RUHR
ECONOMIC PAPERS

Nicole Becker
Kirsten Häger
Jan Heufer

Revealed Notions of Distributive Justice I – Theory

Ruhr Graduate School in Economics
technische universität dortmund

#443

Imprint

Ruhr Economic Papers

Published by

Ruhr-Universität Bochum (RUB), Department of Economics
Universitätsstr. 150, 44801 Bochum, Germany

Technische Universität Dortmund, Department of Economic and Social Sciences
Vogelpothsweg 87, 44227 Dortmund, Germany

Universität Duisburg-Essen, Department of Economics
Universitätsstr. 12, 45117 Essen, Germany

Rheinisch-Westfälisches Institut für Wirtschaftsforschung (RWI)
Hohenzollernstr. 1-3, 45128 Essen, Germany

Editors

Editorial Office

Ruhr Economic Papers #443

Responsible Editor: Wolfgang Leininger

ISSN 1864-4872 (online) – ISBN 978-3-86788-500-3

Ruhr Economic Papers #443

Nicole Becker, Kirsten Häger, and Jan Heufer

Revealed Notions of Distributive Justice I – Theory

Ruhr Graduate School in Economics
technische universität dortmund

Bibliografische Informationen der Deutschen Nationalbibliothek

Die Deutsche Bibliothek verzeichnet diese Publikation in der deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über: *http://dnb.d-nb.de* abrufbar.

http://dx.doi.org/10.4419/86788500
ISSN 1864-4872 (online)
ISBN 978-3-86788-500-3

Nicole Becker, Kirsten Häger, and Jan Heufer[1]

Revealed Notions of Distributive Justice I – Theory

Abstract

We provide a framework to decompose preferences into a notion of distributive justice and a selfishness part and to recover individual notions of distributive justice from data collected in appropriately designed experiments. "Dictator games" with varying transfer rates used in Andreoni and Miller (2002) and Fisman et al. (2007) can be used to assess individuals' preferences, but – with the help of simple new axioms – also to recover some part of individuals' notion of justice. "Social planner" experiments or experiments under a "veil of ignorance" (Rawls 1971) can be used to recover larger parts of the notion of justice. The axioms also allow a simple test for the validity of such an experimental approach, which is not necessarily incentive-compatible, and to recover a greater part of an individual's preference relation in dictator experiments than before. Interpersonal comparison of the individual intensity of justice (or fairness) similar to the suggestions in Karni and Safra (2002b) are possible, and we can evaluate the intensity based on an individual's own notion of justice. The approach is kept completely non-parametric. As such, this article is in the spirit of Varian (1982) and Karni and Safra (2002a).

JEL Classification: C14, C91, D11, D12, D63, D64

Keywords: Altruism; distributive justice; nonparametric analysis; preference decomposition; revealed preference; social preferences

October 2013

1 Nicole Becker, RGS Econ and TU Dortmund University, Kirsten Häger, Friedrich Schiller University Jena, Jan Heufer, TU Dortmund University – Views expressed represent exclusively the authors' own opinions. *Thanks to Wolfgang Leininger for his support. Thanks to Christian Bredemeier, Katharina Eckartz, Dirk Engelmann, Jörg Franke, Yiqu an Gu, Burkhard Hehenkamp, Linda Hirt, Oliver Kirchkamp, Eliane Lambertz, and Karl Schlag for helpful comments and discussions, and to the participants of the European Meeting of the Econometric Society 2011 in Oslo, the Dortmund Brown Bag Seminar 2011, the Brown Bag Seminar 2012 at the Max Planck Institute of Economics, and the Research Institute of Industrial Economics (IFN) in Stockholm 2013. Funding from TU Dortmund University, the Ruhr Graduate School in Economics, and Friedrich Schiller University Jena is gratefully acknowledged.* *– All correspondence to: Jan Heufer, TU Dortmund University, Department of Economics and Social Science, 44221 Dortmund, Germany. E-Mail: jan.heufer@tu-dortmund.de*

1 INTRODUCTION

1.1 Overview

This article provides a framework in the spirit of Varian (1982) and Karni and Safra (2002a) to decompose preferences into a *notion of distributive justice* (or simply *notion of justice*) and a *selfishness* part and to recover individual notions of distributive justice from data collected in induced budget experiments, such as "dictator games" with varying transfer rates. With the help of simple new axioms, it is also possible to recover a greater part of an individual's preference relation than before (e.g., in the analysis of Andreoni and Miller 2002 and Fisman et al. 2007). The methodology allows for interpersonal comparison of the individual strength of justice similar to the suggestions in Karni and Safra (2002b) and for the definition of a money-metric equivalence to measure this strength.

Afriat (1967) shows how to construct utility functions from expenditure data and provides a condition, called cyclic consistency, which is both necessary and sufficient for the existence of a non-trivial utility function which is maximised by the observed choices of a consumer. Varian (1982) introduces a condition called the Generalised Axiom of Revealed Preference (GARP), which is equivalent to Afriat's condition. Varian shows how to use the revealed preference framework to recover all that can be said about a consumer's preference if one accepts GARP as a solid basis for demand analysis.

Andreoni and Miller (2002) use Varian's GARP to analyse data collected in an experimental "dictator game" with varying transfer rates. They apply the demand theory framework, treating donations and own payoff as two different goods. The transfer rate thus corresponds to prices, and the choice sets participants were asked to make decisions on correspond to standard competitive budget sets. The data generated in such an *induced budget experiment* is, by all practical means, just as the theory of revealed preference assumes. A very similar experiment was conducted by Fisman et al. (2007). In Choi et al. (2007), the authors use Varian's methods to recover preferences graphically by constructing upper and lower bounds of indifference curves and the examples they offer neatly illustrate different types of preferences that can be observed.

Karni and Safra (2002a) introduce an axiomatic framework for the choices of individuals who are both self-interested and motivated by some notion of justice. Their model concerns choices over random allocation procedures for some indivisible prize. They decompose preferences according to which an individual chooses allocation procedures into a self-interest and moral value judgement component and provide conditions for representation of preferences by (additively separable) utility functions. Heufer (2013b) uses a similar framework for a non-parametric analysis of choices over probabilities and provides methods to recover preferences in the spirit of Varian's analysis.

In this article, the existing theoretical and experimental literature is combined. The analysis is based on the idea that a participant's preference can be recovered in a dictator experiment[1] (cf. Andreoni and Miller 2002) and that a participant's individual notion of distributive justice

[1]The first dictator experiment was conducted by Forsythe et al. (1994). Since then this kind of experiment is often called a "dictator game", as it can be thought of as an ultimatum game (see, e.g., Güth and Tietz 1990, Güth et al. 2003) without veto power. It is therefore a degenerate game without strategic interaction, which is why we prefer not to call it a game.

can be recovered using an appropriately designed social planner experiment (cf. Dickinson and Tiefenthaler 2002, Traub et al. 2009) or an experiment similar to the dictator experiment, but behind a "veil of ignorance" (cf. Harsanyi 1955, Rawls 1971). Here, an individual's notion of distributive justice is basically an impartial preference about payoff distributions which is unbiased by the individual's preferences about his own payoff. The analysis is based on two simple axioms: *Symmetry* of the notion of justice, based on the idea that a decision maker in an anonymous experiment is impartial with respect to the receivers of payoff, and *Agreement* of the preference with the notion of justice and own payoff, i.e. a participant prefers an allocation if (not only if) the allocation is both more just and offers a higher own payoff than an alternative. The Agreement axiom also allows an equivalent interpretation: if a participant prefers an allocation even though this allocation offers a lower own payoff than an alternative, this must be because the allocation is more just.

The axioms allow to recover a large part of a participant's preference and notion of justice. In fact, in dictator experiments, they allow to recover some parts of the participants' notion of justice and a larger part of the preference than before. Combining dictator and social planner experiments, the approach provides a new way to compare the intensity of an individual's sense of justice which complements previous analysis.

Part II of this investigation will focus on testing a participant's notion of justice and his actual preference separately in a single experiment. This will make it possible to measure more precisely than before how strong a sense of justice of a participant is, to analyse different notions of justice in detail, and to recover a larger part of participants' preferences than in previous experiments. The analysis in the article is the theoretical starting point for such examinations.

1.2 Related Literature

This paper is related to the experimental literature on impartial notions of distributive justice in simple distribution experiments, specifically the comparative performance of several well known distribution mechanisms, the comparison of impartial and self-biased notions of distributive justice, the introduction and evaluation of parametric and non-parametric models to analyse individual preferences, and to the literature on dictator experiments with varying transfer rates.

Yaari and Bar-Hillel (1984) address the topic of distributive justice in the context of ethical judgements. In their experiment participants were asked to choose between the outcomes of several distribution mechanisms within a pre-specified framework. The participants were not paid for or affected by their choices, so the results are unbiased by selfishness but should be interpreted with caution.[2] Varying only the context of the distributional situation but holding the mathematical representation constant, choices are expected to be identical in the different treatments. The results, however, show substantial differences: In the 'needs' treatment most participants repeatedly choose the maxmin (Rawlsian) solution, whereas in the 'tastes' treatment efficiency (maximising the sum of all payoffs) is the most prominent motive.

Dickinson and Tiefenthaler (2002) provide the results of a modified dictator experiment. In their experiment, the dictator is asked to choose an allocation of inputs for two unknown

[2]The authors replicated their results after assigning the same experimental task to associates with economic background.

individuals which is then transferred according to two asymmetric payoff functions. More than half of the dictators choose inputs that equalize final payoffs and eleven percent choose inputs which maximize joint final payoffs. The authors did not include a standard dictator experiment. They also did not test if the participants' choices could be the result of utility maximisation. Traub et al. (2009) include one social planner treatment in their experiment, in which a participant has to choose between two income distributions which then became effective for other participants.

Engelmann and Strobel (2004) compare different notions of distributive justice in three-player dictator experiments with role uncertainty. Each player is asked to make a choice between three allocations. Subsequently the choice of one randomly chosen participant (the dictator) per group is implemented. In order to assess an impartial notion of distributive justice, the income of the dictator is kept constant while the different preference motives are isolated by systematically varying the income of the other two participants. Comparing the performance of the models of Fehr and Schmidt (1999, henceforth FS) and Bolton and Ockenfels (2000) with maxmin (Rawlsian) preferences, selfishness, and efficiency concerns, their data shows that individual preferences can be rationalised to a great extent by a combination of maxmin preferences, selfishness, and efficiency concerns while inequity aversion and FS preferences have only little explanatory power. This result is confirmed by Engelmann and Strobel (2007) who refer to their results in an extensive internet experiment and who provide an overview of the experimental literature on distributional preferences in dictator experiments.

In their experiment Cappelen et al. (2010) compare the fairness ideas of two groups of participants: impartial spectators and stakeholders. They address the question whether individuals favour equality of ex ante-opportunities or equality of ex-post outcomes. They implemented a two-stage design, where on the first stage stakeholders engage in a risk-taking phase and in stage two both parties are asked to redistribute the joint income of two individuals respectively. Their data shows that for both groups the notions of distributive justice diverge: Spectators focus more on ex-post equality whereas stakeholders favour equality of ex-ante opportunities.

Cappelen et al. (2007) introduce a parametric model to analyse participants' fairness ideals and the weight they attach to their own income and their fairness ideal. Experimental evidence from their two-stage dictator experiment is presented where on the first stage production takes place followed by a distribution phase on the second stage. Exploring their data the authors are able to show the prevalence of multiple fairness ideals: egalitarianism, liberal egalitarianism, and libertarianism. They further conclude that an impartial notion of distributive justice alone is not able to fully account for an individual's choice behaviour but that a mixture of several motives is.

Cox et al. (2008) develop a non-parametric model to analyse an individual's preferences over own monetary payoffs and payoff of others. They introduce a partial ordering called "more altruistic then", based on marginal rates of substitution and willingness to pay. They interpret their model with common parametric models and analyse the observable magnitudes of their model in two player sequential games.

In the context of induced budget experiments with varying transfer rates Tan and Bolle (2006) introduce a parametric model to analyse a participant's notions of selfishness, altruism, and inequality aversion. Experimental results of four dictator experiments are presented and show that fairness motives and altruism co-exist. They further conclude that varying transfer rates have important implications: A transfer rate of less than one induces fairness motives whereas a transfer rate greater than one leads to fairness violations. As the transfer rate increases so does the amount of giving. A transfer rate of zero induces money burning.

1.3 Outline

The rest of the article is organised as follows. In Section 2 we provide theoretical preliminaries: In Section 2.2 preferences, notions of justice, and the central axioms are defined. In Section 2.3 we describe the experimental setup we have in mind. In Section 2.4 revealed preferences and revealed notions of justice are defined, based on the data collected in an experiment. In Section 2.5 we show how the application of the central axioms can be used to extend the revealed relations. Section 3 provides further tools for data analysis. In particular, in Section 3.1 rationalisability of the data by well behaved functions is discussed. In Section 3.2 we show how the data can be used to recover large parts of a participant's preference and notion of justice. In Section 3.3 we show how interpersonal comparisons are possible based on intersections of "revealed preferred" and "revealed less just" sets. Furthermore, we define money metric functions which can be used to measure the individual strength of the notion of justice. Section 4 concludes. The main analysis focuses on the two-dimensional case, which simplifies the analysis and is sufficient for most applications. The higher dimensional case is treated in Appendix A. All proofs can be found in the Appendix B.

2 THEORETICAL FOUNDATIONS: PRELIMINARIES

2.1 General Definitions

Let $\mathbb{A} = \mathbb{R}^L_+$.[3] Let $N_\varepsilon(x) = \{a \in \mathbb{A} : d(a,x) < \varepsilon\}$ be the open epsilon neighbourhood of $x \in \mathbb{A}$, where $d : \mathbb{A} \times \mathbb{A} \to \mathbb{R}_+$ is the Euclidean distance function. Let $\sigma(a)$ be the set of all $L!$ permutations of the elements of the vector a, and let $\sigma_\ell(a)$, $\ell = 1, \ldots, L!$ denote the respective permutation, with $\sigma_1(a) = a$.

For any set $S \subseteq \mathbb{A}$, the interior of S, denoted int$S$, is the set of all points $a \in S$ for which there exists an $\varepsilon > 0$ such that $N_\varepsilon(a) \subset S$. The closure of S, denoted cl$S$, is the set of all $a \in \mathbb{A}$ such that for all $\varepsilon > 0$, $N_\varepsilon(a) \cap S \neq \emptyset$. The boundary of a set S is denoted ∂S.

For a finite set $\{s^1, \ldots, s^n\} = S \subset \mathbb{A}$, the *convex hull* of S is defined as

$$CH(S) = \left\{ a \in \mathbb{A} : \exists \lambda \in [0,1]^n, \sum_{i=1}^n \lambda_i = 1, x = \sum_{i=1}^n \lambda_i s^i \right\},$$

and the *convex monotonic hull* is

$$CMH(S) = CH\left(\left\{a \in \mathbb{A} : a \geq s^i \text{ for some } i = 1, \ldots, n\right\}\right).$$

[3] We use the following notation: For all $x, y \in \mathbb{R}^L$, $L \geq 2$, we denote $x \geqq y$ if $x_i \geq y_i$ for all $i = 1, \ldots, L$; $x \geq y$ if $x \geqq y$ and $x \neq y$; $x > y$ if $x_i > y_i$ for all $i = 1, \ldots, L$. We denote $\mathbb{R}^L_+ = \{x \in \mathbb{R}^L : x \geqq (0, \ldots, 0)\}$ and $\mathbb{R}^L_{++} = \{x \in \mathbb{R}^L : x > (0, \ldots, 0)\}$. Note that we adopt the convention to use superscripts for indices and subscripts for coordinates.

