with or without Strictness) and generally lies nested between these two primary ones is also shown in Table 1.[18] Notably, the output for

[17]The entries in Table 1 were confirmed by two independent computer programs that used the *Gecode* (`https://github.com/Gecode/gecode`) and *Z3Prover* (`https://github.com/Z3Prover/z3`) constraint solvers, respectively. The programs are available from the author upon request.

[18]Example 1 illustrates this computational output with two strict intensity orderings $\succsim_1$ and $\succsim_2$ that are preference-intensity representable but either (i) violate Concatenation and hence Cancellation (cf $\succsim_1$), or (ii) satisfy Concatenation but violate Cancellation (cf $\succsim_2$).

$|X| = 3$ on the one hand proves that Cancellation coincides with Concatenation under Weak Order and Reversal, and on the other hand demonstrates that these axioms are stronger than Consistency even in that special case. Finally, in addition to highlighting the much wider domain of application of the proposed preference intensity model, the computations shown in Table 1 might be thought of as providing -for the first time- some formal vindication of Samuelson's (1938) severe criticism of the utility-difference models' triangular additivity property in general -and its Concatenation implication in particular- as *"arbitrary"* and *"infinitely improbable"*.

4.2 Special Case 2: Fechnerian and Simply Scalable Stochastic Choice

The lateral consistency condition (5c) of preference intensity functions bears a structural similarity with -but is, of course, logically and conceptually distinct from- what is known in the random choice literature as *Strong Stochastic Transitivity* (SST). Before stating this 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 $p : X \times \mathcal{B}_X \rightarrow [0,1]$ 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 in Section 2, such $\succsim$ is a special case of a preference intensity relation on the set X that is defined in a data-driven way by simply taking the algebraic differences between the particular real numbers that correspond to choice probabilities.

Moreover, as was also noted previously, Debreu (1958) imposed sufficient structure on the -necessarily infinite- set X and on the binary random choice model p for the relation $\succsim$ induced by it to admit a non-trivial 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[19] 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

$$p(a, \{a, b\}) = F\big(u(a) - u(b)\big). \tag{8}$$

A more general class of binary random choice models are known as *simply 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

$$p(a, \{a, b\}) = F\big(u(a), u(b)\big). \tag{9}$$

The monotonicity properties of this function F, together with the fact that the pair (u, F) is no longer required to preserve utility differences as in (8), immediately establishes the greater generality of simply scalable relative to Fechnerian models.

The class of simply scalable binary random choice models on a finite set X was axiomatically characterized by Tversky and Russo (1969) by means of SST* and the (almost always 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 Cancellation-like *Acyclicity* axiom that we will refer to as *FIS-Acyclicity*. 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 simply scalable representations that invite simple new 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 simply scalable.*

3. *The p-induced intensity relation $\succsim$ 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.*

[19] This may be thought of as the strictly-stochastic special case to the general model of non-cardinal utility differences that was discussed earlier.

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

As also suggested by the preceding 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, simply 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 binary stochastic choice model p that induces it violates SST*.

4.3 Extension: Infinite Domains

We now turn to the case where X is an infinite set by endowing it with the structure of a connected and separable metric space. This implies that finite products of that space -once endowed with the topologies induced by their respective product metrics- are also connected and separable metric spaces. The following continuity axiom on such domains is standard and, in the special case where the primitive is a preference intensity relation on $X = \mathbb{R}_+$, it appears, for example, in Shapley (1975).

Continuity
If $(a_n, b_n) \succsim (c_n, d_n)$ for all n and $(a_n, b_n) \to (a, b)$, $(c_n, d_n) \to (c, d)$, then $(a, b) \succsim (c, d)$.

Theorem 2
The following are equivalent for a binary relation $\succsim$ on a connected metric space $X \times X$:

1. $\succsim$ satisfies Weak Order, Reversal, Consistency and Continuity.

2. $\succsim$ is representable by a continuous general preference intensity function.

Example 2

Let $X := \mathbb{R}^2_{++}$ and define $s_1, s_2, s_3 : X \times X \to \mathbb{R}$ by

$$\begin{aligned} s_1(a,b) &:= a_1a_2 - b_1b_2 \\ s_2(a,b) &:= a_1a_2 - b_1b_2 + \frac{a_1}{b_1} - \frac{b_1}{a_1} + \frac{a_2}{b_2} - \frac{b_2}{a_2} \\ s_3(a,b) &:= \frac{a_1}{b_1} - \frac{b_1}{a_1} + \frac{a_2}{b_2} - \frac{b_2}{a_2}. \end{aligned}$$

Although s_1 and s_3 emerge from s_2 by subtracting this function's ratio and product terms, respectively, we first observe that while s_1 is utility-difference decomposable because $s_1(a,b) \equiv u(a) - u(b)$ for $u(x) := x_1x_2$, the intensity functions s_2 and s_3 are not.[20] Moreover, s_1 and s_2 actually represent the same intensity relation $\succsim_1 = \succsim_2 = \succsim$, which in turn induces the symmetric Cobb-Douglas preference ordering $\gtrapprox$ on X. These claims follow by noticing that, for $i = 1, 2$ and $\beta > 0$,

$$s_i(a,b) = s_i\left(\beta a_1,\ \frac{a_2}{\beta},\ \beta b_1,\ \frac{b_2}{\beta}\right) = s_i\left(\beta,\ \frac{b_1b_2}{\beta},\ b_1,\ b_2\right) \tag{11}$$

The first equation defines the (common) intensity-equivalence classes of s_1 and s_2, while the equality between the first and third terms does so for the indifference curves of the induced preferences. Therefore, even though any positive linear transformation $\widehat{s}_1 := \alpha \cdot s_1$ of s_1 represents $\succsim$ *and* preserves its utility-difference decomposition in a cardinal manner, arbitrary strictly increasing transformations $\widetilde{s}_1 := f \circ s_1$ generally represent $\succsim$ only in the sense of (6), while any such transformation that is also *odd* represents $\succsim$ in the sense of (5). This, in particular, implies that $s_3 = f \circ s_1$ is true for some f that is odd and strictly increasing in $s_1(X \times X)$. The intensity function s_3, finally, is homogeneous of degree 0 and it is easy to verify that there exist $a, b, c, d \in X$ such that $s_1(a,b) > s_1(c,d)$ and $s_3(a,b) < s_3(c,d)$. In addition to representing an intensity relation $\succsim_3$ that is distinct from $\succsim$, Figure 1 shows that the preference relation $\gtrapprox_3$ induced by $\succsim_3$ is also distinct from $\gtrapprox$. $\Diamond$

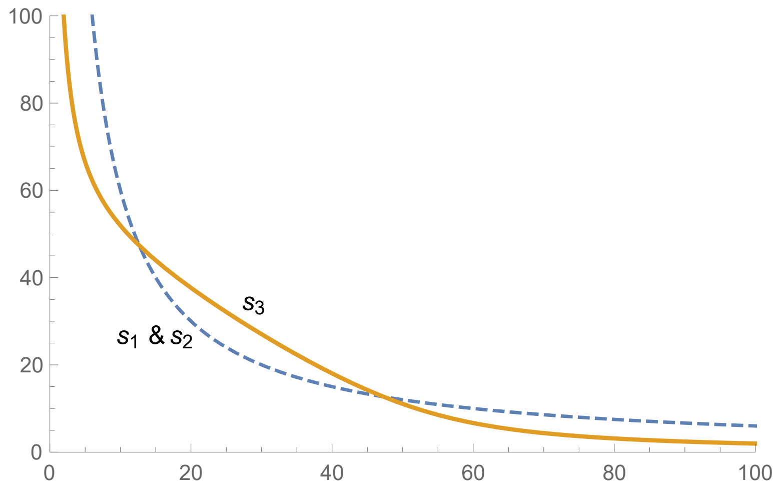
5 Applications

5.1 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

[20] Indeed, it is easy to see that $s_i(a,c) > s_i(a,b) + s_i(b,c)$ or $s_i(a,c) < s_i(a,b) + s_i(b,c)$ are generally true for both $i = 2, 3$. But since -as in the finite environment of Corollary 1- triangular additivity of each s_i is necessary and sufficient for its utility-difference decomposition in this infinite domain too (see *Sincov's* functional equation in, for example, Theorem 2, p. 356 in Aczél, 1966 or pp. 97-98 in Falmagne, 2002), the claim follows immediately.