A binary relation Q on $\mathbb{A}$ is a set of ordered pairs of elements of $\mathbb{A}$. We will mostly use the usual notation, $x \mathrm{Q} y$ instead of $(x, y) \in \mathrm{Q}$, but for some concepts it is helpful to remember that Q is a set. A binary relation Q is *transitive* if whenever $x \mathrm{Q} y$ and $y \mathrm{Q} z$ then $(x, z) \in \mathrm{Q}$; *complete* if for all $x, y \in \mathbb{A}$, either $x \mathrm{Q} y$ or $y \mathrm{Q} x$; *continuous* if for all $x \in \mathbb{A}$ the sets $\{a \in \mathbb{A} : x \mathrm{Q} a\}$ and $\{a \in \mathbb{A} : a \mathrm{Q} x\}$ are closed; *convex*[4] if for all $x, y, z \in \mathbb{A}$ with $x \mathrm{Q} z$ and $y \mathrm{Q} z$, $([1 - \mu]x + \mu y) \mathrm{Q} z$ for all $\mu \in [0, 1]$; *monotonic* if for all $x, y \in \mathbb{A}$ with $x \geq y$, $x \mathrm{Q} y$. Let $\mathcal{L}$ denote the set of elements $a \in A$ such that $x \mathrm{Q} a$, that is, $\mathcal{L}(\mathrm{Q}, x) = \{a \in \mathbb{A} : x \mathrm{Q} a\}$. Reversely, let $\mathcal{U}$ be defined as $\mathcal{U}(\mathrm{Q}, x) = \{x \in \mathbb{A} : a \mathrm{Q} x\}$.

2.2 Preferences, Notions of Distributive Justice, and Two Central Axioms

Our hypothesis concerns the preference and notion of distributive justice underlying the decision making of a participant—the *decision maker* (DM)—who is asked to allocate money between two or more individuals, one of whom may be the DM. We assume that there are $L \geq 2$ individuals. In the context of consumer theory, $\mathbb{A}$ is usually called the commodity space, and an element $a \in \mathbb{A}$ is called a consumption bundle. In our case, we call $\mathbb{A}$ the *payoff space* and an element $a \in \mathbb{A}$ an *allocation*. Whenever the DM who is asked to make the allocation is one of these individuals, his payoff is given by a_1, and $a_2, \ldots, a_L$ is the payoff of the remaining individuals.

Our hypothesis is that a DM can be represented by transitive, complete, and continuous binary relations on $\mathbb{A}$. The first relation, $\succsim \in \mathbb{A} \times \mathbb{A}$, represents his actual *preference* according to which he decides whenever he has his own monetary stake in the choice situation (i.e., whenever his own payoff is given by a_1). The second relation, $\succsim_{\mathrm{J}} \in \mathbb{A} \times \mathbb{A}$, represents his *notion of distributive justice*, or simply *notion of justice*. The interpretation of his preference is as usual, i.e. $x \succsim y$, means that to the DM x is at least as good as y, while $x \succsim_{\mathrm{J}} y$ means that the DM considers x to be at least as just as y. The strict preference and justice relations, $\succ$ and $\succ_{\mathrm{J}}$, and the indifference and iso-justice relations, $\sim$ and $\sim_{\mathrm{J}}$, are defined as usual, that is, as the asymmetric and symmetric parts of $\succsim$ and $\succsim_{\mathrm{J}}$, respectively.

We will also use a third transitive, complete, continuous binary relation which we will call the *self payoff* relation, denoted $\succsim_{\mathrm{S}} \in \mathbb{A} \times \mathbb{A}$. This relation is the same for all DMs; we will try to recover a subset of the relations $\succsim$ and $\succsim_{\mathrm{J}}$ empirically, but not $\succsim_{\mathrm{S}}$. The relation $\succsim_{\mathrm{S}}$ is defined as $x \succsim_{\mathrm{S}} y$ if $x_1 \geq y_1$, with $\succ_{\mathrm{S}}$ and $\sim_{\mathrm{S}}$ as its asymmetric and symmetric part, respectively. If the DM has no own monetary stake, i.e. if none of the elements of $x \in \mathbb{A}$ represent his own payoff, then $x \sim_{\mathrm{S}} y$ for all $x, y \in \mathbb{A}$. Below we will also introduce the assumption that both $\succsim$ and $\succsim_{\mathrm{J}}$ are convex and monotonic. Monotonicity in particular requires some explicit justification in the context of this paper which we will provide below.

We now introduce two simple axioms about the preferences. At the end of this subsection, we will provide an example based on CES-preferences that illustrates the axioms and which provides a parametric interpretation. One of these axioms will be presented in two equivalent forms, which is helpful for the interpretation. The first axiom is called *Symmetry* (SY) and concerns the justice relation. It is based on the idea that DMs are not aware about the identity of the other individuals and that no individual is per se more deserving than others. The second axiom, called *Agreement* (AG), postulates that if an allocation x is at least as just as y and x gives the DM a higher own payoff than y, then to the DM x is at least as good as y; that is, if both $\succsim_{\mathrm{J}}$ and $\succ_{\mathrm{S}}$ agree, than $\succsim$ agrees with

[4] A convex preferences is sometimes also called quasi-concave.

them as well. An equivalent reformulation of this axiom then postulates that if a DM prefers x over y but y gives the DM a higher payoff, then the reason why the DM prefers x must be because to the DM x is more just than y. The two formulations are referred to as AG1 and AG2.

$$(a, \sigma_\ell(a)) \in \succsim_J \qquad \text{for all} \quad \ell = 1, \dots, L! \qquad \text{and all} \quad a \in \mathbb{A} \tag{SY}$$

$$[\succsim_J \cap \succ_S] \subseteq \succsim \tag{AG1}$$

$$[\succsim \cap \prec_S] \subseteq \succsim_J \tag{AG2}$$

Note that AG1 states that $x \succsim_J y$ and $x \succ_S y$ implies $x \succsim y$, and AG2 states that $x \succsim y$ and $y \succ_S x$ implies $x \succsim_J y$. See Figure 1 for an illustration; note that the indifference curves are drawn assuming monotonicity and convexity of preferences. Axiom AG formalises the idea that the preference is *between* pure selfishness and pure justice. The reason why we use the strict relation $\succ_S$ is that given the possibility that $\succsim = \succsim_S$, the AG1 and AG2 would not be equivalent if they were based on $\succsim_S$.[5] Given our definition of AG, the two versions are indeed equivalent, as the following fact shows. All proofs can be found in the appendix.

Fact 1 *AG1 and AG2 are equivalent.*

The Axiom SY can also be interpreted as a special case of a "more deserving than" relation: A more general formulation would account for different degrees of deservingness of individuals, whereas here all individuals are equally deserving. Note that, combined with monotonicity, SY is a special case of second order stochastic dominance: If an $a \in \mathbb{A}$ is interpreted as a portfolio of assets of which only one will pay off, then—assuming equal probabilities for all assets—second order stochastic dominance together with monotonicity is equivalent to SY; see also Heufer (2011).

We also assume that both the preference and the notion of justice are convex and monotonic. While convexity is a plausible and usual assumption, monotonicity—in particular of the notion of justice—may require some justification. In a revealed preference framework based on choices from competitive budget sets, monotonicity is not falsifiable, as Afriat's Theorem (see below) shows. Furthermore, as shown below in Fact 3, monotonicity of $\succsim$ in conjunction with SY and AG implies monotonicity of $\succsim_J$. It also extends AG as shown in Fact 2. When we recover revealed preferred (more just) and worse (less just) sets, abandoning monotonicity raises the question with what kind of condition it should be replaced. Assuming satiated preferences or notions of justice (e.g., a preference with a "bliss point") appears unreasonable. One candidate would be first order stochastic dominance or a kind of mean-variance preference to reflect possible inequality aversion. But then again, when choices are to be made on budget sets, the allocation $x \in \partial B$ with $x_i = x_j$ for all $i, j \in \{1, \dots, L\}$ is always available. We therefore believe that for the purpose of this paper, monotonicity is plausible enough and not too restrictive. See Heufer (2012b, 2013b) for an analysis of revealed preference data without monotonicity.

Fact 2 *If $\succsim$ and $\succsim_J$ are monotonic*

- *and $\succsim_J \neq \succsim_S$ or $\succsim_J$ satisfies SY, then $[\succsim_J \cap \succ_S] \subseteq \succsim$ is equivalent to $[\succsim_J \cap \succsim_S] \subset \succsim$;*

[5]Thanks to Dirk Engelmann for pointing out the close relationship between the two axioms used in a previous draft of this paper, which led to the current definition of AG.

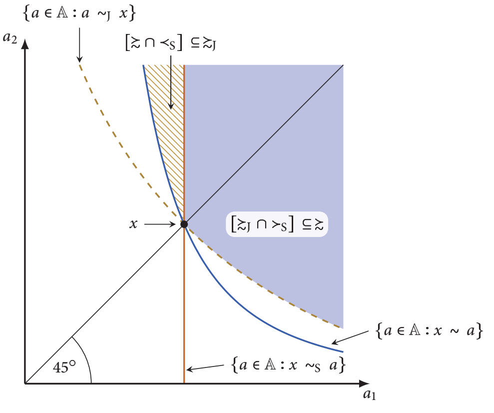


Figure 1: Reasoning about preferences and notions of justice: If an allocation offers a higher payoff to the DM and is also more just than an alternative allocation, then it must also be preferred. If it is preferred even though it offers a lower payoff to the DM, then it must be more just.

- *then* $[\succsim \cap \prec_S] \subseteq \succsim_J$ *is equivalent to* $[\succsim \cap \precsim_S] \subset \succsim_J$.

Fact 3 *If $\succsim_J$ is monotonic, then Sy and Ag together imply that $\succsim$ is monotonic. If $\succsim$ is monotonic, then Sy and Ag together imply that $\succsim_J$ is monotonic.*

Fig. 2 shows an example of what we can learn about the notion of justice based on Ag2 and Sy if we know a DM's preferences, based on a function $u(a) = a_1^{3/5} a_2^{2/5}$ representing the preference. Fig. 2.(a) shows an indifference curve and an isojustice curve based on the function $v(a) = a_1^{1/2} a_2^{1/2}$ representing the notion of justice. It can be easily verified that these preferences satisfy the Ag axiom. Fig. 2.(b) shows what we can learn about the "more just" and "less just" region of the indicated allocation. In Fig. 2.(c), symmetry is added; in Fig. 2.(d), monotonicity is added. Fig 2.(e) fills some gaps based on the assumption that $\succsim_J$ is convex. Finally, Fig 2.(f) combines the sets constructed in (b) - (e) and also shows the isojustice curve to demonstrate that we can learn quite a lot about the notion of justice.

Note that the construction shown in Figure 2 also indicates something about the *strength* of a DM's *sense of justice*. We will define this concept formally in Section 3 based on the ideas in Karni and Safra (2002b). For now, think of the strength of the sense of justice as the weight the DM's attaches to his notion of justice when trading off own payoff against justice. The construction in Figure 2 is based on taking the DM's preference $\succsim$ and "pulling out" the selfishness part $\succsim_S$. We can then analyse what we are left with. In the example shown in the figure, there is quite a bit left. Now imagine instead that the DM's preferences are perfectly selfish, that is, $\succsim = \succsim_S$. In that case, the indifference curves would be straight vertical lines, and pulling out $\succsim_S$ would leave us with no information at all about the DM's notion of justice. Generally, the more we learn about the DM's notion of justice based on the Ag-axiom, the stronger his sense of justice, and vice versa.

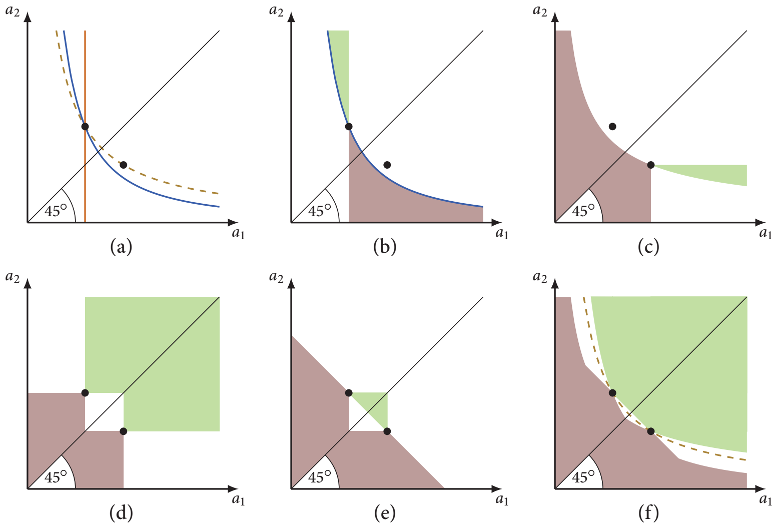


Figure 2: The construction of bounds on the notion of justice based on the preference, the axioms, and in (d) and (e) also on monotonicity and convexity.

We say that a utility function $u : \mathbb{A} \to \mathbb{R}$ *represents* a preference $\succsim$ if $u(a) \geq u(a')$ whenever $a \succsim a'$. Similarly, a distributive justice function, or simply justice function, $v : \mathbb{A} \to \mathbb{R}$ represents a notion of distributive justice $\succsim_J$ if $v(a) \geq v(a')$ whenever $a \succsim_J a'$.

Consider the CES-family (*constant elasticity of substitution*) of utility functions for the two-dimensional case: Suppose $\succsim$ can be represented by $u(a) = (\alpha a_1^r + [1-\alpha]a_2^r)^{1/r}$, and $\succsim_J$ can be represented by $v(a) = (\beta a_1^s + [1-\beta]a_2^s)^{1/s}$, with $\alpha, \beta \in [0,1]$ and $r, s \leq 1$. Andreoni and Miller (2002) use data from a dictator experiment to estimate parameters of u, and interpret α as an indication of selfishness and r as an indicator of the convexity of preferences via the elasticity of substitution $\eta_S = 1/(r-1)$. Similarly, the parameter β can be interpreted as the weight a DM puts on the payoff of the first individual, while s captures the convexity and indicates the willingness of the DM to trade off total payoff and equality.

Clearly, $\succsim_J$ satisfies Sy if and only if $\beta = 1/2$ (i.e., if and only if the DM places the same weight on both individuals). Then Ag implies $\alpha \in [1/2, 1]$, that is, the DM attaches at least the same weight to his own payoff as to the payoff of other individuals. We can then show that Ag and Sy are equivalent to $r = s$, that is, the elasticity of substitution is the same for the utility and the justice function. This also holds in the general higher dimensional case with more than two individuals, where Sy implies that in the utility function, the weight attached to all individuals other than the DM are the same. Thus, for the CES model, the axioms postulate that the trade-off between total payoff and equality is the same irrespective of whether or not the DM has a personal monetary stake. What differs is the weight attached to the DM's own payoff. Therefore, we can interpret r and s as a measure for equality concerns. If two different DM's then do indeed both satisfy Ag, their respective αs alone indicate how selfish they are.

Fact 4 *Suppose $\succsim$ and $\succsim_J$ are represented by CES functions, with $\tilde{u}(a) = \left(\sum_{i=1}^{L} \alpha_i a_i^r\right)^{1/r}$ for $\succsim$ and $\tilde{v}(a) = \left(\sum_{i=1}^{L} \beta_i a_i^s\right)^{1/s}$ for $\succsim_J$, $r, s \leq 1$ and $\alpha_i, \beta_i > 0$. Then* Ag *and* Sy *are satisfied if and only if $\succsim$ can be represented by $u(a) = \left(\alpha a_1^t + \sum_{i=2}^{L} \frac{1-\alpha}{L-1} a_i^t\right)^{1/t}$ and $\succsim_J$ can be represented by $v(a) = \left(\sum_{i=1}^{L} a_i^t\right)^{1/t}$ with some $\alpha \in (1/L, 1)$ and $t = r = s$.*

2.3 *Experimental Setup*

Here we only describe how the experimental setup should look like. In our companion paper (Becker et al. 2013) we analyse the results of such an experiment. As the experimental analysis is carried out based on choices from competitive budget sets which are standard in the theory of demand, we will focus on this case here. We would like to emphasise though that the entire analysis can also be carried out using more general budget sets. Such an analysis could be based on the results of Yatchew (1985), Matzkin (1991), or Forges and Minelli (2009). A very general treatment can be found in Heufer (2012b,a).