Figure 1: The preferences induced by s_1 and s_2 (cf. dotted indifference curve) are distinct from those induced by s_3 (cf. thick indifference curve).



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 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 (2014b), for example.[21] Importantly, this is generally not an *absolute* rating that conveys information about the desirability of the chosen alternative in some pseudo-cardinal scale; instead it is a *relative* rating that provides information about the strength of preference between two alternatives chosen at two binary menus relative to the alternatives that were rejected at those menus.[22]

2. *Willingness to pay.* The decision maker can be asked to state the amount she would be willing to spend in order to receive the alternative she chose at each menu. This was envisaged, for example, 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.[23] Alternatively, as in Butler, Isoni, Loomes, and Navarro-Martinez (2014a) 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

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

[22]Absolute ratings are often elicited in empirical and experimental studies, but effectively force the decision maker to form a pseudo-cardinal ranking of the alternatives directly. Depending on the complexity of the task for the individual in question, a structured ranking elicited in this way may reflect her true intensity ordering in varying degrees of accuracy. In light of the previous analysis and the entries of Table 1, in particular, providing such ratings in an accurate way might be viewed as a more demanding task for her than providing the simpler relative ratings we consider as a primitive here. As the results of this section demonstrate, however, in the special case where these simpler relative ratings are sufficiently structured and consistent, they can also be used to recover a pseudo-cardinal ranking of the alternatives.

[23]This approach is also commonly followed in applications, for example in the very specific case where all agents are assumed to have quasilinear preferences and their differing valuations for a good are interpreted as differences in preference intensities across the agents.

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* binary stochastic choice model (DDM) and its various applications and generalizations.[24] This model postulates the existence of cardinally unique utilities for all alternatives, which are discovered by the decision maker -possibly with error- via sequential sampling. A key prediction of that model is a negative relationship between the decision maker's response time at a menu and the utility difference between the alternatives at that menu, which is in line with the intuition that the decision is easier and hence faster when the intensity difference is higher.

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. Focusing on the case where the additional observable is response times, the first study that analyzed such extended data in the deterministic revealed-preference tradition was Echenique and Saito (2017). For the testable case of finite data, these authors proposed a notion of rationalizability that builds on the pseudo-cardinal utility-difference model in (2), and identified a testable axiom on such data that is necessary and sufficient for them to be *utility-difference rationalizable*.[25] This notion requires the existence of a utility function $u : X \to \mathbb{R}$ and of a strictly decreasing[26] function $f : \mathbb{R} \to \mathbb{R}$ such that

$$C(\{a,b\}) = \{a\} \implies u(a) > u(b) \tag{12a}$$

$$u(a) - u(b) = f\big(t(a,b)\big), \tag{12b}$$

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 of a over b. The *Strong Compensation*[27] axiom that characterizes the model effectively provides the first non-parametric theoretical revealed preference intensity test for this class of behavioural data.

[24] See Ratcliff (1978); Ratcliff and McKoon (2008); Konovalov and Krajbich (2019); Baldassi, Cerreia-Vioglio, Maccheroni, Marinacci, and Pirazzini (2019); Alós-Ferrer, Fehr, and Netzer (2018) and references therein.

[25] This model was labelled *"response-time representation"* by the authors. We use the term *"utility-difference rationalization"* instead as it is more general and also in line with the broader view about potential non-choice intensity-revealing data that we adopt in this study.

[26] The case where f is instead strictly increasing was also covered.

[27] This will be defined sequentially, as follows. First, a set of sequences $\{(a_i)_{i=1}^{n_k} : k = 1,\ldots,K\}$ is a *collection of cycles* if $x_1^k = x_{n_k}^k$ and for all $i = 1,\ldots,n_k - 1$, either $a_i^k \ggcurly^B a_{i+1}^k$ or $a_{i+1}^k \ggcurly^B a_i^k$. A *(decreasing) overcompensation* on such a collection $\{(a_i)_{i=1}^{n_k} : k = 1,\ldots,K\}$ is a one-to-one function π that maps any (a_i^k, a_{i+1}^k) with $a_i^k \ggcurly^B a_{i+1}^k$ into some pair $\pi(a_i^k, a_{i+1}^k) \equiv (a_{i'}^{k'}, a_{i'+1}^{k'})$ with $a_{i'+1}^{k'} \ggcurly^B a_{i'}^{k'}$ and $r_{\{a_{i'+1}^{k'}, a_{i'}^{k'}\}} \leqslant r_{\{a_i^k, a_{i+1}^k\}}$, and, in addition, either there are i and k s.t. $r_{\pi(a_i^k, a_{i+1}^k)} < r_{\{a_i^k, a_{i+1}^k\}}$ or there are i' and k' with $a_{i'+1}^{k'} \ggcurly^B a_{i'}^{k'}$ and $(a_{i'}^{k'}, a_{i+1}^{k'}) \notin range(\pi)$. The *Strong (Decreasing) Compensation* axiom then is the requirement that there be no collection of cycles with a decreasing overcompensation.

Despite the novelty of this analysis, however, as the critical discussion of the pseudo-cardinal utility-difference model in Sections 3 and 4 (including Table 1) suggests, the existence of a utility-difference representation that lies at the heart of the utility-difference rationalization is a restrictive requirement that rules out a wide range of perfectly consistent behaviour. Moreover, similar to Cancellation, the Strong Compensation axiom does generate some interpretable restrictions which the authors highlighted, such as *(Decreasing) Monotonicity* $[t(a,c) < t(a,b), t(b,c)]$ and *Time Transitivity* $[t(a,b) \leqslant t(a',b')$ and $t(b,c) \leqslant t(b',c')$ implies $t(a,c) \leqslant t(a',c')]$, which in turn can be thought of, respectively, as the response-time analogs of Strong Stochastic Transitivity and Debreu's (1958) Stochastic Monotonicity axiom on binary stochastic choices. Yet, also similar to Cancellation, there are many additional restrictions imposed by Strong Compensation that are elusive because, again, there seems to be no simple way to summarize what the axiom says about observable behaviour.

In response to these limitations, and also to highlight the greater generality and explanatory power of the purely ordinal model of preference intensities that is laid out in (5), we will now employ this model and follow a purely constructive approach to generalize the analysis in Echenique and Saito (2017) by considerably relaxing the restrictions imposed on the data through a simple *Congruence*-like (Richter, 1966) axiom. In addition, and although this change is primarily notational and intepretational rather than formal, to facilitate applications of these non-parametric tests in experimental/empirical economics we introduce a more general and unifying definition of a behavioural dataset that allows for the additional non-choice observable to be *any* relevant menu-specific variable with intensity-revealing potential such as those mentioned above, with response times being but a special case.

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* $\{a_i, b_i\}$*, 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.*

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. Our proposed model of revealed preference intensity can now be stated formally.

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*

$a, b, c \in X$,

$$a \gg^B b \quad \Longrightarrow \quad s(a,b) > 0 \tag{13a}$$

$$s(a,b) \quad = f\left(r_{\{a,b\}}\right) = \quad -s(b,a) \tag{13b}$$

$$s(a,b), s(b,c) > 0 \quad \Longrightarrow \quad s(a,c) \geqslant s(a,b), s(b,c). \tag{13c}$$

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 general acyclic version of the lateral consistency requirement that still ensures conceptual harmony between and within revealed preferences and revealed preference intensities.

Turning to the model's empirical content, the following axiom –laid out in two versions, 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 $a_1 \gg^B a_2 \cdots \gg^B a_n$ *and* $a_i \gg^B a_{i+h}$ *for some* $i, i+h \leqslant n$ *with* $h > 1$*, then* $a_n \not\gg^B a_1$ *and* $r_{\{a_i,a_{i+h}\}} \leqslant r_{\{a_i,a_{i+1}\}}$.

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

The novel monotonicity part of the axiom predicts that the value of the intensity-revealing resource is increasing or decreasing for every sequence of alternatives where any two consecutive elements in the sequence are related by revealed preference. Specifically, for such a sequence $a_1, \ldots, a_n$ of alternatives, $a_i \gg^B a_{i+h}$ and $h > 1$ suggests that, with a_i and a_{i+h} being further apart from each other in the decision maker's revealed preference relation than a_i and a_{i+1} are, a_i is preferred to a_{i+h} no less than a_i is to a_{i+1}. This is reflected in the way in which the resource/intensity-revealing values at these menus are required to be ordered, which in turn imposes a cross-modal consistency requirement whereby the agent's revealed preferences and intensities are intuitively aligned. 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.

Theorem 3

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 3

Suppose that $X = \{a, b, c, d\}$ and consider two behavioural datasets $\mathcal{D}_1$ and $\mathcal{D}_2$ with the common binary choices shown below, and with generally distinct menu-specific 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):

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

Notice that the orderings of the $r^1_{\{\cdot\}}$ and $r^2_{\{\cdot\}}$ values are as dictated by the preference intensity relations $\gtrsim_1$ and $\gtrsim_2$ of Example 1, and that 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 a utility-difference rationalization, and hence that Strong Compensation must be violated. ♢

We conclude this analysis by clarifying the formal relationship between utility-difference and preference-intensity rationalizations.

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.

In our terminology, the equivalence between the first two statements is due to Echenique and Saito (2017). Recalling also the relevant discussion in Section 4.1, the equivalence between the last two statements is immediate.

5.2 Interpersonal Intensity Comparisons and Intensity-Efficient Allocations

In the preceding section we formally argued that a decision maker's simple preference intensity comparisons can in principle be revealed to a third party who observes that individual's choices and some additional intensity-relevant information that is also conveyed by her during the decision process. We now turn to the question of what are the welfare implications when a social planner/policy maker is assumed to have such additional information about agents. In particular, we consider a social allocation problem where monetary transfers are infeasible or prohibited, as is the case, for example, in several matching markets.[28] We therefore rule out the agents' willingness to pay as a potential source of information about their preference intensities over alternatives.[29] Among the other intensity-revealing variables that were mentioned above, choice frequencies and response times are also ruled out on the basis that they are generally associated with noisiness"/bounded rationality in the decision-making process *and* can potentially bias the intensity-informed allocation in favour of agents with higher cognitive abilities. We will therefore assume that the additional information about the agents' preference intensities comes from simple and focused questions of a purely ordinal kind, such as *"do you prefer a to b more than you prefer c to d?"*.[30]

Formally, we 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 and hence satisfies Weak Order, Reversal and Consistency. A preference intensity profile $\succsim$ is *strict* if each relation $\succsim_i$ also satisfies Strictness. The ordinary preference profile induced by $\succsim$ is denoted by $\gtrapprox$. An *allocation* is an n-tuple $x \in X^n$, where x is a permutation of the elements of X and $x_i \in X$ corresponds to the alternative allocated to agent i under x.

Recall now that, by Theorem 1, every intensity relation $\succsim_i$ in a profile $\succsim$ is representable by a canonical preference intensity function s_i. Importantly, if $\succsim$ is also strict, then we also have

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

[28]We return to this in more detail in the next section.

[29]Even if willingness to pay was a "permissible" variable, however, in order for it to be used legitimately in such a problem one would have to make the restrictive assumption that all agents have equal incomes.

[30]Analogous -but in important ways different- types of questionnaires have long been used in the happiness research literature. In this body of work, respondents state, for example, whether they feel "not too happy", "pretty happy" or "very happy", while analysts attempt to estimate from these responses the mean level of happiness across different groups under assumptions on the cardinalization of these responses and the distributions over all happiness states (see Bond and Lang, 2019). The questions that we have in mind for our purposes are qualitatively different because they do not require neither the respondents nor the analysts to identify the "utility level" of an alternative on some scale, but instead to rank-order their preference intensities on two pairs of alternatives where the first element in each pair is preferred to the second.

where k is the number of distinct pairs of distinct alternatives in X. That is, all agents' preference intensity functions are *onto* the same range of consecutive integers. This implies, in particular, that the values of the agents' canonical intensity functions at a given pair reflect the positions of those pairs in the agents' strict intensity rankings. Such a common-range representation of a profile $\succsim$ in turn allows for a novel kind of meaningful *ordinal* interpersonal comparisons of preference intensities to be assumed *without* also assuming interpersonally comparable utilities, cardinal or otherwise.

Equally Weighted Interpersonally Comparable Intensities

Given a strict preference intensity profile $\succsim = (\succsim_1, \ldots, \succsim_n)$ *that is canonically represented by* $s = (s_1, \ldots, s_n)$*, the statement*

$$s_i(a, b) > s_j(a, b) > 0$$

is assumed to imply that agent i prefers a to b more *than j does.*

Towards motivating this assumption, let us first recall that our underlying model of preference intensities at the level of the individual decision maker effectively assumes that no intensity comparison of any agent can be quantified with any precision beyond the level of an ordinal ranking. Whether this prediction (which is the polar opposite to that of perfectly quantifiable intensity comparisons that is afforded by the cardinal utility model) is descriptively accurate or not is of course ultimately an issue to be settled empirically. From a theoretical point of view, however, the important question here is whether, under the maintained assumption of simple/non-quantifiable intensity comparisons, all agents' intensity-difference rankings should be treated *equally* by the social planner or not. In particular, is it the case that, in the absence of any reliable information regarding the degree to which agents i and j would suffer if they received b instead of a, the planner should declare that i prefers a to b *more* than j does if all that she knows is that the former intensity difference lies higher in i's ranking than the latter does in j's? Since the intensity orderings convey all the available welfare-relevant information about the agents, treating them in any way other than equal would call for a justification that appears elusive. Our equal-weighting assumption might therefore be thought of as a reasonable benchmark for interpersonal comparisons in such an environment, by analogy to the uniform-prior assumption that is often made in the information economics literature.

We further note that the canonical interpersonal comparisons of *preference intensities* that are postulated in the above assumption can be thought of as constituting an ordinal analog to the standard *relative utilitarianism* assumption that rests on interpersonal comparisons of normalized von Neumann-Morgenstern *utilities* whose range is the unit interval for every agent.[31] It is worth stressing, however, that although canonical preference intensity representations can deliver an analogously unique and well-defined normalization that allows for interpersonal comparisons of intensities rather than utilities, it is an implication of well-known results [see Bossert and Weymark (2004) and Fleurbaey and Hammond (2004) and relevant references therein] that this is not the case with riskless and non-cardinally unique utility-difference representations. That is, 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 (we return to this point at the end of Example 4 below). Therefore, the interpersonal comparisons of intensities that we postulate here are operational in an enormously bigger domain of rational preference intensity profiles than those restricted to contain only utility-difference representable relations (cf Table 1), and they are also generally applicable at the same time when a similar normalization is typically impossible for the more richly structured utility-difference-representable profiles.

Retaining the equally-weighted interpersonally comparable intensities assumption throughout the rest of this section and the next, we 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 $\succsim = (\succsim_1, \ldots, \succsim_n)$ with canonical preference-intensity representation $s = (s_1, \ldots, s_n)$ if:*

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

(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 $\succsim$ if it is not intensity-dominated.*

In words, an 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[32] it to b more (less) than the agent receiving it under y;

(ii) all agents not belonging to such a pair are weakly better off at x than at y;

[31] See, for example, Dhillon and Mertens (1999) and references therein.
[32] By "*a* is *dis-preferred* to *b*" we mean that *a* is considered inferior to *b*.

(iii) at least one *inter*personal comparison in (i) or *intra*personal comparison in (ii) is strict.

Therefore, if allocations x and y are Pareto efficient and x intensity-dominates y, then the interpersonal preference trade-offs in all pairs of agents that receive the same two alternatives -in reverse order- under x and y are always resolved by x in favour of the agent in the pair who prefers the relevant alternative *most*.

Proposition 4

The following are true for any strict intensity profile $\gtrsim = (\gtrsim_1, \ldots, \gtrsim_n)$ *on a finite set* X *with* n *alternatives:*

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* $\gtrapprox$.