The participant is asked to make decisions on several choice sets, called *budgets*. Budgets are of the form $B^i = B(\rho^i) = \{a \in \mathbb{A} : \rho^i a \leq 1\}$, where each *price vector* ρ is an element of the *price space* $\mathbb{R}_{++}^L$. We will also refer to the budgets by the characterising price vector.[6] The *demand correspondence* $D : 2^{\mathbb{A}} \to 2^{\mathbb{A}}$ of a participant assigns to each choice set the set of allocations or bundles $D(\rho^i)$ demanded by the participant when asked to make a decision on $B(\rho^i)$. In an experiment, we *observe* one of these allocations, $a^i \in D(\rho^i)$. This demand is assumed to be the only observable in the model. Note that the definition of σ allows to let $\sigma(\rho)$ denote the set of all $L!$ permutations of the elements of the vector ρ.

An experiment consists of N different choice sets, indexed by $i = 1, \ldots, N$. An observation is then a pair (a, ρ), where $a \in D(\rho)$. The set of observations on a participant can then be denoted $\{(a^i, \rho^i)\}_{i=1}^N$, which is short hand notation for $\bigcup_{i=1}^N \{(a^i, \rho^i)\}$. A *dictator choice experiment*, or simply dictator experiment, is an experiment in which the first element of each bundle, a_1, is the payoff to the participant, and all other elements, $a_2, \ldots, a_L$, are the payoffs to the other individuals. A *social planner choice experiment*, or simply social planner experiment, is an experiment in which the participant has no personal monetary stake, that is, an experiment in which the participants allocates the payoff of two or more other individuals. A *choice experiment under veil of ignorance*, or simply veil of ignorance experiment, is an experiment in which the participant has personal monetary stake but does not know with certainty which of the elements of an $a \in \mathbb{A}$ gives his own payoff. More precisely, the participant allocates payoff to "persons" labelled $j = 1, \ldots, L$, and every individual $i = 1, \ldots, L$, including the participant, is labelled person j with the same probability. The participant does not know the labels in advance. We will call these experiments D-experiments, P-experiments, and V-experiments, respectively.

To avoid confusion, we will use the following notation: We have N_x observations from a D-experiment, and this set of observations is denoted $\Omega_x = \{(x^i, p^i)\}_{i=1}^{N_x}$, and we have N_y observations from a P- or V-experiment, and this set is denoted $\Omega_y = \{(y^j, q^j)\}_{j=1}^{N_y}$, with each $x^i, y^j \in \mathbb{A}$ and

[6] Price vectors are normalised such that expenditure equals 1 for choices for which the budget constraint is binding. This definition is usual (see, e.g. Varian 1982) and relies on homogeneity of demand. Strictly speaking, the researcher does not only observe price vectors which characterise the budget and the corresponding choice or demand, but also the expenditure, which can then be used to normalise prices.

$p^i, q^j \in \mathbb{R}^L_{++}$. In a slight abuse of notation, let $\Omega_x = \varnothing$ and $\Omega_y = \varnothing$ if $N_x = 0$ or $N_y = 0$, respectively. Furthermore it is assumed that observed demand on budgets is exhaustive (or that it satisfies budget-balancedness, or that the budget constraints are binding).

The main idea behind these experiments is that we can use the observations from the D-experiment to elicit a participant's preferences and from the P-experiment or the V-experiment to elicit a participant's individual notion of distributive justice. In an experimental implementation each participant would thus have to participate in the D-experiment and either the P- or V-experiment. The way in which this inference about preferences is made is described in Section 2.4 below.

2.4 *Revealed Preference and Revealed Notions of Justice*

We can now define the observable revealed preference relation R and revealed notion of distributive justice relation $\mathrm{R_J}$.

The *transitive closure* $(\mathrm{Q})^+$ of a binary relation Q is defined as the smallest transitive relation that contains Q, that is, $x(\mathrm{Q})^+ y$ if there are $x', \ldots, x'''$ such that $x \,\mathrm{Q}\, x'$, $x' \,\mathrm{Q}\, x''$, …, $x''' \,\mathrm{Q}\, y$.[7]

Given a set of observations $\Omega_x = \{(x^i, p^i)\}_{i=1}^{N_x}$ in a D-experiment, we say that the allocation x^i is *directly revealed preferred* to an allocation a, written as $x^i \,\mathrm{R}^0\, a$, if $p^i x^i \geq p^i a$; it is (indirectly) *revealed preferred* if $x^i \,\mathrm{R}\, a$, where $\mathrm{R} = (\mathrm{R}^0)^+$. The allocation x^i is *strictly directly revealed preferred* to a, written as $x^i \,\mathrm{P}^0\, a$, if $p^i x^i > p^i a$; it is (indirectly) *strictly revealed preferred*, written as $x^i \,\mathrm{P}\, a$, if $x^i \,\mathrm{R}\, x^j$, $x^j \,\mathrm{P}^0\, x^k$, and $x^k \,\mathrm{R}\, a$ for some observations x^j, x^k. Given a set of observations $\Omega_y = \{(y^j, q^j)\}_{j=1}^{N_y}$ in a P-experiment or V-experiment, the relations R^0_J, $\mathrm{R_J}$, P^0_J, and $\mathrm{P_J}$ are defined in the same way.

While a dictator experiment fits well into the standard theory of consumer behaviour and data obtained from these experiments have been analysed with methods based on revealed preference (e.g., Andreoni and Miller 2002), the idea of "revealed justice" in experiments may require some justification. The "veil of ignorance" is one of the central ideas in Rawls' (1971) theory of justice, where it is argued that to consider the morality of an issue, individuals must not know their position in society.[8] Similar conceptions have been discussed by Harsanyi (1953, 1955). In the context of our approach to experimentally assess notions of justice by participants, the veil of ignorance approach has the advantage of being incentive compatible. However, if participants are mostly self-interested with little regard for others, a veil of ignorance experiment is very similar to an experiment in which participants are asked to invest in two or more risky assets, and Traub et al. (2009), for example, find that participants' risk preferences do not adequately reflect their inequality attitudes. Thus, an experiment that eliminates all influence of self-interest and risk preferences, such as a P-experiment, might be better suited to assess participants' notion of justice. However a P-experiment has the disadvantage of not being incentive compatible in a narrow sense: If participants have little or no interest in choosing a just allocation, then they might shun the cognitive effort of finding this allocation. However, the task does not require a lot of effort, and if participants do care about justice, they should have an incentive to put in the required effort.

[7] Note that $(\mathrm{Q})^+$ is a *closure operator* on a binary relation Q. A closure operator is a function C which is extensive ($\mathrm{Q} \subseteq C(\mathrm{Q})$), increasing ($\mathrm{Q} \subseteq \mathrm{Q}' \Rightarrow C(\mathrm{Q}) \subseteq C(\mathrm{Q}')$), and idempotent ($C(C(\mathrm{Q})) = C(\mathrm{Q})$).

[8] For recent experimental applications and discussions, see for example Becker and Miller (2009), Krawczyk (2010), and Schildberg-Hörisch (2010). An application to ethics in the health sector can be found in Andersson and Lyttkens (1999).

Whether or not results from a P-experiment in the framework of this article differ from a V-experiment is an empirical question that will be a focus of the experimental analysis in the companion paper (Becker et al. 2013).

2.5 *Extensions of Revealed Relations Based on the Central Axioms*

Let $N = N_x + N_y$ and

$$\{z^k\}_{k=1}^{N} = \{x^i\}_{i=1}^{N_x} \cup \{y^j\}_{j=1}^{N_y}.$$

Let R^0 and R_J^0 be the revealed relations based on Ω_x and Ω_y, with $\mathrm{R} = (\mathrm{R}^0)^+$ and $\mathrm{R}_\mathrm{J} = (\mathrm{R}_\mathrm{J}^0)^+$.

Consider the Axiom AG, which states that $[\succsim_\mathrm{J} \cup \succ_\mathrm{S}] \subseteq \succsim$. If we know a DM's complete preference and notion of justice, we can directly test if the axiom is satisfied. But if it is, we do not learn anything new about the DM's relations, as we already know the complete relations. Given a finite set of observations, however, we will only observe a subset of an individual's preference and notion of justice. Suppose that $\mathrm{R} \subseteq \succsim$ and $\mathrm{R}_\mathrm{J} \subseteq \succsim_\mathrm{J}$, and that the choices do not violate AG. Then we can extend R by adding the part of $\mathrm{R}_\mathrm{J} \cup \succ_\mathrm{S}$ that is not already contained in R, and similarly for R_J.

First, by imposing axiom SY on the data, we obtain that z^i is revealed more just than a if some permutation of z^i is revealed more just than a. Thus, we define the *SY-closure* $C_{\mathrm{SY}}(\mathrm{Q})$ of a binary relation Q as

$$C_{\mathrm{SY}}(\mathrm{Q}) = \bigcup_{k=1}^{L!}\bigcup_{\ell=1}^{L!}\{(\sigma_k(x), \sigma_\ell(y)) \in \mathbb{A} \times \mathbb{A} : x \,\mathrm{Q}\, y\}. \tag{1}$$

By imposing axiom AG2 on the data we obtain that z^i is revealed more just than a if either z^i is revealed more just than a in the P- or V-experiment, or z^i is revealed preferred to a in the D-experiment even though a gives the DM a higher payoff than z^i. Thus, we define the AG2-extension of R_J^0 as

$$\tilde{\mathrm{R}}_\mathrm{J}^0 = \mathrm{R}_\mathrm{J}^0 \cup [\mathrm{R}^0 \cap \prec_\mathrm{S}], \tag{2}$$

and of the strict relation as

$$\tilde{\mathrm{P}}_\mathrm{J}^0 = \mathrm{P}_\mathrm{J}^0 \cup [\mathrm{P}^0 \cap \prec_\mathrm{S}]. \tag{3}$$

See Figure 3.(a) for an illustration. Let $\tilde{\mathrm{R}}_\mathrm{J} = (\tilde{\mathrm{R}}_\mathrm{J}^0)^+$. We also define the relation $\tilde{\mathrm{P}}_\mathrm{J}$ in the usual way, that is, $z^i \,\tilde{\mathrm{P}}_\mathrm{J}\, z^j$ if $z^i \,\tilde{\mathrm{R}}_\mathrm{J}\, z^k \,\tilde{\mathrm{P}}_\mathrm{J}^0\, z^\ell \,\tilde{\mathrm{R}}_\mathrm{J}\, z^j$ for some observations z^k and z^ℓ.

By imposing axiom AG1 on the data we obtain that z^i is revealed preferred to a if either z^i is revealed preferred to a in the D-experiment, or z^i is both revealed more just than a and gives the DM a higher payoff. Thus, we define the AG1-extension as

$$\tilde{\mathrm{R}}^0 = \mathrm{R}^0 \cup [\mathrm{R}_\mathrm{J}^0 \cap \succ_\mathrm{S}], \tag{4}$$

$$\tilde{\mathrm{P}}^0 = \mathrm{P}^0 \cup [\mathrm{P}_\mathrm{J}^0 \cap \succ_\mathrm{S}]. \tag{5}$$

See Figure 3.(b) for an illustration. Again, $\tilde{\mathrm{R}}$ and $\tilde{\mathrm{P}}$ are defined in the usual way.

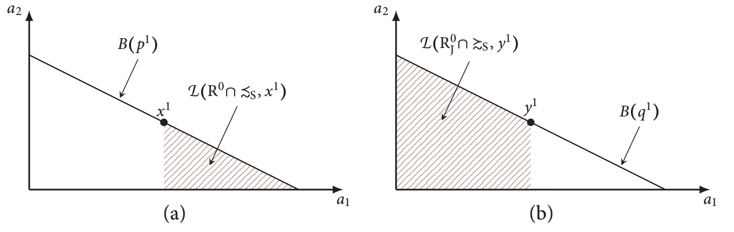


Figure 3: An illustration of $\tilde{\mathrm{R}}_\mathrm{J}$ and $\tilde{\mathrm{R}}$. In (a), we have $x^1\ \mathrm{R}^0\ a$ and $x^1 \precsim_\mathrm{S} a$ for all a in the indicated region and thus $x^1\ \tilde{\mathrm{R}}_\mathrm{J}\ a$. In (b), we have $y^1\ \mathrm{R}_\mathrm{J}^0\ a$ and $y^1 \succsim_\mathrm{S} a$ for all a in the indicated region and thus $y^1\ \tilde{\mathrm{R}}\ a$.

We define the Sy-Ag2-extensions as

$$\hat{\mathrm{R}}_\mathrm{J}^0 = C_{\mathrm{Sy}}(\tilde{\mathrm{R}}_\mathrm{J}^0), \tag{6}$$
$$\hat{\mathrm{P}}_\mathrm{J}^0 = C_{\mathrm{Sy}}(\tilde{\mathrm{P}}_\mathrm{J}^0) \tag{7}$$

and the Sy-Ag1-extension as

$$\hat{\mathrm{R}}^0 = \tilde{\mathrm{R}}^0 \cup [\hat{\mathrm{R}}_\mathrm{J}^0 \cap \succ_\mathrm{S}], \tag{8}$$
$$\hat{\mathrm{P}}^0 = \tilde{\mathrm{P}}^0 \cup [\hat{\mathrm{P}}_\mathrm{J}^0 \cap \succ_\mathrm{S}], \tag{9}$$

and again similarly for $\hat{\mathrm{R}}_\mathrm{J}$, $\hat{\mathrm{P}}_\mathrm{J}$, $\hat{\mathrm{R}}$, and $\hat{\mathrm{P}}$.

Note that $\mathrm{R} \subseteq \tilde{\mathrm{R}} \subseteq \hat{\mathrm{R}}$ and $\mathrm{R}_\mathrm{J} \subseteq \tilde{\mathrm{R}}_\mathrm{J} \subseteq \hat{\mathrm{R}}_\mathrm{J}$. Figure 4 illustrates the construction of the extended relations.

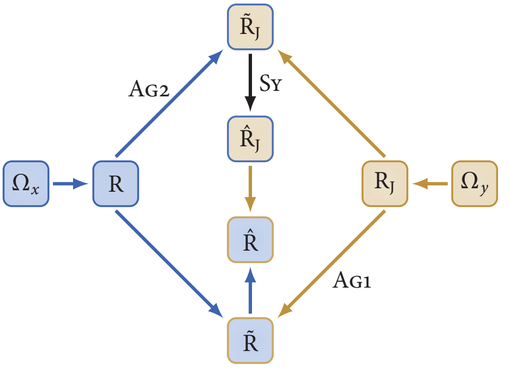


Figure 4: Construction of the extended revealed preference and revealed notion of justice relations.

3 THEORETICAL FOUNDATIONS: DATA ANALYSIS

3.1 *Rationalisability*

3.1.1 *Rationalisability without Extensions*

A utility function u *rationalises* a set of observations Ω_x if $u(a) \geq u(a')$ whenever $a \, \mathrm{R} \, a'$. A utility function is *non-satiated* if for every $a' \in \mathbb{A}$ and every $\varepsilon > 0$, there exists an $a \in N_\varepsilon(a')$ such that $u(a) > u(a')$.

As Afriat's Theorem below shows, there is an easily testable condition which is both necessary and sufficient for the existence of a utility function which rationalises a set of observations. This condition is know as Garp (Varian 1982); we also use a weak form of it, the W-Garp (Banerjee and Murphy 2006).

Definition 1 *A set of observations Ω_x satisfies the* Generalised Axiom of Revealed Preference *(Garp) if whenever $x^i \, \mathrm{R} \, x^j$ then $[$not $x^j \, P^0 \, x^i]$. It satisfies the* Weak Garp *(W-Garp) if whenever $x^i \, R^0 \, x^j$ then $[$not $x^j \, P^0 \, x^i]$.*

Theorem 1 (Afriat 1967, Diewert 1973, Varian 1982, Fostel et al. 2004) *Given a set of observations Ω_x, the following conditions are equivalent:*

1. *The set of observations satisfies Garp.*
2. *There exists a non-satiated, continuous, concave, and monotonic utility function which rationalises the set of observations.*

Note that Afriat's Theorem implies that if there exists a non-satiated utility function which rationalises the data, there also exists a monotonic utility function which rationalises the data. Thus, demand generated by a monotonic utility function is indistinguishable from data generated by a non-monotonic utility function when the experiment only involves budgets as choice sets. In two dimensions, testing for the existence of a rationalising utility function is simpler, as the next proposition shows.