That intensity-efficient allocations refine Pareto efficient ones readily follows from the definition of intensity-dominance, which reduces to ordinary Pareto dominance when statement (i) is suppressed. Moreover, as it turns out, the intensity-dominance relation is generally incomplete but always transitive. This in turn ensures that the set of intensity-efficient allocations is always nonempty. Finally, as the following example illustrates, although it gives rise to a generally incomplete social ranking over the set of allocations, the refinement that intensity-dominance offers on the set of Pareto efficient allocations can be quite substantial.

Example 4

Suppose that $X = \{a, b, c, d\}$ and consider the strict intensity profile $\gtrsim = (\gtrsim_1, \gtrsim_2, \gtrsim_3, \gtrsim_4)$ that is represented canonically by

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

and induces the following preference profile $\gtrapprox$:

$$\begin{array}{ccccccc}
a & \gg_1 & b & \gg_1 & c & \gg_1 & d \\
d & \gg_2 & c & \gg_2 & b & \gg_2 & a \\
a & \gg_3 & b & \gg_3 & c & \gg_3 & d \\
d & \gg_4 & c & \gg_4 & b & \gg_4 & a.
\end{array}$$

Observe that the preferences of agents 1,3 and those of 2,4 coincide while their intensity orderings differ. Observe also that, although the intensities of agents 3 and 4 are utility-difference representable, those of agents 1 and 2 are not, due to violations of Concatenation. Note, finally, that the Pareto efficient allocations with respect to the induced preference profile $\gtrapprox$ are

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

whereas the *unique* intensity-efficient allocation with respect to the intensity profile $\gtrsim$ is z. Indeed, z intensity-dominates -and w is intensity-dominated by- all other allocations, while x and y are incomparable.

Notice, finally, that from the two intensity orderings $\gtrsim_3$ and $\gtrsim_4$ that are representable as in (2), only the utility differences of the latter relation can be put in canonical form via some such representation. Indeed, we have $u_4(d) - u_4(a) = 6$, $u_4(c) - u_4(a) = 5$, $u_4(d) - u_4(b) = 4$, $u_4(c) - u_4(b) = 3$, $u_4(b) - u_4(a) = 2$, $u_4(d) - u_4(c) = 1$ if, for example, $u_4(a) = 0$, $u_4(b) = 2$, $u_4(c) = 5$ and $u_4(d) = 6$, whereas the relevant linear system for a canonical utility-difference representation of $\gtrsim_3$ has no solution. $\Diamond$

Intensity efficiency is a normative welfare criterion. In addition to respecting Pareto efficiency, it uses the ordinal information on the agents' preference intensity relations to select allocations where, for every pair of agents who have the same preference over two alternatives, the agent in the pair who receives the commonly preferred alternative is the one who would be hurt most if she did not in fact receive it. As such, it conforms with intuitive principles of distributive justice. Moreover, this concept appears to be the first normative refinement of Pareto efficiency that is operational in an environment where neither the agents' *utilities* are required to be inter- and intra-personally comparable nor monetary transfers between agents are assumed to be feasible. Among the leading examples of normative refinements of Pareto efficiency that are applicable in settings with such richer informational structures are those suggested by the *utilitarian* and *leximin* rules.[33] The latter satisfies the *Pigou-Dalton* (Pigou, 1912; Dalton, 1920) transfer principle of equitability,[34] while a version of the former was recently shown by Piacquadio (2017) to satisfy a multi-commodity extension[35] of that principle for agents who have the same preferences. These refinements assume cardinal utility-difference representations and are therefore logically distinct from those prescribed by the intensity efficiency criterion, whose only informational requirement is purely ordinal preference intensity relations that satisfy Weak

[33] Different axiomatizations of these rules were provided, respectively, in D'Aspremont and Gevers (1977); Deschamps and Gevers (1978); Maskin (1978); Piacquadio (2017) and Hammond (1976); Deschamps and Gevers (1978); Roberts (1980), among others.
[34] See, for example, Moulin (2004).
[35] See also Fleurbaey and Maniquet (2011).

Order, Reversal, Consistency and Strictness. Finally, unlike riskless versions of the utilitarian and leximin models, because intensity efficiency is applicable on a finite set of arbitrary choice alternatives that may well consist of indivisible goods, it is a potentially relevant property in allocation and matching problems where the analyst is able to elicit information about the agents' preferences as well as their preference intensities. We now turn to the investigation of a classic such problem through this lens.

5.3 The Intensity-Efficient Allocation of Individuals to Positions

Existing matching mechanisms that have explicitly aimed to incorporate the agents' generally differing preference intensities over the assignable objects have done so by introducing randomness in the allocation process and by subsequently assuming that intensities can be captured accurately by cardinal vNM utilities. Important examples include the *Boston*; *pseudo-market* (Hylland and Zeckhauser, 1979; Budish, 2011); and *choice-augmented deferred-acceptance* (Abdulkadiroğlu, Che, and Yasuda, 2015, 2011; Gale and Shapley, 1962) mechanisms. The practical limitations of the cardinal-utility assumption that are associated with them, however, has not gone unnoticed by matching theorists. While discussing the pseudo-market mechanism, for example, Che (2013, p. 80) pointed out that *"[it] requires agents to formulate their cardinal preferences. This can be challenging for the agents, and a mistake in formulating preferences may result in misallocation."* This legitimate concern, which is reinforced and empirically substantiated in Budish and Kessler (2018), is *in addition* to the critical remarks raised in Section 3 regarding the general unsuitability of vNM cardinal utilities to accurately reflect the agents' intensities at the more basic individual decision-theoretic level. At the same time, even when this assumption is made, it is not sufficient to ensure -in the context of these mechanisms- that it is a dominant strategy for the agents to reveal their intensities truthfully. In the case of the Boston mechanism, in particular, the exact opposite has in fact been shown to be true in practice (Abdulkadiroğlu and Sönmez, 2003).[36] Motivated by these observations, in this section we propose a simple algorithm that delivers an intensity-efficient matching by assuming only that the agents form and truthfully revel purely ordinal preference intensity relations of the kind studied above. We also argue that such ordinal preference-intensity elicitation can take place in a minimally demanding way that can help avoid mistakes in the formulation and reporting of these relations from the agents.

We start by recalling that, in the *house allocation problem* of Hylland and Zeckhauser (1979), a set of agents are to be assigned to a set of indivisible goods ("houses"). A well-known and widely used matching mechanism for this problem is *random serial dictatorship* (RSD), whereby a pri-

[36]See also the experimental findings in Klijn, Pais, and Vorsatz (2013).

ority ordering over the agents is drawn uniform-randomly, following which the agents -according to their position in the ordering- are sequentially assigned their most preferred house from those that remain in the market. Assuming that the agents have strict preferences over houses and expected-utility preferences over random allocations, this mechanism satisfies strategy-proofness and ex post Pareto efficiency. Abdulkadiroğlu and Sönmez (1998) analysed the house allocation problem by treating it as a generalization of Shapley and Scarf's (1974) *house market problem*. The latter problem differs from the former only in the structure of property rights: every agent comes to the market initially endowed with a house. As shown in Shapley and Scarf (1974), application of David Gale's *top trading cycles* (TTC) algorithm in every such market results in a core allocation. Moreover, as shown in Roth and Postlewaite (1977), TTC corresponds to the unique matching in the core of every house market, while as shown in Roth (1982), the core as a direct mechanism is strategy-proof. Noticing that, upon fixing an artificial initial allocation (possibly randomly) of the houses to the agents one could subsequently analyse the house allocation problem as a house market problem, Abdulkadiroğlu and Sönmez (1998) showed that this *core from assigned endowments* matching mechanism is in fact equivalent to RSD.