Proposition 1 (Banerjee and Murphy 2006) *When $L = 2$, Garp is equivalent to W-Garp.*

A justice function v *rationalises* a set of observations Ω_y if $v(a) \geq v(a')$ whenever $a \, \mathrm{R_J} \, a'$. A justice function is *symmetric* if for every $a \in \mathbb{A}$, $v(a) \geq v(\sigma_\ell(a))$ for all $\ell = 1, \ldots, L!$.

It is obvious that Afriat's Theorem will also hold for rationalisation of a revealed justice relation by a justice function, given data from a p- or v-experiment. It is also easy to see that Garp is not sufficient for rationalisation by a symmetric justice function. However, a simple extension of Garp can be shown to be both necessary and sufficient.

Definition 2 *A set of observations Ω_y satisfies the Sy-Garp if whenever $y^i [C_{\mathrm{Sy}}(\mathrm{R}_J)]^+ y^j$ then $y^j \, C_{\mathrm{Sy}}(P^0_J) \, y^i$.*

Then it is easy to proof the following corollary:

Corollary 1 *The following conditions are equivalent:*

1. *The set Ω_y satisfies* Sy-Garp.
2. *There exists a symmetric, non-satiated, continuous, concave, and monotonic justice function which rationalises Ω_y.*

We omit the rather straightforward proof of Corollary 1, as it is a simple variation of, for example, Varian's (1982) proof of Afriat's Theorem.

Unfortunately, we cannot simply apply an analogous version of Garp to the extended revealed preference relation and obtain a result like Afriat's Theorem. The problem is that an allocation a can be revealed preferred to (or revealed worse than) an allocation a' according to one of the extended relations, but a might not be observed as a choice on budget sets in a D-experiment. For example, if we observe a choice y^1 on a budget $B(q^1)$ in a P- or V-experiment, then for all $a \in B(q^1) \cap \mathcal{L}(\succ_S, y^1)$, $y^1 \tilde{R}\, a$. But then we might also have m observations $\{x^i\}_{i=1}^m$ in a D-experiment, such that no x^i is in $B(q^1) \cap \mathcal{L}(\succ_S, y^1)$, yet a convex combination of the x^i is in $B(q^1) \cap \mathcal{L}(\succ_S, y^1)$. Clearly, if preferences are convex, then y^1 must be preferred to at least one of the x^i. Yet this fact is not captured by the relation $\tilde{R}$, and therefore the violation is not detected when applying Garp to $\tilde{R}$. A similar problem exists for $\tilde{R}_J$, as for all $a \in B(p^1) \cap \mathcal{U}(\succ_S, x^1)$, $x^1 \tilde{R}_J\, a$.

One way of dealing with this problem relies on generalisations of Afriat's Theorem, motivated by choice problems with rationing (e.g., Varian 1983, Yatchew 1985, Fleissig and Whitney 2011). In the appendix (Section A), we introduce conditions which are necessary and sufficient for the context of this paper. However, we are mostly interested in analysing data from a two-dimensional experiment, and as it turns out, these conditions in the appendix are not required for the data analysis we carry out in Becker et al. (2013). We will therefore focus on the two-dimensional case in the main part of this paper and relegate the general case to the appendix.

3.1.2 Extended Rationalisability

We will need notation which is somewhat involved. Let $\Omega = \Omega_x \cup \Omega_y$ and $N = N_x + N_y$. Furthermore, let $M = N\,L!$ and

$$\begin{aligned} \{z^k\}_{k=1}^N &= \{x^i\}_{i=1}^{N_x} \cup \{y^j\}_{j=1}^{N_y}, \\ \{\xi^m\}_{m=1}^M &= \bigcup_{\ell=1}^{L!} \bigcup_{i=1}^{N_x} \{\sigma_\ell(x^i)\} \cup \bigcup_{\ell=1}^{L!} \bigcup_{j=1}^{N_y} \{\sigma_\ell(y^j)\}, \end{aligned} \tag{10}$$

The first definition is a reminder of the definition in Eq. (1) and merely puts together the choices from the two experiments. The second definition adds all permutations of all the choices.

Given a set of observations Ω_x in a D-experiment, and a set of observations Ω_y in a P- or V-experiment, and $\Omega = \Omega_x \cup \Omega_y$,

- a utility function u Ag-*rationalises* Ω if $u(a) \geq u(a')$ whenever $a \tilde{R}\, a'$;
- a justice function v Ag-*rationalises* Ω if $v(a) \geq v(a')$ whenever $a \tilde{R}_J\, a'$;
- a utility function u Ag-Sy-*rationalises* Ω if $u(a) \geq u(a')$ whenever $a \hat{R}\, a'$;
- a justice function v Ag-Sy-*rationalises* Ω if $v(a) \geq v(a')$ whenever $a \hat{R}_J\, a'$.

The following definitions introduce conditions which we show are necessary and sufficient for extended rationalisability.

Definition 3 *A set of observations Ω satisfies*

- *Ag1-Garp if Ω_x and Ω_y satisfy Garp and if whenever $z^i \tilde{\mathrm{R}} z^j$ then $[$not $z^j \tilde{P}^0 z^i]$;*
- *Ag1-W-Garp if Ω_x and Ω_y satisfy W-Garp and if whenever $z^i \tilde{\mathrm{R}}^0 z^j$ then $[$not $z^j \tilde{P}^0 z^i]$;*
- *Ag2-Garp if Ω_x and Ω_y satisfy Garp and if whenever $z^i \tilde{\mathrm{R}}_J z^j$ then $[$not $z^j \tilde{P}^0_J z^i]$;*
- *Ag2-W-Garp if Ω_x and Ω_y satisfy W-Garp and if whenever $z^i \tilde{\mathrm{R}}^0_J z^j$ then $[$not $z^j \tilde{P}^0_J z^i]$;*
- *Ag1-Sy-Garp if Ω_x satisfies Garp and $\hat{\Omega}_y$ satisfies Sy-Garp and if whenever $\xi^i \hat{\mathrm{R}} \xi^j$ then $[$not $\xi^j \hat{P}^0 \xi^i]$;*
- *Ag1-Sy-W-Garp if Ω_x satisfies W-Garp and $\hat{\Omega}_y$ satisfies Sy-Garp and if whenever $\xi^i \hat{\mathrm{R}}^0 \xi^j$ then $[$not $\xi^j \hat{P}^0 \xi^i]$;*
- *Ag2-Sy-Garp if Ω_x satisfies Garp and $\hat{\Omega}_y$ satisfies Sy-Garp and if whenever $\xi^i \hat{\mathrm{R}}_J \xi^j$ then $[$not $\xi^j \hat{P}^0_J \xi^i]$;*
- *Ag2-Sy-W-Garp if Ω_x satisfies W-Garp and $\hat{\Omega}_y$ satisfies Sy-Garp and if whenever $\xi^i \hat{\mathrm{R}}^0_J \xi^j$ then $[$not $\xi^j \hat{P}^0_J \xi^i]$.*

Similarly to the relationship between Garp and W-Garp, we can show that the Ag-extensions are equivalent. What is also interesting is that the Ag1 and Ag2-versions are also equivalent to each other, mirroring the result stated in Fact 1 for complete relations.

Proposition 2 *When $L = 2$,*

- *Ag1-Garp, Ag2-Garp, Ag1-W-Garp, and Ag2-W-Garp are all equivalent;*
- *Ag1-Sy-Garp, Ag2-Sy-Garp, Ag1-Sy-W-Garp, and Ag2-Sy-W-Garp are all equivalent.*

Given Proposition 2, we will also refer to both versions as simply Ag-(W-)Garp and Ag-Sy-(W-)Garp, respectively.

Theorem 2 *Given observations Ω_x from a d-experiment and Ω_y from a p- or v-experiment with two dimensions $(L = 2)$, the following conditions are equivalent:*

1. *Ω satisfies Ag-W-Garp [Ω satisfies Ag-Sy-W-Garp].*
2. *There exists a non-satiated, continuous, concave, and monotonic utility function u which Ag-rationalises [Ag-Sy-rationalises] Ω, and a non-satiated, continuous, concave, and monotonic [and symmetric] justice function v which Ag-rationalises Ω, such that for all $a, a' \in \mathbb{A}$ with $a >_S a'$, $u(a) \leq u(a')$ implies $v(a) \leq v(a')$ and $v(a) \geq v(a')$ implies $u(a) \geq u(a')$.*

Note that the last condition of Theorem 2 is stronger than mere Ag-rationalisation. It states that not only does there exist u and v which Ag-rationalise Ω, but these two functions also represent complete preferences which satisfy Ag as stated in Eq. (Ag1) and (Ag2).

3.2 *Recoverability*

Following Varian (1982), we now turn to the question of recoverability of preferences. Given some allocation x^0 which was not necessarily observed as a choice, the set of prices which *support* x^0 is defined as

$$\Phi_{\mathrm{R}}(x^0) = \left\{p^0 \in \mathbb{R}^L_{++} : \ \{(x^i, p^i)\}_{i=0}^{N_x} \text{ satisfies } \textsc{Garp} \text{ and } p^0 x^0 = 1\right\}. \tag{11}$$

This definition can then be used to describe the set of all allocations which are revealed worse and revealed preferred to an allocation x^0: If for all price vectors at which x^0 can be demanded without violating Garp x^0 must be revealed preferred to x, then x is in the set of all allocations revealed worse to x^0. If for all price vectors at which some x is demanded – given that it does not violate Garp – the price vector will make x revealed preferred to x^0, then x is in the set of all allocations revealed preferred to x^0. Thus, the set of all allocations which are *revealed worse* than x^0 is given by

$$\mathcal{RW}(x^0) = \left\{a \in \mathbb{A} : \ \text{for all } p^0 \in \Phi_{\mathrm{R}}(x^0), x^0 \, \mathrm{P} \, a\right\} \tag{12}$$

and the set of all allocations which are *revealed preferred* to x^0 is given by

$$\mathcal{RP}(x^0) = \left\{a \in \mathbb{A} : \ \text{for all } p \in \Phi_{\mathrm{R}}(a), a \, \mathrm{P} \, x^0\right\}. \tag{13}$$

Note that, by definition, $a \in \mathcal{RW}(x^0)$ is equivalent to $x^0 \in \mathcal{RP}(a)$. We can define corresponding *revealed less just* and *revealed more just* sets for revealed notions of justice, and we can extend these sets using the relations $\tilde{\mathrm{R}}$, $\hat{\mathrm{R}}$, $\tilde{\mathrm{R}}_{\mathrm{J}}$, and $\hat{\mathrm{R}}_{\mathrm{J}}$ based on the relations recovered from a d- and a p- or v-experiment.

For brevity, we will only consider the relations $\hat{\mathrm{R}}$ and $\hat{\mathrm{R}}_{\mathrm{J}}$ and the two-dimensional case with $L = 2$. For the higher dimensional case, the axioms provided in the appendix (Section A) can be used. We define

$$\hat{\Phi}_{\mathrm{R}}(x^0) = \left\{p^0 \in \mathbb{R}^L_{++} : \{(x^0, p^0)\} \cup \Omega \text{ satisfies } \textsc{Ag-Sy-Garp} \text{ and } p^0 x^0 = 1\right\}, \tag{14}$$

$$\hat{\Phi}_{\mathrm{J}}(y^0) = \left\{q^0 \in \mathbb{R}^L_{++} : \Omega_x \cup \{(y^0, q^0)\} \cup \Omega_y \text{ satisfies } \textsc{Ag-Sy-Garp} \text{ and } q^0 y^0 = 1\right\}. \tag{15}$$

This leads to the definitions of revealed preferred and revealed worse allocations based on the extended revealed preference relation and corresponding sets of revealed more just and revealed less just allocations:

$$\widehat{\mathcal{RW}}(x^0) = \left\{a \in \mathbb{A} : \ \text{for all } p^0 \in \hat{\Phi}_{\mathrm{R}}(x^0), x^0 \, \hat{\mathrm{P}} \, a\right\}, \tag{16}$$

$$\widehat{\mathcal{RP}}(x^0) = \left\{a \in \mathbb{A} : \ \text{for all } p \in \hat{\Phi}_{\mathrm{R}}(a), a \, \hat{\mathrm{P}} \, x^0\right\}, \tag{17}$$

$$\widehat{\mathcal{RLJ}}(y^0) = \left\{a \in \mathbb{A} : \ \text{for all } q^0 \in \hat{\Phi}_{\mathrm{J}}(y^0), y^0 \, \hat{\mathrm{P}}_{\mathrm{J}} \, a\right\}, \tag{18}$$

$$\widehat{\mathcal{RMJ}}(y^0) = \left\{a \in \mathbb{A} : \ \text{for all } q \in \hat{\Phi}_{\mathrm{J}}(a), a \, \hat{\mathrm{P}}_{\mathrm{J}} \, y^0\right\}. \tag{19}$$

A first simple example is shown in Figure 5. It shows the revealed more and revealed less just sets of an allocation y^0 without any observations, based only on the symmetry assumption. Two more examples are shown in Figure 6, which are based on a single observation in a p- or v-experiment.

In Figure 6.(a), the revealed more and less just allocations of the observation (y^1, q^1) are shown; in Figure 6.(b) these sets are shown for a different observation (y^1, q^1) and some allocation y^0 which has not been observed as a choice.

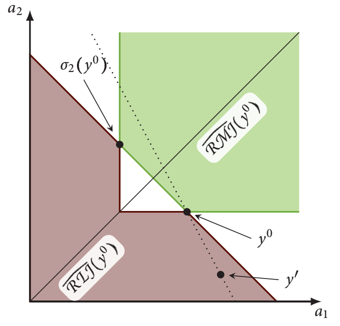


Figure 5: Revealed more just and revealed less just allocations without any observations. For example, if $y^0\ \mathrm{P}_{\mathrm{J}}^0\ y'$ then $y^0\ \mathrm{P}_{\mathrm{J}}^0\ \sigma_2(y^0)$, but by symmetry $\sigma_2(y^0)\ \hat{\mathrm{R}}_{\mathrm{J}}\ y^0$, thus $y^0\ \mathrm{P}_{\mathrm{J}}^0\ y'$ would violate Ag2SyGarp. Therefore we must have $y' \in \widehat{\mathcal{RLJ}}(y^0)$.

What is more interesting than the standard application of Varian's framework to observations collected in a P- or V-experiment is the fact that with the help of Ag2 and Sy, we can deduce notions of justice from a standard D-experiment. Figure 7.(a) shows that if a participant chooses an allocation to the lower left of the 45° line we can deduce more about his notion of justice than without such an observation. Figure 7.(b) shows that some observations do not provide more information than the one already depicted in Figure 5.

Another interesting application is that based on Ag and Sy, we can also deduce more about the revealed preference relation in a D-experiment, even if no observations from a P- or V-experiment are available. Figure 8 illustrates this.

Figure 9 illustrates the revealed preferred and worse sets for an allocation not observed as a choice, and 10 shows an example with several observations.

Conveniently, we can express the revealed preferred and more just sets as the convex monotonic hull of a finite set of points, which makes it very easy to check if a point is revealed preferred to another and to draw the sets. The next proposition shows this.

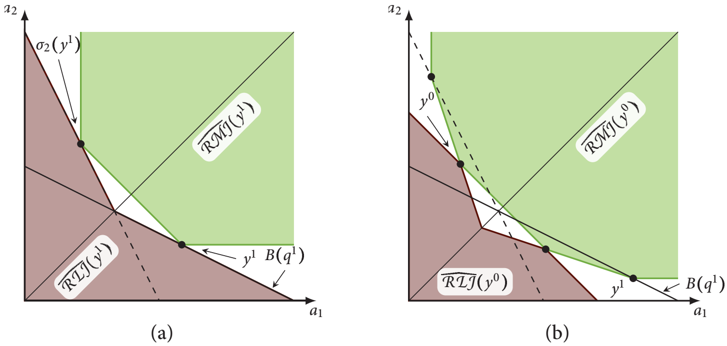


Figure 6: Revealed more just and revealed less just allocations with one observation y^1 in a P- or V-experiment. In (a), the revealed more and less just sets of the observation y^1 are shown. In (b), we use a different observation y^1; it shows the revealed more and less just sets for an allocation y^0 which was not observed as a choice.

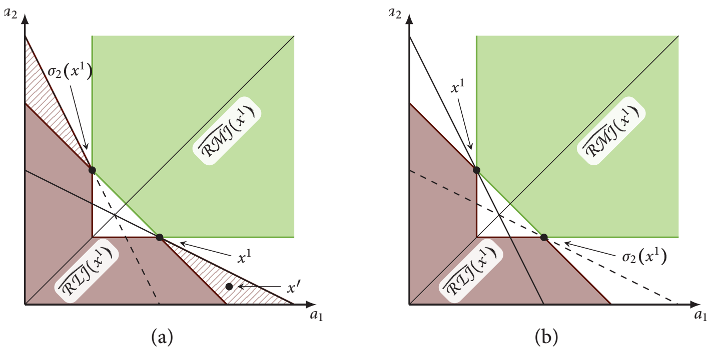


Figure 7: Revealed more and less just allocations with one observation, x^1, in a D-experiment. The dashed region shows the parts which are added by using the AG- and SY-axioms. In (a), we have that $x' \in \widehat{\mathcal{RLJ}}(x^1)$ because $x^1\,[\mathrm{P}^0 \cap \precsim_\mathrm{S}]\,x'$ (compare with Figure 3.(a)). Of course, we might observe $x'\,\mathrm{R_J}\,x^1$ in a P- or V-experiment, but then the AGSYGARP would be violated. The observation in (b) does not provide additional information about the notion of fairness (compare with Figure 5).

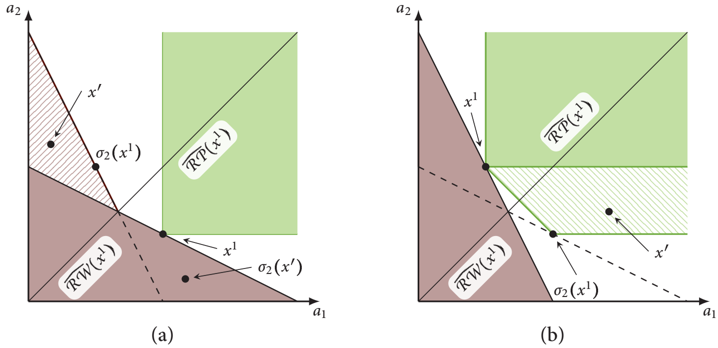


Figure 8: Revealed preferred and revealed worse allocations with one observation, x^1, in a D-experiment. The dashed region shows the parts which are added by using the AG- and SY-axioms. In (a), $x' \in \widehat{\mathcal{RW}}(x^1)$ because $x^1 \, P^0 \, \sigma_2(x')$ and $\sigma_2(x') \, [R_J \cap \succsim_S] x'$. In (b), $x' \in \widehat{\mathcal{RP}}(x^1)$ because $x' \geq \sigma_2(x^1)$ and $\sigma_2(x^1)[R_J \cap \succsim_S] x^1$.

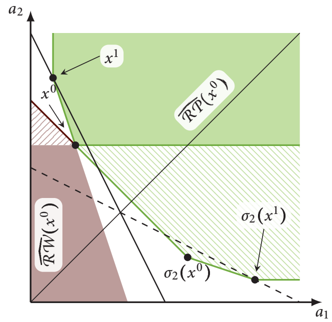


Figure 9: Revealed preferred and revealed worse allocations to an allocation not observed as a choice with one observation, x^1, in a D-experiment.

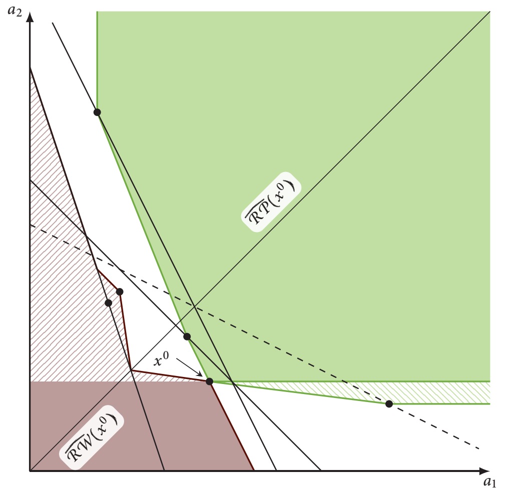


Figure 10: Revealed preferred and revealed worse allocations to an allocation not observed as a choice with several observations in a D-experiment.

Proposition 3 *Suppose Ω satisfies AG1-SY-GARP and AG2-SY-GARP and $L = 2$.*

$$\mathrm{int}CMH(\{x^i : x^i \,\mathrm{R}\, x^0\}) \subseteq \mathcal{RP}(x^0) \subseteq CMH(\{x^i : x^i \,\mathrm{R}\, x^0\}),$$
$$\mathrm{int}CMH(\{\xi^i : \xi^i \,\hat{\mathrm{R}}\, x^0\}) \subseteq \widehat{\mathcal{RP}}(x^0) \subseteq CMH(\{\xi^i : \xi^i \,\hat{\mathrm{R}}\, x^0\}),$$
$$\mathrm{int}CMH(\{\xi^i : \xi^i \,\hat{\mathrm{R}}_J\, x^0\}) \subseteq \widehat{\mathcal{RMJ}}(x^0) \subseteq CMH(\{\xi^i : \xi^i \,\hat{\mathrm{R}}_J\, x^0\}).$$

3.3 Interpersonal Comparisons

3.3.1 Global Comparison: A Stronger Sense of Justice

In this section, we show that the recovery of preferences and notions of justice allows interpersonal comparison of intensity of fairness, similar to the work of Karni and Safra (2002b).[9] Let $\trianglerighteq$, a binary relation on the set of all binary relations, be the *stronger sense of justice than* relation. For two preference-notion of justice pairs $(\succsim^1, \succsim_J^1)$ and $(\succsim^2, \succsim_J^2)$ which satisfy AG, we define

$$(\succsim^1, \succsim_J^1) \trianglerighteq (\succsim^2, \succsim_J^2) \;\text{ if }\; (\succsim^1 \cap \precsim_J^1) \subseteq (\succsim^2 \cap \precsim_J^2). \tag{20}$$

[9] See also Karni and Safra (2002a); related work also include Nguema's (2003) analysis of a sense of impartiality, Heufer's (2013b) revealed preference analysis of Karni and Safra's (2002b) work, and Heufer's (2011) revealed preference approach to interpersonal comparisons of risk aversion.

That is, a DM with the preference-justice pair $(\succsim^1, \succsim^1_J)$ has a stronger sense of justice than a DM with the preference-justice pair $(\succsim^2, \succsim^2_J)$ if, for every allocation $a \in \mathbb{A}$, $\mathcal{U}(\succsim^1 \cap \precsim^1_J, a) \subseteq \mathcal{U}(\succsim^2 \cap \precsim^2_J, a)$. Consider Figure 11, which shows indifference and iso-justice curves of two individuals with the same notion of justice. There are some allocations which both DMs prefer to a, even though these allocations are less just than a. Every allocation preferred by $\succsim^1$ and considered less just by $\succsim^1_J$ is also preferred by $\succsim^2$ and considered less just by $\succsim^2_J$. But there do exist allocations preferred by $\succsim^2$ and considered less just by $\succsim^2_J$, but not preferred by $\succsim^1$. Thus, every instance that can be used to argue that the DM with $\succsim^1$ is "unjust" because he contradicts his own justice ideal can equally be used to construct the same "accusation" against the DM with $\succsim^2$. But there are instances which can be used to accuse $\succsim^2$ but not $\succsim^1$, which is at the core of the definition of a stronger sense of fairness. Put differently, the selfishness of the first DM is a subset of the selfishness of the second DM.

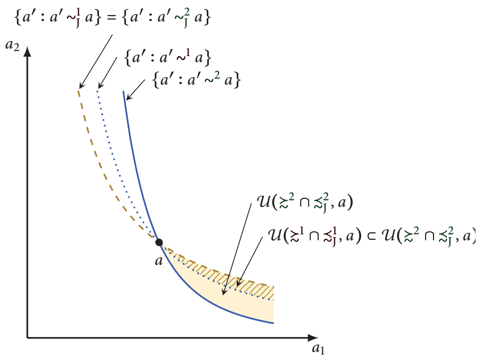


Figure 11: A stronger sense of justice: Both individuals have the same notion of justice. There are allocations which are more just than a and not preferred to a according to $\succsim^1$, but preferred to a according to $\succsim^2$. Thus, $(\succsim^1, \succsim^1_J)$ has a stronger sense of justice than $(\succsim^2, \succsim^2_J)$.

How can $\trianglerighteq$ be made operational given a finite set of observations on two DMs? The major obstacle is that the revealed preference relation is not complete, so $a \,\hat{R}\, a'$ does *not* imply $a' \,\hat{R}\, a$; even for the revealed preferred sets, $a \notin \widehat{\mathcal{RP}}(a')$ does *not* imply $a \in \widehat{\mathcal{RW}}(a')$. We therefore introduce the following concept, which is strongly based on Heufer's (2011) approach to comparative revealed risk aversion: Suppose that for some $a \in \mathbb{A}$, there is an allocation a' which the first DM—call him DM^1—reveals to prefer over a, and that he also reveals that to him, a' is more just than a. A second DM, DM^2, also reveals that he prefers a' over a, but he also reveals that to him, a is more just than a'. We then say that DM^1 has a *partially stronger revealed sense of justice* than DM^2. Then if there is no pair of allocations according to which DM^2 has a partially stronger revealed sense of justice, we conclude that DM^1 has a *stronger revealed sense of justice* than DM^2.

More formally, we define the *partially stronger revealed sense of justice than* relation $\hat{\trianglerighteq}_{\text{Rev}}$ as

$$(\hat{R}^1, \hat{R}^1_J) \,\hat{\trianglerighteq}_{\text{Rev}}\, (\hat{R}^2, \hat{R}^2_J) \text{ if } \exists a \in \mathbb{A}, [\text{cl}\widehat{\mathcal{RW}}^1(a) \cap \widehat{\mathcal{RLJ}}^1(a)] \cap [\text{cl}\widehat{\mathcal{RP}}^2(a) \cap \widehat{\mathcal{RLJ}}^2(a)] \neq \varnothing. \quad (21)$$

If $(\hat{R}^1, \hat{R}_J^1) \hat{\trianglerighteq}_{\text{Rev}} (\hat{R}^2, \hat{R}_J^2)$ and not $(\hat{R}^2, \hat{R}_J^2) \hat{\trianglerighteq}_{\text{Rev}} (\hat{R}^1, \hat{R}_J^1)$, then we say that $(\hat{R}^1, \hat{R}_J^1)$ has a stronger revealed sense of justice than $(\hat{R}^2, \hat{R}_J^2)$, written $(\hat{R}^1, \hat{R}_J^1) \hat{\triangleright}_{\text{Rev}} (\hat{R}^2, \hat{R}_J^2)$; that is, $\hat{\triangleright}_{\text{Rev}}$ is the asymmetric part of $\hat{\trianglerighteq}_{\text{Rev}}$.

Consider two sets of observations on two different DMs, Ω^1 and Ω^2. Let $\{\psi^i\}_{i=1}^{N^1+N^2}$ be the union of the sets of $\{\xi^i\}$ as defined in Eq. (10) for the two DMs. Then define

$$\hat{\delta}(\Omega^1, \Omega^2) = \begin{cases} 1 & \text{if there exist } \psi, \psi' \in \{\psi^i\}_{i=1}^{N^1+N^2} \text{ such that } \psi \in \widehat{\mathcal{RLJ}}^1(\psi') \cap \widehat{\mathcal{RLJ}}^2(\psi') \\ & \text{and } [\psi \in \text{cl}\widehat{\mathcal{RP}}^1(\psi') \cap \text{cl}\widehat{\mathcal{RW}}^2(\psi')]; \\ 0 & \text{otherwise.} \end{cases} \tag{22}$$

We then arrive at the following powerful theorem, which shows that the comparative approach is completely operational as it only requires a finite number of comparisons.

Theorem 3 *Suppose Ω^1 and Ω^2 satisfy AG1-SY-GARP and AG2-SY-GARP and $L = 2$.*

1. *The following conditions are equivalent:*
   - $\hat{\delta}(\Omega^1, \Omega^2) = 1$ *and* $\hat{\delta}(\Omega^2, \Omega^1) = 0$;
   - $(\hat{R}^1, \hat{R}_J^1) \hat{\triangleright}_{\text{Rev}} (\hat{R}^2, \hat{R}_J^2)$.
2. *The following conditions are equivalent:*
   - $\hat{\delta}(\Omega^1, \Omega^2) = \hat{\delta}(\Omega^2, \Omega^1) = 1$;
   - $(\hat{R}^1, \hat{R}_J^1) \hat{\trianglerighteq}_{\text{Rev}} (\hat{R}^2, \hat{R}_J^2)$ *and* $(\hat{R}^2, \hat{R}_J^2) \hat{\trianglerighteq}_{\text{Rev}} (\hat{R}^1, \hat{R}_J^1)$.
3. *The following conditions are equivalent:*
   - $\hat{\delta}(\Omega^1, \Omega^2) = \hat{\delta}(\Omega^2, \Omega^1) = 0$;
   - *neither* $(\hat{R}^1, \hat{R}_J^1) \hat{\triangleright}_{\text{Rev}} (\hat{R}^2, \hat{R}_J^2)$ *nor* $(\hat{R}^2, \hat{R}_J^2) \hat{\trianglerighteq}_{\text{Rev}} (\hat{R}^1, \hat{R}_J^1)$.

We omit the proof of Theorem 3, as it is practically the same as the proof in Heufer (2011, Theorem 3). The difference is that the revealed justice relation is not the same for all individuals, whereas in Heufer (2011) the common notion of risk in terms of stochastic dominance is the same for all. This is captured by the condition that $\psi \in \widehat{\mathcal{RLJ}}^1(\psi') \cap \widehat{\mathcal{RLJ}}^2(\psi')$, that is, both individuals agree on the justice ranking between ψ and ψ'.

Theorem 3 is important and powerful because it shows that it is both necessary and sufficient to compare only allocations which have been observed as choices in one of the experiments, or which are a permutation of one of these choices. Thus, even though $\hat{\trianglerighteq}_{\text{Rev}}$ is defined in terms of revealed sets for *all* elements of $\mathbb{A}$, a finite number of comparisons is enough to check if the condition in Eq. (21) is satisfied. The theorem therefore provides an operational non-parametric way to compare the strength of the sense of justice of two DMs.

Figure 12 shows an example with two DMs. We use a Cobb-Douglas utility function $u(a_1, a_2) = a_1^\alpha a_2^{1-\alpha}$ with $\alpha = 11/20$ for DM^1 and $\alpha = 19/20$ for DM^2 to generate choices on the budgets in 12.(a) for a D-experiment. We furthermore generated choices in a P- or V-experiment using a similar set of budgets and the same justice function for both DMs, in particular, a Cobb-Douglas form with $\alpha = 1/2$. Figure 12.(b) shows the intersection of the two revealed less just sets for the allocation x^0. Figures 12.(c) and (d) show the revealed worse set of DM^1 and the revealed preferred set of DM^2.

Figure 12.(e) shows $[\widehat{\mathcal{RW}}^1(x^0) \cap \widehat{\mathcal{RLJ}}^1(x^0)] \cap [\widehat{\mathcal{RP}}^2(x^0) \cap \widehat{\mathcal{RLJ}}^2(x^0)]$, which clearly indicates that $(\hat{\mathrm{R}}^1, \hat{\mathrm{R}}^1_{\mathrm{J}}) \hat{\triangleright}_{\mathrm{Rev}} (\hat{\mathrm{R}}^2, \hat{\mathrm{R}}^2_{\mathrm{J}})$; obviously, it can also be shown that DM[1]'s revealed sense of justice is not only partially stronger than DM[2]'s.