In what follows, we borrow from Abdulkadiroğlu and Sönmez (1998) the idea of treating a house allocation problem as a house market problem by assigning agents with some artificial initial endowment. Unlike these authors and the above-cited literature more generally, however, we don't assume that the agents have expected-utility (or indeed, any) preferences over random allocations. Instead, we assume that they have the kind of simple preference intensity relations over houses that were studied above. In so doing, we effectively assume away the question of strategy-proofness in our version of the problem. However, given the well-documented tensions between various notions of efficiency and strategy-proofness that pertain to the house allocation problem (Zhou, 1990; Bogomolnaia and Moulin, 2001; Ehlers, 2002; Budish, 2011; Nesterov, 2017; Zhang, 2019), and considering also that our interest here is on whether there exists a matching that always produces intensity-efficient -hence Pareto efficient- allocations in a risk-free setting, this omission might be thought of as a reasonable simplifying assumption for the purposes of this normative analysis that focuses on the criterion encapsulated by intensity efficiency. Yet, if such a matching mechanism does exist, then whether or not this omission is detrimental for the purposes of applying it in practical problems clearly depends on whether it is manipulable at the preference and/or preference-intensity level in an *obvious* way (as was the case with the Boston mechanism, for example) that would invite systematic misrepresentation of this information by the agents, which would indeed render its practical application unreliable and hence undesirable.

We adapt the notation in Abdulkadiroğlu and Sönmez (2013) to describe and analyse the house allocation problem with *strict ordinal preference intensities*. In particular, the problem we consider consists of a triple $\langle I, H, \succsim \rangle$, where $I = \{1, \ldots, n\}$ is a set of agents, H is a set of indivisible objects ("houses"), and $\succsim = (\succsim_1, \ldots, \succsim_n)$ is a preference intensity profile on H for which, as in the previous section, every $\succsim_i$ is assumed to satisfy Weak Order, Reversal, Consistency and Strictness. Again, the (strict) preference profile that is induced by $\succsim$ is denoted by $\gg$. Finally, we assume that $|I| = |H| = n$. The outcome of a house allocation problem is an assignment of houses to agents so that each agent receives exactly one house. The assignment is formally described by a *matching* function $\mu : I \to H$ that is one-to-one and onto. Recall that, by Theorem 1 and the Strictness assumption, the intensity relation of each agent i is represented canonically by some onto preference intensity function $s_i : X \times X \to \mathbb{R}$ with $s_i(X \times X) = \{-k, \ldots, -1, 0, 1, \ldots, k\}$. Let $s = (s_1, \ldots, s_n)$ be the profile of canonical preference intensity functions corresponding to $\succsim$. A matching μ on $\langle I, H, \succsim \rangle$ is *intensity-efficient* if $\big(\mu(1), \ldots, \mu(n)\big) \in X^n$ is an intensity-efficient allocation with respect to $\succsim$ or, equivalently, s.

We now introduce an algorithm that defines a matching mechanism for the problem we consider, which we refer to as *intensity efficiency from assigned endowments* (IEAE).

Step 1. Fix a strict ordering on I, and let any allocation $z = (z_1, \ldots, z_n) \in X^n$ be thought of as assigning item z_j to the j-th agent in that ordering.

Step 2. Fix a matching μ that specifies the agents' artificial initial endowments. This step defines the house market problem $\langle I, H, \gg, \mu \rangle$.

Steps $3, \ldots, k$. Apply the TTC algorithm on the house market problem $\langle I, H, \gg, \mu \rangle$. That is, start by letting every agent j point to the agent that is endowed with j's unique most preferred item (possibly j's own endowment). There is at least one ordered list (a *"cycle"*) of agents such that every agent points to the next one in the list, and the last agent points to the first. Implement the exchanges suggested by each such cycle, remove the respective agents from the market and proceed by repeating this process, each time with the remaining agents. Let ν be the (core) matching that results when this TTC algorithm terminates.

Steps $k+1, \ldots, k+s$. For the house market problem $\langle I, H, \succsim, \nu \rangle$ with initial endowments as in ν:

(i) Search for the first agent $j \neq 1$ such that $s_1(\nu(j), \nu(1)) > s_j(\nu(j), \nu(1))$. If such an agent j exists, swap the items assigned to 1 and j and remove both agents from the market; then repeat this step starting with the first remaining agent.

(ii) If such an agent $j \neq 1$ does not exist, search for the first agent $j \neq 2$ such that $s_2(\nu(j), \nu(2)) > s_j(\nu(j), \nu(2))$. If such an agent j exists, repeat the above step, then move to agent 3, and so on.

(iii) Stop when no pair of agents i, j with $s_i(\nu(j), \nu(i)) > s_j(\nu(j), \nu(i))$ remains in the market.

Proposition 5

The IEAE algorithm specifies an intensity-efficient matching in every house allocation problem with a strict preference intensity profile.

Indeed, note first that once an initial matching μ is specified in Step 2, Gale's TTC algorithm applied on the derived house market problem with the induced preference profile is guaranteed to generate a core matching ν in a finite number of steps. Moreover, since the number of agents in the market is finite and *either* some pair of agents is removed in each of the steps of the IEAE algorithm that follow TTC *or* the search is forced to stop if it identifies no pairs of remaining agents for whom the assigned items are swappable, IEAE terminates in finitely many steps at some allocation κ that generally differs from ν. It remains to be shown that κ is intensity-efficient. Suppose to the contrary that there is a pair of agents i, j such that $s_i(\kappa(j), \kappa(i)) > s_j(\kappa(j), \kappa(i))$. Without loss of generality, suppose $i < j$. By part $(k+1, \ldots, k+s)-(iii)$ of the IEAE algorithm, κ cannot be the terminal matching. In particular, by part $(k+1, \ldots, k+s)-(i)$, there exists a step in the swap-and-remove process where items $\kappa(j)$ and $\kappa(i)$ are swapped between agents j and i. A contradiction therefore obtains.

Example 5

Let $I = \{1, 2, 3, 4\}$ and consider the intensity profile of Example 3.

Step 1: Assume the ordering $1 \gg 2 \gg 3 \gg 4$ over I.

Step 2: For simplicity, let the initial matching μ specify the core allocation (a, c, b, d) (cf allocation w in Example 3).

Step 3: Observe that $s_2(\mu(4), \mu(2)) = s_2(d, c) = 5 > 1 = s_4(d, c) = s_4(\mu(4), \mu(2))$.
Swap houses in the pair $(2, 4)$ and remove this pair from the market.

Step 4: Observe that $s_3(\mu(1), \mu(3)) = s_3(a, b) = 4 > 2 = s_1(a, b) = s_1(\mu(1), \mu(3))$.
Swap houses in the pair $(1, 3)$ and remove this pair from the market.

The resulting matching κ specifies the (unique) intensity-efficient allocation (b, d, a, c) (cf allocation z in Example 3).[37] $\Diamond$

As far as the related issues of complexity in the agents' ability to formulate and report their preference intensities and the practical implementation of IEAE are concerned, one of the possible procedures that may achieve the latter by minimizing the former could be along the following lines. The matching platform takes the form of a computer program (e.g. a web application)

[37] Evidently, if multiple intensity-efficient allocations are associated with an intensity profile $\succsim$, then the one selected by the IEAE algorithm is sensitive to the agent ordering and initial assignment specified in Steps 1 and 2, respectively.

where n agents are registered for the allocation of n objects. Following registration, everyone is informed about the ordering over agents and the artificial initial assignment that are required by the first two steps of the algorithm. At that point, each agent i is asked to submit their strict preference ranking over the n objects:

$$i(1) \ggcurly_i i(2) \ggcurly_i \ldots \ggcurly_i i(n).$$

Following receipt and processing of agent i's preference ordering, the platform then prompts agent i to also submit information about her preference intensities by sequentially presenting her between the following pairs of objects and asking whether or not she strictly prefers the first to the second element in the first pair strictly more than she strictly prefers the first to the second element in the second pair:

$$\begin{array}{ccc} \big(i(1), i(2)\big) & \& & \big(i(2), i(3)\big) \\ \big(i(2), i(3)\big) & \& & \big(i(3), i(4)\big) \end{array}$$