Note that in the definition of $\hat{\trianglerighteq}_{\mathrm{Rev}}$ and for Theorem 3 we have assumed consistency with Ag and Sy. This is not a necessity for such a construction; one can also justify different definitions which do not rely on these axioms. While using the extended relations $\hat{\mathrm{R}}$ and $\hat{\mathrm{R}}_{\mathrm{J}}$ instead of the regular R and R_{J} obviously provide more information under the axioms, there is another advantage: Ag assures that if $\psi \succ_{\mathrm{S}} \psi'$ and both DMs agree that ψ' is less just than ψ then both DMs will prefer ψ over ψ'. Thus, when we find that DM[1] has a stronger sense of justice than DM[2], this conclusion cannot be based on an observation where DM[2] violates his justice ideal in favour of the other person (instead of himself), and we can extend the interpretation to the statement that DM[1] is less selfish than DM[2].

However, there might be cases where not assuming the axioms is desirable. A practical consideration is that the computation of many different $\widehat{\mathcal{RP}}$ and $\widehat{\mathcal{RMJ}}$ sets can be very resource intensive and might be infeasible. One might also argue that the basic idea of the comparison is interesting enough even when the validity of Ag or Sy are doubted, or when there are minor violations of these axioms. We therefore suggest an alternative and define the relation $\trianglerighteq_{\mathrm{Rev}}$ as

$$
\begin{aligned}
(\mathrm{R}^1, \mathrm{R}^1_{\mathrm{J}}) \trianglerighteq_{\mathrm{Rev}} (\mathrm{R}^2, \mathrm{R}^2_{\mathrm{J}}) &\text{ if } \exists a \in \mathbb{A}, \\
&[\mathrm{cl}\mathcal{RW}^1(a) \cap \mathcal{RLJ}^1(a)] \cap [\mathrm{cl}\mathcal{RP}^2(a) \cap \mathcal{RLJ}^2(a)] \cap \mathcal{U}(\succ_{\mathrm{S}}, a) \neq \varnothing,
\end{aligned}
\tag{23}
$$

where $\mathcal{U}(\succ_{\mathrm{S}}, a)$ is added to exclude cases where a DM prefers a less just alternative that has a lower payoff. Eq. (22) can be redefined accordingly as

$$
\delta(\Omega^1, \Omega^2) = \begin{cases} 1 & \text{if there exist } \psi, \psi' \in \{\psi^i\}_{i=1}^{N^1+N^2} \text{ such that } \psi \succ_{\mathrm{S}} \psi' \text{ and} \\ & \psi \in \mathcal{RLJ}^1(\psi') \cap \mathcal{RLJ}^2(\psi') \text{ and } [\psi \in \mathrm{cl}\mathcal{RP}^1(\psi') \cap \mathrm{cl}\mathcal{RW}^2(\psi')]; \\ 0 & \text{otherwise.} \end{cases}
\tag{24}
$$

The results of Theorem 3 based on Eq. (23) and (24) then still hold.

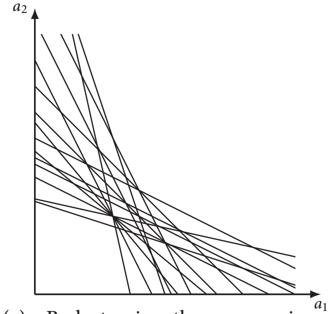
3.3.2 *Money Metric Based Comparisons*

Varian (1982) introduces approximations of Samuelson's (1974) money-metric utility function based on revealed preferred sets. The "exact" money-metric utility of an allocation $x^0 \in \mathbb{A}$ at prices p, given the utility function u, is defined as $m(x^0, p) = \inf p\, a$ such that $u(a) > u(x^0)$. Since we do not observe the true utility function u, we have to rely on approximations or upper and lower bounds. Varian (1982) defines the upper bound as

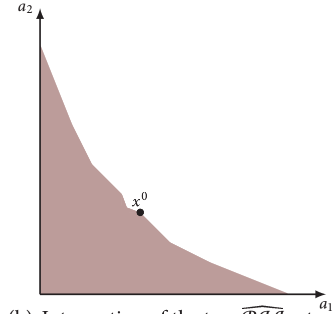
$$
m^+(x^0, p) = \inf_{a \in \mathcal{RP}(x^0)} pa
$$

With the help of Knoblauch's (1992) result and Proposition 3 above, we can write

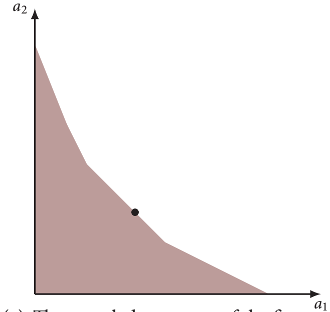
$$
m^+(x^0, p) = \min_{\{i:\, x^i \mathrm{R} x^0\}} px^i,
\tag{25}
$$



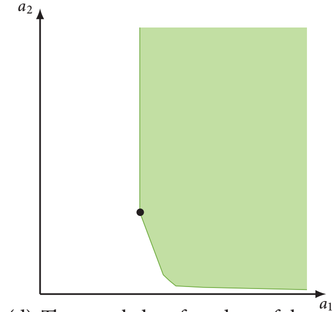
(a) Budgets in the D-experiment (clipped)



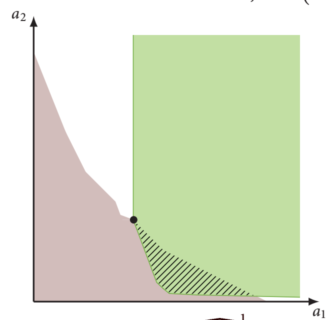
(b) Intersection of the two $\widehat{\mathcal{RLJ}}$ sets of the DMs.



(c) The revealed worse set of the first DM, $\widehat{\mathcal{RW}}^1(x^0)$.



(d) The revealed preferred set of the second DM, $\widehat{\mathcal{RP}}^2(x^0)$.



(e) Intersection of $\widehat{\mathcal{RW}}^1(x^0)$ and $\widehat{\mathcal{RP}}^2(x^0)$ with the intersection of the two $\widehat{\mathcal{RLJ}}$ sets.

Figure 12: Sense of justice with revealed sets. Choices were generated on the budgets in (a) using two Cobb-Douglas utility functions with $\alpha = 11/20$ for DM^1 and $\alpha = 19/20$ for DM^2. A similar set of budget was used for choices in a P- or V-experiment. Both DMs have the same symmetric Cobb-Douglas justice function. As the last figure clearly indicates, the first DM has a stronger revealed sense of justice than the second DM.

and the lower bound is given by

$$m^{-}(x^0, p) = \inf_{\{a \notin \mathcal{RW}(x^0)\}} pa. \tag{26}$$

Similarly, we define

$$\hat{m}^{+}(\xi^0, p) = \min_{\{i:\ \xi^i \hat{R} x^0\}} p\xi^i \tag{27}$$

$$\hat{m}^{-}(\xi^0, p) = \inf_{\{a \notin \widehat{\mathcal{RW}}(\xi^0)\}} pa. \tag{28}$$

Analogously, we can define a "money-metric justice value function", i.e.

$$\widehat{mmj}^{+}(\xi^0, q) = \min_{\{i:\ \xi^i \hat{R}_J x^0\}} q\xi^i, \tag{29}$$

$$\widehat{mmj}^{-}(\xi^0, q) = \inf_{\{a \notin \widehat{\mathcal{RLJ}}(y^0)\}} qa. \tag{30}$$

The money metric allows an interesting way to measure the strength of a DM's sense of justice. It allows to give approximate answers to the questions:

1. Given a price vector p and the most just allocation on $B(p)$, what is the money metric utility of the most just allocation?
2. Conversely, given a price vector p and the most preferred allocation on $B(p)$, what is the money metric justice value of the most preferred allocation?

Suppose we observe $x^i \in D(p^i)$ in a D-experiment and $y^i \in D(q^i)$ in a P- or V-experiment, with $p^i = q^i$. Then $\hat{m}^{+}(y^i, p^i)$ and $\widehat{mmj}^{+}(x^i, q^i)$ give approximate answers to the first and second question, respectively. Noting that $\hat{m}^{+}(x^i, p^i) = 1$ and $\widehat{mmj}^{+}(y^i, p^i) = 1$ if the data satisfies AG1-SY-GARP and AG2-SY-GARP, respectively, there is not necessarily a need to normalise utility among different DMs: $\hat{m}^{+}(y^i, p^i)$ can be interpreted as a percentage loss of utility if the DM were forced to choose y^i instead of x^i. If for two DMs with $\hat{m}_1$ and $\hat{m}_2$ we have $\hat{m}_1^{+}(y^i, p^i) > \hat{m}_2^{+}(\check{y}^i, p^i)$, then the first DM *reveals a stronger money metric sense of justice* than the second DM on the budget $B(p^i)$.

4 CONCLUSION AND DISCUSSION

This paper provides the first step towards a more extensive analysis of experimental data on social preferences and a new framework for original experimental design. The simple axioms on preferences and notions of justice and their empirical counterparts provided and analysed here allow to recover more about participants' preferences based on data collected in generalised dictator games. Combined with social planner or veil of ignorance experiments we can recover large parts of a participant's individual notion of justice. Furthermore it allows to make interpersonal comparisons between participants. The empirical approach is kept completely non-parametric and operational.

This paper, then, is the first step towards an extensive analysis of people's individual notion of justice and their personal strength of sense of justice. The experimental companion paper will test several hypothesis, using a two-dimensional allocation space. Questions of interest include

- Do participants have a well behaved notion of justice at all?
- Is their notion of justice compatible with their choices in the dictator experiment, and vice versa, in the sense of the Agreement-Axiom Ag?
- Are there substantial differences in the revealed justice relation observed in a social planner experiment and a veil of ignorance experiment? If not, this could imply that (i) the social planner experiment results in reasonable choices although it is not incentive-compatible in the usual sense, and (ii) the veil of ignorance is indeed a good concept to recover participants' notions of justice.
- Do participants "rationalise" (in the psychological sense) their choices in a dictator experiment if it is followed by a planner or veil of ignorance experiment? That is, do they choose in the social planner or veil of ignorance experiment to make their choices in the dictator experiment seem more just? Conversely, if the dictator experiment follows a social planner or veil of ignorance experiment, do participants make choices in the dictator experiment which are closer to their revealed notion of justice?
- How many different prototypically notions of justice are there, and how are they distributed? How well can participants be ranked by the strength of their sense of justice, based on the comparative approach?

We did not find satisfying answers to most of these question in the previous literature. We expect that the approach outlined here and the experimental design based on it will allow us to answer most or all of these questions. The results of such an experiment are reported in detail in the companion paper (Becker et al. 2013).

A APPENDIX: THE HIGHER-DIMENSIONAL CASE

In this section, we introduce a necessary and sufficient condition for extended rationalisability which will also allow to carry out the same non-parametric analysis as in Section 3.2 and 3.3. The approach is based on results and ideas by Varian (1983), Yatchew (1985), and Fleissig and Whitney 2011. An alternative approach could be based on the axioms proposed in Heufer (2012a).

A.1 *Extended Rationalisability in the Higher-Dimensional Case*

Let $\mathbf{1}_1 = (1, 0, \dots, 0)$ denote the L-dimensional vector with the first element equal to 1 and all other elements equal to 0. Let $\mathbf{1}_{-1} = (0, 1, \dots, 1)$ denote the L-dimensional vector with the first element equal to 0 and all other elements equal to 1. Let $\{r^k\}_{k=1}^N = \{p^i\}_{i=1}^{N_x} \cup \{q^j\}_{j=1}^{N_y}$ be the union of the two sets of price vectors from the D- and the P- or V-experiment.

For the following definition, we assume that $\sigma_\ell(a)$ for $a \in \mathbb{A}$ is defined for all $\ell = 1, \dots, L!$, even if some elements of a are equal and therefore some permutations are repeated. That is, $\sigma_\ell(a)$ is $(a_1, \dots, a_L)$ ordered by the ℓth permutation of the list $(1, 2, \dots, L)$. Furthermore, $\sigma_\ell(a)_i$ denotes the ith element of the permutation.

Definition 4 *A set of observations Ω satisfies*

C-AG1 if there exist numbers U^i, $\lambda^i > 0$, $\mu^i \geq 0$, for $i = 1, \dots, N$, with $\mu^i = 0$ if $i \leq N_x$ and $\mu^i \geq 0$ if $i > N_x$, such that

$$U^i \leq U^j + [\lambda^j r^j + \mu^j \mathbf{1}_1] \cdot [z^i - z^j].$$

C-AG2 if there exist numbers U^i, $\lambda^i > 0$, $\nu^i \geq 0$, for $i = 1, \dots, N$, with $\nu^i \geq 0$ if $i \leq N_x$ and $\nu^i = 0$ if $i > N_x$, such that

$$U^i \leq U^j + [\lambda^j r^j + \nu^j [\mathbf{1}_{-1} r^i]] \cdot [z^i - z^j].$$

C-AG1-SY if there exist numbers $U^{i,\ell}$, $\lambda^{i,\ell} > 0$, $\mu^{i,\ell} \geq 0$, $\nu^{i,\ell} \geq 0$, for $i = 1, \dots, N$ and $\ell = 1, \dots, L!$, with

$$\mu^{i,\ell} \begin{cases} \geq 0 & \text{if } i > N_x \text{ or } \ell > 1, \\ = 0 & \text{otherwise,} \end{cases}$$

$$\nu^{i,\ell} \begin{cases} \geq 0 & \text{if } i \leq N_x \text{ and } \ell > 1, \\ = 0 & \text{otherwise,} \end{cases}$$

such that

$$U^{i,\ell} \leq U^{j,m} + [\lambda^{j,m} \sigma_m(r^j) + \mu^{j,m} \mathbf{1}_1 + \nu^{j,m} [\sigma_m(\mathbf{1}_{-1}) \sigma_m(r^i)]] \cdot [\sigma_\ell(z^i) - \sigma_m(z^j)],$$
$$U^{i,\ell} \leq U^{i,m} \qquad \text{whenever } \sigma_\ell(z^i)_1 \leq \sigma_m(z^i)_1$$

C-Ag2-Sy *if there exist numbers* $U^{i,\ell}$, $\lambda^{i,\ell} > 0$, $\nu^{i,\ell} \geq 0$, *for* $i = 1, \ldots, N$ *and* $\ell = 1, \ldots, L!$, *with* $\nu^{i,\ell} \geq 0$ *if* $i \leq N_x$ *and* $\nu^{i,\ell} = 0$ *if* $i > N_y$, *such that*

$$U^{i,\ell} \leq U^{j,m} + \left[\lambda^{j,m}\sigma_m(r^j) + \nu^{j,m}[\sigma_m(\mathbf{1}_{-1})\sigma_m(r^i)]\right] \cdot [\sigma_\ell(z^i) - \sigma_m(z^j)],$$
$$U^{i,\ell} = U^{i,m} \qquad \text{for all } \ell, m = 1, \ldots, L!$$

Given these conditions, we can use the test for rationing described in Varian (1983; see also Yatchew 1985 and Fleissig and Whitney 2011) to prove the following general rationalisability theorem.

Theorem 4 *Given observations* Ω_x *from a* D*-experiment and* Ω_y *from a* P*- or* V*-experiment, the following conditions are equivalent*

1. Ω *satisfies Ag1-Garp and C-Ag1 [*Ω *satisfies Ag1-Sy-Garp and C-Ag1-Sy].*
2. Ω *satisfies Ag2-Garp and C-Ag2 [*Ω *satisfies Ag2-Sy-Garp and C-Ag2-Sy].*
3. *There exists a non-satiated, continuous, concave, and monotonic utility function which Ag-rationalises [Ag-Sy-rationalises]* Ω*, and a non-satiated, continuous, concave, and monotonic [and symmetric] justice function which Ag-rationalises [Ag-Sy-rationalises]* Ω*, such that for all* $a, a' \in \mathbb{A}$ *with* $a >_S a'$, $u(a) \leq u(a')$ *implies* $v(a) \leq v(a')$ *and* $v(a) \geq v(a')$ *implies* $u(a) \geq u(a')$.

Before we proof the theorem, we need to recall a rationalisation theorem for choices with more than one constraint.