If $(i(1), i(2)) \succ_i (i(2), i(3))$ and $(i(2), i(3)) \succ_i (i(3), i(4))$ are the agent's responses, then under the Weak Order, Reversal and Consistency axioms that are implicitly assumed and enforced by the implementing platform in the background, this partial intensity ordering over $i(1), i(2), i(3), i(4)$ is automatically extended into the complete ordering that satisfies $(i(1), i(4)) \succ_i (i(1), i(2))$, $(i(2), i(3))$, $(i(3), i(4))$, $(i(2), i(4))$, and $(i(2), i(4)) \succ_i (i(2), i(3))$, $(i(3), i(4))$. In that case, the platform proceeds by asking the agent to respond to the next question about her preference intensity in $\big(i(3), i(4)\big)$ vs. $\big(i(4), i(5)\big)$, and then continues in a similar fashion until the intensity ordering is extended over all n objects. On the other hand, if the agent's responses to the first two questions are, for example, $(i(1), i(2)) \succ_i (i(2), i(3))$ and $(i(3), i(4)) \succ_i (i(2), i(3))$, then, to achieve the desired properties on the agent's intensity ordering, the platform instead proceeds by asking for more information on her intensity comparisons, namely for her to rank $\big(i(1), i(3)\big)$ vs. $\big(i(3), i(4)\big)$ before completing the intensity ranking over $i(1), i(2), i(3), i(4)$ and then proceeding to eliciting information for the next two pairs. A matching platform that is programmed to operate in this way will eventually define complete intensity rankings with the required properties for all agents. From these orderings, finally, the agents' canonical preference intensity functions can also be readily computed and ultimately applied in the IEAE algorithm as described above.

This procedure is structured in a way that makes preference and preference-intensity elicitation complete, internally consistent and minimally demanding on the agent by combining information on her preferences over the objects in the first stage -which one expects to be easier for the agent to understand and report accurately- with as little additional information as possible on her intensity differences between the choice objects in the second stage. Despite this structure, however, such a matching platform is effectively no more paternalistic than existing ones that elicit ordinal preference information only. In the latter case, by asking participating agents to submit their preferences over the objects directly in the form of a strictly ordered list, as in the first stage above, the platform forces agents to submit a complete and internally consistent strict ranking over the objects, even if some of the same agents' self-reported preferences violated these conditions had they been expressed through a different elicitation method. Similarly, the hereby envisioned platform rules out inconsistencies within the agents' preference intensity rankings as well as across their preference and preference-intensity rankings by taking at face value the information on their preferences that was elicited in the first stage and by using this information in conjunction with Weak Order, Consistency and Reversal on the partial intensity comparisons that are elicited in the second stage also in order to arrive at a complete and internally consistent strict intensity ranking, even if the agent's self-reported such ranking is, in principle, also sensitive to the elicitation method.

The approach to the house allocation problem that we are following here is to our knowledge the first that focuses on a refinement of Pareto efficiency that incorporates information on the agents' preference intensities without assuming existence -or requiring elicitation- of cardinal utility functions. Yet, our motivating concern that Pareto efficiency is a weak and potentially unfair normative requirement in matching-theoretic problems has also been raised explicitly in this literature, for example in Che, Gale, and Kim (2013), Lee and Yariv (2018) and Pycia and Ünver (2018). A key methodological difference in this regard between these authors' approach and ours is that we do not employ any social welfare aggregation method that builds on interpersonally comparable utilities, cardinal or otherwise, in order to define a social ranking over matchings. Instead, our approach assumes interpersonally comparability of the agents' ordinal strict preference intensity relations -which is operationalized through the canonical normalization of their preference intensity functions- and puts this interpersonal comparability to use by means of the intensity efficiency refinement of the Pareto criterion. This requires a different kind of information from the agents in addition to their ordinal preferences, and is therefore logically distinct from any efficiency refinement that is derived from any notion of social welfare aggregation.

6 Concluding Remarks

The existing decision-theoretic apparatus for preference-intensity modelling -primarily developed in the 1930'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 generality, simplicity, tractability and transparency in its behavioural foundations. This study has aimed to contribute towards filling this gap by analysing the novel model of preference intensity functions that was claimed to have these desirable features, and also by identifying its empirically testable behavioural content in the spirit of the revealed-preference analysis tradition. The model can be thought of as the simplest analog 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 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. The canonical normalization afforded by the model motivates –under an equal-weighting assumption– a simple way of making purely ordinal interpersonal comparisons of preference intensities that do not presuppose interpersonal comparisons of utilities. This in turn motivates the novel notion of intensity efficiency that has been shown to be well-defined and to refine Pareto efficiency by discarding allocations that are dominated on intensity-difference grounds. The notion has intuitive normative properties and appears suitable for applications in matching-theoretic or other allocation problems where monetary transfers are infeasible or undesirable. It has finally been shown that an intensity-efficient matching in the house allocation problem with assigned endowments exists whenever the agents' intensity relations are strict. Looking ahead, of special interest in this literature may be to investigate the existence of intensity-efficient matchings in other problems where standard sought-after criteria such as the core are not always well-defined. We leave such investigations for future work.

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.

Abdulkadiroğlu, A., Y.-K. Che, and Y. Yasuda (2011): "Resolving Conflicting Preferences in School Choice," *American Economic Review*, 101, 399–410.

——— (2015): "Expanding "Choice" in School Choice," *American Economic Journal: Microeconomics*, 7, 1–42.

Abdulkadiroğlu, A. and T. Sönmez (1998): "Random Serial Dictatorship and the Core from Random Endowments in House Allocation Problems," *Econometrica*, 66, 689–701.

——— (2003): "School Choice: A Mechanism Design Approach," *American Economic Review*, 93, 729–747.

——— (2013): "Matching Markets: Theory and Practice," in *Advances in Economics and Econometrics, Volume I: Economic Theory*, ed. by D. Acemoglu, M. Arellano, and E. Dekel, New York: Cambridge University Press, 3–47.

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.

Aguiar, V. H. and R. Serrano (2019): "Cardinal Revealed Preference, Price-Dependent Utility, and Consistent Binary Choice," *working paper*.

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, New York: Harcourt Brace Jovanovich, chap. 20, (English translation of Alt (1936) by S. Schach).

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

Arrow, K. J. (1951): *Social Choice and Individual Values*, New York: Wiley.

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

Baldassi, C., S. Cerreia-Vioglio, F. Maccheroni, M. Marinacci, and M. Pirazzini (2019): "A Behavioral Characterization of the Drift Diffusion Model and its Multi-alternative Extension for Choice under Time Pressure," *Management Science*, forthcoming.

Basu, K. (1982): "Determinateness of the Utility Function: Revisiting a Controversy of the Thirties," *Review of Economic Studies*, 49, 307–311.

Blume, L. (2010): "Review of Ken Binmore's "Rational Decisions"," *Economic Journal*, 120, F157–F161.

Bogomolnaia, A. and H. Moulin (2001): "A New Solution to the Random Assignment Problem," *Journal of Economic Theory*, 100, 295–328.

Bond, T. N. and K. Lang (2019): "The Sad Truth About Happiness Scales," *Journal of Political Economy*, 127, 1629–1640.

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, Dordrecht: Kluwer, 1099–1177.

Bridges, D. S. and G. B. Mehta (1995): *Representations of Preference Orderings*, Berlin: Springer.

Budish, E. (2011): "The Combinatorial Assignment Problem: Approximate Competitive Equilibrium from Equal Incomes," *Journal of Political Economy*, 119, 1061–1103.

Budish, E. and J. B. Kessler (2018): "Bringing Real Market Participants' Real Preferences into the Lab: An Experiment that Changed the Course Allocation Mechanism at Wharton," *working paper.*

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

Campbell, D. E. (1973): "Social Choice and Intensity of Preference," *Journal of Political Economy*, 81, 211–218.

Caplin, A. (2014): "Rational Inattention and Revealed Preference: The Data-Theoretic Approach to Economic Modelling," *Research in Economics*, 68, 295–305.

Casella, A. and A. Gelman (2008): "A Simple Scheme to Improve the Efficiency of Referenda," *Journal of Public Economics*, 92, 2240–2261.

Casella, A. and T. Palfrey (2019): "Trading Votes for Votes: A Dynamic Theory," *Econometrica*, 87, 631–652.

Casella, A., T. Palfrey, and R. G. Riezman (2008): "Storable Votes," *Quarterly Journal of Political Science*, 3, 165–200.

Che, Y.-K. (2013): "Discussion of "Matching Markets: Theory and Practice"," in *Advances in Economics and Econometrics, Volume I: Economic Theory*, ed. by D. Acemoglu, M. Arellano, and E. Dekel, New York: Cambridge University Press, 76–85.

Che, Y.-K., I. Gale, and J. Kim (2013): "Assigning Resources to Budget-Constrained Agents," *Review of Economic Studies*, 80, 73–107.

Dahl, R. A. (2006/1956): "Equality, Diversity and Intensity," in *A Preface to Democratic Theory*, ed. by R. A. Dahl, Chicago: University of Chicago Press, 90–123.

Dalton, H. (1920): "The Measurement of the Inequality of Incomes," *Economic Journal*, 30, 348–361.

D'Aspremont, C. and L. Gevers (1977): "Equity and the Informational Basis of Collective Choice," *Review of Economic Studies*, 44, 199–209.

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, New York: Wiley, 233–269.

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

DeMeyer, F. and C. R. Plott (1971): "A Welfare Function Using 'Relative Intensity' of Preference," *Quarterly Journal of Economics*, 85, 179–186.

Deschamps, R. and L. Gevers (1978): "Leximin and Utilitarian Rules: A Joint Characterization," *Journal of Economic Theory*, 17, 143–163.

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.

Ehlers, L. (2002): "Coalition Strategy-Proof House Allocation," *Journal of Economic Theory*, 105, 298–317.

Eilenberg, S. (1941): "Ordered Topological Spaces," *American Journal of Mathematics*, 63, 39–45.

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

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

Fishburn, P. C. (1970): "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, Dordrecht: Kluwer, 1179–1285.

Fleurbaey, M. and F. Maniquet (2011): *A Theory of Fairness and Social Welfare*, Econometric Society Monographs, Cambridge: Cambridge University Press.

——— (2018): "Optimal Income Taxation Theory and Principles of Fairness," *Journal of Economic Literature*, 56, 1029–1079.

Fleurbaey, M. and P. Mongin (2005): "The News of the Death of Welfare Economics Is Greatly Exaggerated," *Social Choice and Welfare*, 25, 381–418.

Frisch, R. (1926): "Sur un Problème d'Économie Pure," *Norsk Matematisk Forenings Skrifter*, I-16, 1–40.

——— (1971): "On A Problem in Pure Economics," in *Preferences, Utility and Demand*, ed. by J. S. Chipman, L. Hurwicz, M. K. Richter, and H. F. Sonnenschein, New York: Harcourt Brace Jovanovich, chap. 19, (English translation of Frisch (1926) by J. S. Chipman).

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

Gale, D. and L. S. Shapley (1962): "College Admissions and the Stability of Marriage," *American Mathematical Monthly*, 69, 9–15.

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, Cambridge: Cambridge University Press, 200–254.

Hammond, P. J. (1976): "Equity, Arrow's Conditions and Rawls' Difference Principle," *Econometrica*, 44, 793–804.

Harsanyi, J. C. (1955): "Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility," *Journal of Political Economy*, 63, 309–321.

——— (1990): "Interpersonal Utility Comparisons," in *Utility and Probability*, ed. by J. Eatwell, M. Milgate, and P. Newman, London: Palgrave Macmillan.

Hylland, A. and R. Zeckhauser (1979): "The Efficient Allocation of Individuals to Positions," *Journal of Political Economy*, 87, 293–314.

Klijn, F., J. Pais, and M. Vorsatz (2013): "Preference Intensities and Risk Aversion in School Choice: A Laboratory Experiment," *Experimental Economics*, 16, 1–22.

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

Konovalov, A. and I. Krajbich (2019): "Revealed Strength of Preference: Inference from Response Times," *Judgment and Decision Making*, 14, 381–394.

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

Lee, S. and L. Yariv (2018): "On the Efficiency of Stable Matchings in Large Markets," *working paper.*

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, New York: Wiley, 249–410.

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.

Maskin, E. (1978): "A Theorem on Utilitarianism," *Review of Economic Studies*, 45, 93–96.

Mosteller, F. and P. Nogee (1951): "An Experimental Measurement of Utility," *Journal of Political Economy*, 59, 371–404.

Moulin, H. J. (2004): *Fair Division and Collective Welfare*, Cambridge, MA: MIT Press.

Nesterov, A. S. (2017): "Fairness and Efficiency in Strategy-Proof Object Allocation Mechanisms," *Journal of Economic Theory*, 170, 145–168.

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

Phelps-Brown, E. H. (1934): "Notes on the Determinateness of the Utility Function," *Review of Economic Studies*, 2, 66–69.

Piacquadio, P. G. (2017): "A Fairness Justification of Utilitarianism," *Econometrica*, 85, 1261–1276.

Pigou, A. C. (1912): *Wealth and Welfare*, London: Macmillan.

Pycia, M. and M. U. Ünver (2018): "Arrovian Efficiency in Allocation of Discrete Resources," *working paper.*

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, 635–645.

Roberts, F. S. (1979): *Measurement Theory with Applications to Decisionmaking, Utility and the Social Sciences*, vol. 7 of *Encyclopedia of Mathematics and its Applications (editor: Gian-Carlo Rota)*, Reading, MA: Addison-Wesley.

Roberts, K. W. S. (1980): "Possibility Theorems with Interpersonally Comparable Welfare Levels," *Review of Economic Studies*, 47, 409–420.

Roth, A. E. (1982): "Incentive Compatibility in a Market with Indivisibilities," *Economics Letters*, 9, 127–132.

Roth, A. E. and A. Postlewaite (1977): "Weak versus Strong Domination in a Market with Indivisible Goods," *Journal of Mathematical Economics*, 4, 131–137.

Roth, A. E. and M. A. O. Sotomayor (1990): *Two-Sided Matching: A Study in Game-Theoretic Modelling and Analysis*, Econometric Society Monographs, Cambridge: Cambridge University Press.

Rustichini, A. (2018): "Stochastic Choice, Time to Choose and Individual Characteristics," *working paper.*

Sákovics, J. (2013): "Revealed Cardinal Preference," *Economic Theory Bulletin*, 1, 39–45.

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

Schoemaker, P. J. H. (1982): "The Expected Utility Model: Its Variants, Purposes, Evidence and Limitations," *Journal of Economic Literature*, 20, 529–563.

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.

Sen, A. (2017/1970): *Collective Choice and Social Welfare*, UK: Penguin, (expanded edition).

Shapley, L. and H. Scarf (1974): “On Cores and Indivisibility,” *Journal of Mathematical Economics*, 1, 23–37.

Shapley, L. S. (1975): “Cardinal Utility from Intensity Comparisons,” *Technical Report R-1683-PR, The RAND Corporation.*

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

Suppes, P. and M. Winet (1955): “An Axiomatization of Utility Based on the Notion of Utility Differences,” *Management Science*, 1, 259–270.

Tversky, A. (1967a): “Additivity, Utility and Subjective Probability,” *Journal of Mathematical Psychology*, 4, 175–201.

——— (1967b): “A General Theory of Polynomial Conjoint Measurement,” *Journal of Mathematical Psychology*, 4, 1–20.

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 ed.

Webb, R., I. Levy, S. C. Lazzaro, and P. W. Glimcher (2019): “Neural Random Utility: Relating Cardinal Neural Observables to Stochastic Choice Behavior,” *Journal of Neuroscience, Psychology and Economics*, 12, 45–72.

Zhang, J. (2019): “Efficient and Fair Assignment Mechanisms Are Strongly Group Manipulable,” *Journal of Economic Theory*, 180, 167–177.

Zhou, L. (1990): “On a Conjecture by Gale about One-Sided Matching Problems,” *Journal of Economic Theory*, 52, 123–135.