Theorem 5 (Varian 1983, Yatchew 1985) *Suppose we observe choices* χ^i, $i = 1, \ldots, N_\chi$, *which might have been generated by a model of the form* $\max u(a)$ *such that* $\mathbf{G}^i a \leq \mathbf{c}^i$, *where* $\mathbf{G}^i$ *is an* $L \times M$ *matrix,* $\mathbf{c}^i$ *is an* M *vector,* $M \geq 1$, *each* L*-vector* $\mathbf{G}^i_m \in \mathbb{R}^L_+$, *and each element* $\mathbf{c}^i_m \geq 0$. *Then the following conditions are equivalent:*

1. *There exist scalars* U^i *and* M*-vectors* $\boldsymbol{\kappa}^i_m \geq 0$ *for* $i = 1, \ldots, N_\chi$ *and* $j = 1, \ldots, M$, *with* $\boldsymbol{\kappa}^i_j = 0$ *if* $\mathbf{G}^i_j a = \mathbf{c}^i_j$, *such that*

$$U^i \leq U^j + [\boldsymbol{\kappa}^j \mathbf{G}^j] \cdot [\chi^i - \chi^j].$$

2. *There exists a non-satiated, continuous, concave, and monotonic utility function which rationalises the set of observations.*

Proof of Theorem 4 The proof is based on the fact that for the extended revealed preference relations, choices from a P- or V-experiment are used *as if* they were choices made under rationing. Consider Figure 3.(b). Given the directly revealed worse set of y^1, this choice from the P- or V-experiment can be interpreted as a choice from a D-experiment with rationing, that is, a choice were the DM was not allowed to demand more than y^1_1 of the first commodity. The situation is similar for choices from a D-experiment and the revealed justice relation. Consider Figure 3.(a). This choice x^1 from a D-experiment can be interpreted as a choice from a P- or V-experiment with fixed cost, that is, a choice where the DM had to spend at least $p^1_1 x^1_1$ on the first commodity.

Recall that in the standard version of Afriat's Theorem, GARP is equivalent to the existence of numbers $U^i, \lambda^i > 0$, for $i = 1, \ldots, N$, such that $U^i \leq U^j + [\lambda^j p^j] \cdot [x^i - x^j]$. The $[\lambda^j p^j]$ part is a result of the standard budget constraint $p^j a \leq 1$.

C-AG1 We first introduce artificial constraints for the choices $\{z^i\}_{i=1}^N$ from the D- and the P- or V-experiment, which match the constraints used in Theorem 5. Let $M = 2$, $\mathbf{G}_1^i = r^i$, $\mathbf{c}_1^i = 1$, $\boldsymbol{\kappa}_1^i = \lambda^i$, and $\boldsymbol{\kappa}_2^i = \mu^i$ for $i = 1, \ldots, N$. Let $\mathbf{G}_2^i = (0, \ldots, 0)$ and $\mathbf{c}_2^i = 1$ for $i = 1, \ldots, N_x$. Let $\mathbf{G}_2^i = \mathbf{1}_1$ and $\mathbf{c}_2^i = \mathbf{1}_1 z^i$ for $i = N_x + 1, \ldots, N$. Then the first condition of Theorem 5 reduces to the existence of numbers U^i, λ^i, μ^i, such that $U^i \leq U^j + [\lambda^j r^j + \mu^j \mathbf{1}_1] \cdot [z^j - z^i]$ with $\mu^j = 0$ for $j = 1, \ldots, N_x$.

This is exactly the condition C-AG1. The existence of a utility function which rationalises the choices $\{z^i\}_{i=1}^N$ observed under the artificial constraints follows from Theorem 5. That this utility function also AG1-rationalises Ω without the artificial constraints follows from the fact that if z^i is one of the choices from a P- or V-experiment (i.e., $i > N_x$), we have $\mathcal{L}(\mathrm{R}_\mathrm{J}^0 \cap \gtrsim_\mathrm{S}, z^i) = \{a \in \mathbb{A} : r^i \cdot a \leq r^i \cdot z^i \text{ and } a_1 \leq z_1^i\}$, where $a_1 \leq z_1^i \Leftrightarrow \mathbf{1}_1 a \leq \mathbf{1}_1 z^i \Leftrightarrow z^i \gtrsim_\mathrm{S} a$. Thus, C-AG1 is equivalent to the existence of a AG1-rationalising utility function. The existence of a rationalising justice function follows from AG1-GARP. Finally, necessity of AG1-GARP already follows from Theorem 2 (see Section B).

C-AG2 Let $M = 2$, $\mathbf{G}_1^i = r^i$, $\mathbf{c}_1^i = 1$, $\boldsymbol{\kappa}_1^i = \lambda^i$, and $\boldsymbol{\kappa}_2^i = \nu^i$ for $i = 1, \ldots, N$. Let $\mathbf{G}_2^i = \mathbf{1}_{-1} r^i$ and $\mathbf{c}_2^i = [\mathbf{1}_{-1} r^i] z^i$ for $i = 1, \ldots, N_x$. Let $\mathbf{G}_2^i = (0, \ldots, 0)$ and $\mathbf{c}_2^i = 1$ for $i = N_x + 1, \ldots, N$. Then the first condition of Theorem 5 reduces to C-AG2.

If z^i is one of the choices from a D-experiment (i.e., $i \leq N_x$), we have $\mathcal{L}([\mathrm{R}^0 \cap \precsim_\mathrm{S}] \cup \geqq, z^i) = \{a \in \mathbb{A} : r^i \cdot a \leq r^i \cdot z^i \text{ and } (a_1 \geq z_1^i \text{ or } a \leqq z^i)\}$, where $(a_1 \geq z_1^i \text{ or } a \leqq z^i) \Leftrightarrow [\mathbf{1}_{-1} r^i] a \leq [\mathbf{1}_{-1} r^i] z^i \Leftrightarrow (z^i \precsim_\mathrm{S} a \text{ or } z^i \geqq a)$. Thus, C-AG2 is equivalent to the existence of a AG2-rationalising justice function. The existence of a rationalising justice function follows from AG2-GARP, and necessity of AG2-GARP already follows from Theorem 2.

C-AG2-SY Note that we proof necessity and sufficiency of C-AG2-SY before C-AG1-SY.

Let $M = 2$, $\mathbf{G}_1^{i,\ell} = \sigma_\ell(r^i)$, $\mathbf{c}_1^{i,\ell} = 1$, $\boldsymbol{\kappa}_1^{i,\ell} = \lambda^{i,\ell}$, and $\boldsymbol{\kappa}_2^{i,\ell} = \nu^{i,\ell}$ for $i = 1, \ldots, N$ and $\ell = 1, \ldots, L!$. Let $\mathbf{G}_2^{i,\ell} = \sigma_\ell(\mathbf{1}_{-1}) \sigma_\ell(r^i)$ and $\mathbf{c}_2^{i,\ell} = [\sigma_\ell(\mathbf{1}_{-1}) \sigma_\ell(r^i)] \sigma_\ell(z^i)$ for $i = 1, \ldots, N_x$ and $\ell = 1, \ldots, L!$. Let $\mathbf{G}_2^{i,\ell} = (0, \ldots, 0)$ and $\mathbf{c}_2^{i,\ell} = 1$ for $i = N_x + 1, \ldots, N$. Then the first condition of Theorem 5 reduces to the first part of C-AG2-SY.

Assuming symmetry, if y^i from a P- or V-experiment is the choice on $B(q^i)$, the DM would have chosen $\sigma_\ell(y^i)$ on $B(\sigma_\ell(y^i))$. Furthermore, we must have $u(a) = u(\sigma_\ell(a))$ for all $\ell = 1, \ldots, L$ and all utility functions which AG2-SY-rationalise the choices. This is assured by the second condition, $U^{i,\ell} = U^{i,m}$, as the $U^{i,\ell}$ are the utility values assigned to $\sigma_\ell(z^i)$ in the construction of the utility function in the proof of Theorem 5. For $i > N_x$ we have $\nu^{i,\ell} = 0$. Thus, the case is the same as for C-AG2, except that all the permutations of all y^i are added.

For $i \leq N_x$ we have $\nu^{i,\ell} > 0$. Therefore, we only need to consider the choices x^i from the D-experiment and their permutations. Recall that the interpretation is that z^i from a D-experiment chosen on $B(r^i)$ is as if the DM demanded z^i in a P- or V-experiment with an additional fixed costs constraint, that is, $[\mathbf{1}_{-1} r^i] a \leq [\mathbf{1}_{-1} r^i] z^i$. Then by symmetry, the DM would have chosen $\sigma_\ell(z^i)$ on

$B(\sigma_\ell(r^i))$ with the additional constraint $[\mathbf{1}_{-1}\sigma_\ell(r^i)]a \leq [\mathbf{1}_{-1}\sigma_\ell(r^i)]\sigma_\ell(z^i)$. Thus, again, the case is the same as for C-Ag2 with the permutations of x^i added.

C-Ag1-Sy Let $M = 3$, $\mathbf{G}_1^{i,\ell} = \sigma_\ell(r^i)$, $\mathbf{c}_1^{i,\ell} = 1$, $\boldsymbol{\kappa}_1^{i,\ell} = \lambda^{i,\ell}$, $\boldsymbol{\kappa}_2^{i,\ell} = \mu^{i,\ell}$, and $\boldsymbol{\kappa}_3^{i,\ell} = \nu^{i,\ell}$ for $i = 1, \ldots, N$ and $\ell = 1 \ldots, L!$. Let

$$\mathbf{G}_2^{i,\ell} = \begin{cases} \mathbf{1}_1 & \text{if } i > N_x \text{ or } \ell > 1, \\ (0, \ldots, 0) & \text{otherwise,} \end{cases}$$

$$\mathbf{c}_2^{i,\ell} = \begin{cases} \mathbf{1}_1\sigma_\ell(z^i) & \text{if } i > N_x \text{ or } \ell > 1, \\ 1 & \text{otherwise,} \end{cases}$$

$$\mathbf{G}_3^{i,\ell} = \begin{cases} \sigma_\ell(\mathbf{1}_{-1})\sigma_\ell(r^i) & \text{if } i \leq N_x \text{ and } \ell > 1, \\ (0, \ldots, 0) & \text{otherwise,} \end{cases}$$

$$\mathbf{c}_3^{i,\ell} = \begin{cases} [\sigma_\ell(\mathbf{1}_{-1})\sigma_\ell(r^i)]\sigma_\ell(z^i) & \text{if } i \leq N_x \text{ and } \ell > 1, \\ 1 & \text{otherwise.} \end{cases}$$

Then the first condition of Theorem 5 reduces to the first part of C-Ag1-Sy.

We only need to distinguish two cases: (1) $i > N_x$ with $\mu^{i,\ell} \geq 0$ and $\nu^{i,\ell} = 0$ and (2) $[i \leq N_x$ and $\ell > 1]$ with $\mu^{i,\ell} \geq 0$ and $\nu^{i,\ell} \geq 0$. Case (1) with $\ell = 1$ is the same as in C-Ag1, and with $\ell > 1$ the same reasoning for symmetry applies as in the proof of C-Ag2-Sy. In case (2) the artificial constraints for choices from a d-experiment for Ag2-rationalisation are combined with the artificial constraints for choices from a p- or v-experiment for Ag1-rationalisation. This is because, by the definition of $\hat{\mathrm{R}}$, the only part of the directly revealed worse set of a choice x^i in the d-experiment which is to be permuted (or "mirrored") is the part that would also be revealed worse when x^i were a choice in the p- or v-experiment (recall that $\hat{\mathrm{R}}$ is constructed based on $\hat{\mathrm{R}}_\mathrm{J}$). So suppose that hypothetically x^i, observed as a choice in a d-experiment, were a choice in the p- or v-experiment. Then the constraint $[\sigma_1(\mathbf{1}_{-1})\sigma_1(p^i)]a \leq [\sigma_1(\mathbf{1}_{-1})\sigma_1(p^i)]\sigma_1(x^i)$ determines the part of the budget that is revealed less just than x^i, and the constraint $\mathbf{1}_1 a \leq \mathbf{1}_1 x^i$ determines the part of the budget that is revealed worse. This is obviously redundant, which is why $\mu^{i,1} = \nu^{i,1} = 0$. But for $\sigma_\ell(x^i)$ with $\ell > 1$ the constraints are not redundant and determine the part of the budget described by the permutation of p^i that is both less just and provides less a_1 to the dm (i.e., $a_1 \leq x_1^i$). Then we can apply the same reasoning as for C-Ag1 and C-Ag2-Sy. ■

A.2 *Recoverability in the Higher-Dimensional Case*

In this section, we will sketch a way to do the same kind of analysis as in Section 3 in the higher dimensional case. The approach we present is based on the idea of finding *virtual price vectors* for observations from a p- or v-experiment when constructing the revealed preference relation, and observations from a d-experiment when constructing the revealed justice relation. The idea is based on the work of Fleissig and Whitney (2011), who compute such virtual price vectors for observed demand with rationing. We will not provide any formal proofs, but the approach is rather straightforward. We only consider the AG1 and AG2 cases here, but the approach can be generalised for AG1-SY and AG2-SY as well.

Consider the extended revealed preference relation based on AG1. Recall the necessary and sufficient condition C-AG1 for AG1-rationalisation and suppose that it is satisfied. We have $\lambda^i r^i + \mu^i \mathbf{1}_1 = \lambda^i(r_1^i + [\mu^i/\lambda^i], r_2^i, \ldots, r_L^i)$. Let $\theta^i = \mu^i/\lambda^i$, which is well defined as $\lambda^i > 0$. We have $\theta^i > 0$ if and only if $\mu^i > 0$. Then we can construct the *virtual price vectors* $\bar{r}^i = (r_1^i + \theta^i, r_2^i, \ldots)$. By Theorem 4 (also see Fleissig and Whitney (2011), who use the same construction based on Varian's (1983) Theorem 7), the data $\bar{\Omega} = \{z^i, \bar{r}^i\}_{i=1}^N$ satisfies GARP.

Also note that *all* virtual price vectors such that $\bar{\Omega}$ satisfy GARP and the utility function which rationalises $\bar{\Omega}$ also AG1-rationalises Ω must be of the above form. If not, then either $z^i \notin B(\bar{r}^i)$ or $\mathcal{L}(\mathrm{R}_\mathrm{J}^0 \cap \succsim_\mathrm{S}, z^i) \notin B(\bar{r}^i)$. We can therefore compute the minimal and maximal θ^i for each z^i, which together completely describe the set of feasible virtual prices. Then based on the same reasoning as for the sets $\mathcal{RP}$ and $\mathcal{RW}$ in Section 3.2, we can conclude that if for some z^j we have that $\bar{r}^i \cdot z^i \geq \bar{r}^i \cdot z^j$ for all feasible $\bar{r}^i$, then we must have $u(z^i) \geq u(z^j)$ for all continuous, monotonic, and concave utility functions u which AG1-rationalise the data.

This then provides us with a way to compute all z^k which must be preferred to z^j, and we can apply the same way to construct the extended revealed preferred set based on the convex monotonic hull. This follows from the same arguments as in the proof of Proposition 3 (see Section B below).

Now consider the extended revealed justice relation based on AG2. We have $\lambda^i r^i + \nu^i[\mathbf{1}_{-1} r^i] = \lambda^i(r_1^i, r_2^i + [(\nu^i r_2^i)/\lambda^i], r_3^i + [(\nu^i r_3^i)/\lambda^i], \ldots)$. Let $\theta^i = (\nu^i r_2^i)/\lambda^i$; then again we can construct virtual price vectors $\bar{r}^i = (r_1^i, r_2^i + \theta^i, \ldots)$. Then we can use a similar approach as the one sketched for AG1.

B APPENDIX: PROOFS

B.1 *Proof of Fact 1*

Suppose AG1 is not satisfied. Then there exist $x, y \in \mathbb{A}$ such that $x[\succsim_\mathrm{J} \cap \succ_\mathrm{S}]y$ and $y \succ x$, thus $y[\succ \cap \prec_\mathrm{S}]x$. Then AG2 implies that $y \succsim_\mathrm{J} x$, which implies $y \sim_\mathrm{J} x$. By continuity, $y \succ x$ implies that there exists an $\varepsilon > 0$ such that for all $z \in N_\varepsilon(x)$, $y \succ z$. Furthermore, $y \sim_\mathrm{J} x$ implies that for small enough $\varepsilon > 0$ there exist a $z \in N_\varepsilon(x)$ such that $z \succ_\mathrm{J} y$, and $x \succ_\mathrm{S} y$ implies that for all $z \in N_\varepsilon(x)$, $z \succ_\mathrm{S} y$. Thus $y[\succ \cap \prec_\mathrm{S}]z$ and $z \succ_\mathrm{J} y$, which violates AG2. Thus, AG2 implies AG1, and analogously for the reverse.