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, $(b, a) \nsucc (d, c)$ 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

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

and

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

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 (5b) by construction. We will show that s also satisfies (5a), and it will then follow from Claim 3 that s obeys (5c) 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 $\gtrsim$ 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 $(a_j, a_l) \gtrsim (a_m, a_n)$ 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 $(a_n, a_m) \gtrsim (a_l, a_j)$. 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 $a_j, a_l, a_m, a_n \in X$, $(a_j, a_l) \gtrsim (a_m, a_n)$ 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 $(a_j, a_l) \not\gtrsim (a_m, a_n)$. Since $\gtrsim$ 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 $\gtrsim$ as in (5). Moreover, by construction, $s(X \times X)$ is a symmetric set of consecutive integers. Hence, s constitutes a canonical preference intensity representation of $\gtrsim$.

$3 \Rightarrow 4$. 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 $\gtrsim$ 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 $(a, b) \gtrsim (c, d)$. 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 $\gtrsim$ 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 $\gtrsim$. 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 $\gtrsim$. 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 $\gtrsim$ as in (5), then s also represents $\gtrsim$ as in (6). 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 $\gtrsim$ as in (6).

$5 \Rightarrow 1$. Weak Order is implied by (6a) and Reversal is implied by (6b). To show that Consistency is also implied by (6), 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 that, from (6a) and (6b), $g(b,a) > g(a,b)$ 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 (6b), $g(b,a) > g(a,a) > g(c,a)$. It follows then that $g(b,a), g(a,c) > g(a,a)$ and, by (6c), $g(b,c) > g(a,c)$. 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 (6c), $g(b,a) > g(a,a)$ 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 (6b), this implies $g(c,b) > g(a,a)$. By (6c), moreover, this and $g(b,a) > g(a,a)$ together imply $g(c,a) > g(c,b)$. By (6b) 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 (6c), $g(b,a) > g(a,a)$ and $g(a,c) > g(a,a)$ implies $g(b,c) > g(b,a)$, a contradiction. ■

Proof of Corollary 1.

1 ⇔ 2. See Theorem 3.2 in Scott (1964).

2 ⇒ 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 *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 $\succsim$ is represented by the triangularly additive preference intensity function $s : X \times X \to \mathbb{R}$. We have

$$\begin{aligned} a \gtrapprox b &\iff s(a,b) \geqslant s(b,a) \\ &\iff s(a,b) \geqslant 0 \\ &\iff s(a,b) + s(b,z) \geqslant s(b,z) \\ &\iff s(a,z) \geqslant s(b,z) \end{aligned} \tag{14}$$

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

$$u(a) := s(a,z). \tag{15}$$

It follows from (14) and (15) that

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

Thus,

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

■

Proof of Corollary 2.

1 ⇔ 2: See Tversky and Russo (1969).

2 ⇔ 3: Omitted (simple). ■

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.

1 ⇒ 2. Since $X^2 \times X^2$ is connected, it follows from Eilenberg's Theorem (Eilenberg, 1941; Bridges and Mehta, 1995) that Weak Order and Continuity together imply the existence of a continuous $s : X \times X \to \mathbb{R}$ such that $(a,b) \succsim (c,d) \Leftrightarrow s(a,b) \geqslant s(c,d)$. This fact and Reversal further imply that $s(d,c) \geqslant s(b,a)$ whenever $s(a,b) \geqslant s(c,d)$. Finally, given the above and Consistency, it follows from Claim 3 that $s(a,b), s(b,c) \geqslant s(d,d) \Rightarrow s(a,c) \geqslant s(a,b), s(b,c)$.

2 ⇒ 1. Continuity of $\succsim$ is directly implied by continuity of s. The remaining argument is as in the last part of the proof of Theorem 1. ■

Proof of Theorem 3.

With a slight abuse of notation, we write $\{a,b\} \in \mathcal{D}$ and $r_{\{a,b\}} \in \mathcal{D}$ 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 $a \not\gg^B b$ and $b \not\gg^B a$ (to be written $a \parallel^B b$), we 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 $a_i \gg^B a_{i+1}$ for all $i = 1, \ldots, k-1$ or $a_{i+1} \gg^B a_i$ for all $i = 1, \ldots, k-1$ (the latter two situations will be denoted by $a \gg^{\hat{B}} b$ and $b \gg^{\hat{B}} a$, respectively). Moreover, if $a \gg^B b$ or $b \gg^B a$, we write $[a, b] := \{a, b\}$. Finally, if $a \not\gg^{\hat{B}} b$ and $b \not\gg^{\hat{B}} a$, we write $a \parallel^{\hat{B}} b$. In light of these definitions, we have

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

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

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

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

$$s(a, b) := \begin{cases} \left(\min\limits_{\{a_i\}_{i=1}^k \in [a,b]} \min\limits_{\substack{\{a_i, a_{i+l}\} \in \mathcal{D}, \\ i+l \leqslant k}} r_{\{a_i, a_{i+l}\}} \right)^{-1} - \epsilon, & \text{if } a \gg^{\hat{B}} b \\ 0, & \text{if } a \parallel^{\hat{B}} b \\ \epsilon - \left(\min\limits_{\{a_i\}_{i=1}^k \in [a,b]} \min\limits_{\substack{\{a_i, a_{i+l}\} \in \mathcal{D}, \\ i+l \leqslant k}} r_{\{a_i, a_{i+l}\}} \right)^{-1}, & \text{if } b \gg^{\hat{B}} a \end{cases} \tag{16}$$

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

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

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 $s(a,b) > 0 \Leftrightarrow a \ggcurly^{\hat{B}} b$ also hold by construction (for the latter claim, recall also the definition of ϵ). The above implies, in particular, that $a \ggcurly^{B} b \Rightarrow s(a,b) > 0$. Therefore, (13a) and (13b) 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 $a \ggcurly^{\hat{B}} b$ and $b \ggcurly^{\hat{B}} c$, which in turn implies $a \ggcurly^{\hat{B}} c$. Therefore, $s(a,c) \geqslant 0$. 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

$$\min_{\{x_i\}_{i=1}^{k} \in [a,c]} \min_{\substack{\{x_i,x_{i+l}\} \in \mathcal{D} \\ i+l \leqslant k}} r_{\{x_i,x_{i+l}\}} \leqslant \min_{\{y_i\}_{i=1}^{m} \in [a,b]} \min_{\substack{\{y_i,y_{i+l}\} \in \mathcal{D} \\ i+l \leqslant m}} r_{\{y_i,y_{i+l}\}}, \min_{\{z_i\}_{i=1}^{n} \in [b,c]} \min_{\substack{\{z_i,z_{i+l}\} \in \mathcal{D} \\ i+l \leqslant n}} r_{\{z_i,z_{i+l}\}}. \tag{17}$$

It is immediate that the minimum operators in (17) are redundant if $[a,b] = \{a,b\}$, $[b,c] = \{b,c\}$ and $[a,c] = \{a,c\}$, i.e. when $a \ggcurly^{B} b \ggcurly^{B} c$ and $a \ggcurly^{B} c$ are true. Similarly, the claim that (17) 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

$$\{a_i\}_{i=1}^{k} = \{b_i\}_{i=1}^{m} \cup \{c_i\}_{i=1}^{n} \qquad \text{for some } \{b_i\}_{i=1}^{m} \in [a,b],\ \{c_i\}_{i=1}^{n} \in [b,c]. \tag{18}$$

Suppose to the contrary that

$$\min_{\{a_i\}_{i=1}^{k} \in [a,c]} \min_{\substack{\{a_i,a_{i+l}\} \in \mathcal{D} \\ i+l \leqslant k}} r_{\{a_i,a_{i+l}\}} > \min_{\{b_i\}_{i=1}^{m} \in [a,b]} \min_{\substack{\{b_i,b_{i+l}\} \in \mathcal{D} \\ i+l \leqslant m}} r_{\{b_i,b_{i+l}\}}. \tag{19}$$

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 $a' \ggcurly^{B} b'$. By (18) 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 (19). The case where $[a,b]$ is replaced by $[b,c]$ on the right hand side of (19) is similarly ruled out. Therefore, (17) 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 4.

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 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 (5c) that $s_i(x_i, z_i) \geqslant 0$. Thus, the second condition in Definition 8 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 (5c), 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. But this is obviously true because, by definition, intensity dominance encompasses Pareto dominance. ■

Proof of Proposition 5.

In the main text. ■