B.2 *Proof of Fact 2*

Note that SY implies that $\succsim_\mathrm{J} \neq \succsim_\mathrm{S}$. It is obvious that $[\succsim_\mathrm{J} \cap \succsim_\mathrm{S}] \subseteq \succsim$ implies $[\succsim_\mathrm{J} \cap \succsim_\mathrm{S}] \subset \succsim$. So suppose $x[\succsim_\mathrm{J} \cap \sim_\mathrm{S}]y$. Then by monotonicity of $\succsim_\mathrm{J}$ and because $\succsim_\mathrm{J} \neq \succsim_\mathrm{S}$, $x \geq y$. But then by monotonicity of $\succsim$, $x \succsim y$. Suppose that $\succsim = \succsim_\mathrm{S}$. Then $[\succsim \cap \prec_\mathrm{S}] = \varnothing$. So suppose $\succsim \neq \succsim_\mathrm{S}$. Then if $x[\succsim \cap \sim_\mathrm{S}]y$, then $x \geq y$, and monotonicity of $\succsim_\mathrm{J}$ implies $x \succsim y$.

B.3 *Proof of Fact 3*

Suppose $\succsim_J$ is monotonic and $y \geq x$. Then $y \succsim_S x$ and by monotonicity $y \succsim_J x$. Then by Ag and Fact 2, $y \succsim x$ follows.

Suppose that $\succsim = \succsim_S$. Then $[\succsim \cap \prec_S] = \varnothing$. So suppose $\succsim$ is monotonic, $\succsim \neq \succsim_S$, and $y \geq x$ with $y_j = x_j$ for at least one and at most $L-1$ indices $j \in \{1, \ldots, L\}$. Then $\sigma_\ell(y) \geq \sigma_\ell(x)$ and therefore by monotonicity $\sigma_\ell(y) \succsim \sigma_\ell(x)$ for all $\ell = 1, \ldots, L!$. Then for at least one $k \in \{1, \ldots, L!\}$, $\sigma_k(y) \sim_S \sigma_k(x)$, and therefore $\sigma_k(y) [\succsim \cap \sim_S] \sigma_k(x)$. Then by Ag and Fact 2, $\sigma_k(y) \succsim_J \sigma_k(x)$ and by Sy $y \succsim_J x$ follows.

Suppose instead that $y > x$. Then let $\tilde{x} = x + (y_1 - x_1, 0, \ldots, 0)$. Then $y \geq \tilde{x} \geq x$, but neither $y > \tilde{x}$ nor $\tilde{x} > x$. Then it follows from the arguments in the preceding paragraph that $y \succsim_J \tilde{x}$ and $\tilde{x} \succsim_J x$, and by transitivity $y \succsim_J x$ follows.

B.4 *Proof of Fact 4*

When $u(a)$ and $v(a)$ are representations of $\succsim$ and $\succsim_J$, then Ag and Sy follow immediately. We show the reverse here to establish equivalence.

Sy immediately implies that $\beta_i = \beta_j$ for all i, j, and $v(a) = \left(\sum_{i=1}^L a_i^s\right)^{1/s}$ represents the same $\succsim_J$ as $\tilde{v}(a)$. Assume for simplicity that for $\tilde{u}(a)$ the parameters are normalised such that $\sum_{i=1}^L \alpha_i = 1$ with $\alpha_i > 0$.

If $\succsim = \succsim_J$, the fact follows immediately. If $\succsim = \succsim_S$, then $\tilde{u}(a) = (\alpha_1 a_1^r)^{1/r}$, and $\succsim$ is also represented by $u(a) = (\alpha_1 a_1^s)^{1/s}$. Then the convexity parameter is the same for both functions and the fact follows. So suppose for the rest of the proof that $\succsim \neq \succsim_J$ and $\succsim \neq \succsim_S$.

Suppose for some $a \in \mathbb{A}$ with $a_i \neq a_j$ for some $i, j \neq 1$, $a' \in \mathbb{A}$ is obtained by exchanging the ith and jth entry, that is, $a'_i = a_j$ and $a'_j = a_i$, and suppose that $v(a) = v(a')$. Then Ag and Sy imply $u(a) = u(a')$. But if $\alpha_i \neq \alpha_j$ then $u(a) \neq u(a')$, a contradiction. Therefore, $\alpha_i = \alpha_j$ for all $j, k \neq 1$.

Suppose for some $a \in \mathbb{A}$ with $a_1 > a_i$ for some $i \neq 1$, $a' \in \mathbb{A}$ is obtained by exchanging the first and ith entry, that is, $a'_1 = a_i$ and $a'_i = a_1$, , and suppose that $v(a) = v(a')$. Then Ag and Sy imply $u(a) < u(a')$. But if $\alpha_1 < \alpha_i$ then $u(a) > u(a')$, a contradiction. Therefore, $\alpha_1 > \alpha_i$ for all $i \neq 1$.

It then follows that $u(a) = \left(\alpha a_1^r + \sum_{i=2}^L \frac{1-\alpha}{L-1} a_i^r\right)^{1/r}$ with $\alpha = 1 - \sum_{i=2}^L a_i > 1/L$ represents the same $\succsim$ as $\tilde{u}(a)$. What is left to show is that $r = s$, that is, the convexity parameter of both functions has the same value t as stated.

We proceed by showing that if $r \neq s$, then for some $a \in \mathbb{A}$, u and v have the same gradient. Consider $a \in \mathbb{A}$ with $a_i = a_j$ for all $i, j \neq 1$, and suppose $r \neq s$. Then $\nabla u = \nabla v$ at a only implies equality of the marginal rate of substitution between a_1 and any other a_i for both u and v:

$$\frac{\partial u / \partial a_1}{\partial u / \partial a_i} = \frac{\partial v / \partial a_1}{\partial v / \partial a_i} \quad \text{for all } i \neq 1,$$

and we obtain

$$\left(\frac{a_1}{a_i}\right)^{s-r} = \frac{\alpha(L-1)}{1-\alpha}.$$

With $\alpha \in (1/L, 1)$, the right hand side is defined and strictly greater than 1. Thus, for $s \neq r$, there exist $a \in A$ such that $\nabla u = \nabla v$: For $a_i = \lambda > 0$ for all $i \neq 1$, we obtain $a_1 = \lambda[([L-1]\alpha)/(1-\alpha)]^{1/(s-r)}$.

Suppose $r < s$. Then with $\hat{a} = ([([L-1]\alpha)/(1-\alpha)]^{1/(s-r)}, 1, \ldots, 1)$, there exists an $a' \neq \hat{a}$ with $a'_1 > \hat{a}_1$ such that $v(\hat{a}) = v(a')$ and $u(a') = u(\hat{a}/[1+\varepsilon])$ for some $\varepsilon > 0$. But as $u(\hat{a}/[1+\varepsilon]) < u(\hat{a})$, $u(\hat{a}) > u(a')$, which violates Ag. Suppose instead $r > s$. Then there exists an $a' \neq \hat{a}$ with $a'_1 > \hat{a}_1$ such that $u(\hat{a}) = u(a')$ and $v(a') = v(\hat{a}/[1+\varepsilon])$ for some $\varepsilon > 0$. But as $v(\hat{a}/[1+\varepsilon]) < v(\hat{a})$, $v(\hat{a}) > v(a')$, which violates Ag.

B.5 *Proof of Proposition 2*

The equivalence of the weak forms of Ag1-Garp and Ag2-Garp can be shown by exploiting the possibility that in the two-dimensional case, price vectors can be uniquely sorted after normalising one of the two prices or alternatively, that the vertices (see below) of a convex monotonic hull can be sorted from "right" to "left", that is, by $\succ_S$. The result here is a bit more tedious to prove due to the Ag-extensions, but it is not particularly surprising given previous results. For example, Rose (1958) and Heufer (2013a) show that in the two-dimensional case, the Weak Axiom of Revealed Preference is equivalent to the Strong Axiom of Revealed Preference. Rose's proof is based on normalising price vectors, while Heufer's proof uses basic geometric definitions. Similar to Rose (1958), Banerjee and Murphy (2006) show that W-Garp is equivalent to Garp if $L = 2$ (see Proposition 1).

We start with two Lemmata, which will also be helpful for the proof of Theorem 2 below. We will only proof the first statement of the proposition. The second statement is a straightforward extension based on symmetry of the revealed more just relation. We also need some geometrical definitions. A *vertex* is a corner point of polytope. In two dimensions and on a convex monotonic hull C, a vertex is the intersection of two edges of C. Suppose C is the convex monotonic hull of a set of points $\{a^i\}_{i=1}^m$. Then the set of vertices of C is a (not necessarily proper) subset of $\{a^i\}_{i=1}^m$. Two vertices a^j and a^k of C are *adjacent* if the line connecting a^j and a^k is an edge on the boundary of C. Then in two dimensions, this line forms a supporting hyperplane of C. See for example Brøndsted (1983) or Grünbaum (2003) for more detailed definitions and Heufer (2013a) for an application to revealed preference.

Lemma 1 *Suppose $L = 2$. If Ω satisfies Ag1-W-Garp, then $z^0 \in \partial CMH(\{z^i : z^i \tilde{R}\, z^0\})$. If Ω satisfies Ag2-W-Garp, then $z^0 \in \partial CMH(\{z^i : z^i \tilde{R}_J\, z^0\})$.*

Proof of Lemma 1 We only proof the statement based on Ag1-W-Garp; the proof for Ag2-W-Garp works analogously. Let $C = CMH(\{z^i : z^i \tilde{R}\, z^0\})$. Let $Z = \{z^i : z^i \tilde{R}\, z^0\} \backslash \{z^0\}$, $Z^0 = \{z^i : z^i \tilde{R}^0 z^0\}$ and $\bar{Z} = Z \cap \partial C$. Without loss of generality with respect to the indices, set $\{z^i\}_{i=1}^{\ell} = \bar{Z}$ be such that $z^i \succ_S z^{i+1}$ for $i = 1, \ldots, \ell - 1$. Refer to Figure 13 for an illustration which might be helpful to understand the proof.

Step 1 Suppose Ag1-W-Garp is satisfied. For $Z = \varnothing$ the statement is trivially true. So suppose $Z \neq \varnothing$ and $z^0 \in \text{int}C$. Suppose there is no $z^m \in Z^0$ such that $z^m \in \text{int}C$. Then we must have $z^i \tilde{R}^0 z^0$ for at least one $z^i \in \bar{Z}$; otherwise, $z^i \notin C$. Suppose a set of observed choices $z^m, z^{m'}, \ldots \in Z$ is in the interior of C, such that at least one $z^m \tilde{R}^0 z^0$. Then either $z^i \tilde{R}^0 z^0$ or $z^i \tilde{R}^0 z^m$ for at least one $z^i \bar{Z}$;

otherwise, $z^i \notin C$. Thus, at least one $z^i \in \bar{Z}$ must be such that $z^i\, \tilde{\mathrm{R}}^0\, z^m$, with $z^m \in \mathrm{int}C$ (and possibly $z^m = z^0$).

We now show that $z^i\, \tilde{\mathrm{R}}^0\, z^{i+1}$ for $i = 1, \ldots, \ell - 1$.

Step 2 Consider z^1. As $z^1 \in \partial C$, we must have $z^1\, \tilde{\mathrm{R}}^0\, z^j$ for some $z^j \in C$. Then $z^j \in B(r^1)$, which implies that $B(r^1) \cap C \neq \{z^1\}$. As $z^1 \succ_S z^2$, we must have $z^1\, \tilde{\mathrm{R}}^0\, z^2$.

Step 3 Consider z^2. By Ag1-W-Garp, we cannot have $z^2\, \tilde{\mathrm{P}}^0\, z^1$, and if $z^2\, \tilde{\mathrm{R}}^0\, z^1$, then we cannot have $z^1\, \tilde{\mathrm{P}}^0\, z^2$. Then if $z^1\, \tilde{\mathrm{R}}^0\, z^2\, \tilde{\mathrm{R}}^0\, z^1$, then $\{z^1, z^2\} \subset \partial B(r^1)$ and $B(r^1) = B(r^2)$. Because $\{z^1, z^2\} \in \partial C$, $B(r^1)$ and $B(r^2)$ are supporting hyperplanes of C. Then both z^1 and z^2 can only be preferred to other choices in C if both are preferred to z^3 and $z^3 \in \partial B(r^2)$, and as $z^2 \succ_S z^3$, this implies $z^2\, \tilde{\mathrm{R}}^0\, z^3$. If instead [not $z^2\, \tilde{\mathrm{R}}^0\, z^1$], then again we must have $z^3 \in B(r^2)$. As $z^2 \succ_S z^3$, this implies $z^2\, \tilde{\mathrm{R}}^0\, z^3$. In either case, $z^2\, \tilde{\mathrm{R}}^0\, z^3$ follows. Simple induction then shows that we must have $z^i\, \tilde{\mathrm{R}}^0\, z^{i+1}$ for all $i = 1, \ldots, \ell - 1$.

Next we show that $z^i\, \tilde{\mathrm{R}}^0\, z^{i-1}$ for $i = 2, \ldots, \ell$.

Step 4 Consider z^ℓ, which is the "left most" element in $\bar{Z}$. As $z^\ell \in \partial C$, we must have $z^\ell\, \tilde{\mathrm{R}}^0\, z^j$ for some $z^j \in C$. As $z^\ell \precsim_S z$ for all $z \in C$, we cannot have $z^\ell[\mathrm{R}_J^0 \cap \succ_S]z^j$. But as $z^\ell\, \tilde{\mathrm{R}}^0\, z^j$, z^ℓ must be an observation from a d-experiment, that is, $z^\ell\, \mathrm{R}^0\, z^j$. But then $z^\ell\, \mathrm{R}^0\, z^{\ell-1}$.

Step 5 Based on Step 4, it is now obvious that we can follow the same arguments as in Step 3 to show that $z^i\, \tilde{\mathrm{R}}^0\, z^{i-1}$ for $i = 2, \ldots, \ell$.

Step 6 In Step 1, we have shown that for at least one z^i on ∂C, $z^i\, \tilde{\mathrm{R}}^0\, z^m$ with $z^m \in \mathrm{int}C$. If $i = 1$, then $z^1\, \tilde{\mathrm{R}}^0\, z^m$ implies $z^1\, \tilde{\mathrm{P}}^0\, z^2$, but by Step 5, we have $z^2\, \tilde{\mathrm{R}}^0\, z^1$, which contradicts Ag1-W-Garp. If $i = \ell$, then $z^\ell\, \tilde{\mathrm{R}}^0\, z^m$ implies $z^\ell\, \tilde{\mathrm{P}}^0\, z^{\ell-1}$, but by Step 3, we have $z^{\ell-1}\, \tilde{\mathrm{R}}^0\, z^\ell$, which contradicts Ag1-W-Garp. Suppose $1 < i < \ell$ and $z^i\, \tilde{\mathrm{R}}^0\, z^m$. Then $z^i \succ_S z^m$ implies $z^i\, \tilde{\mathrm{P}}^0\, z^{i+1}$, while $z^i \prec_S z^m$ implies $z^i\, \tilde{\mathrm{P}}^0\, z^{i-1}$, but again by Steps 5 and 3, respectively, this contradicts Ag1-W-Garp. ■

Lemma 2 *Suppose $L = 2$. Let $C = CMH(\{z^i : z^i\, \tilde{\mathrm{R}}\, z^0\})$ and $C_J = CMH(\{z^i : z^i\, \tilde{\mathrm{R}}_J\, z^0\})$. Without loss of generality with respect to the indices, let the set $\{z^i\}_{i=1}^{\ell}$ be the set of observed choices on ∂C or ∂C_J such that $z^i \succ_S z^{i+1}$ for $i = 1, \ldots, \ell-1$. If Ω satisfies Ag1-W-Garp, then $z^1\, \tilde{\mathrm{R}}^0\, z^2\, \tilde{\mathrm{R}}^0\, z^3\, \tilde{\mathrm{R}}^0 \ldots \tilde{\mathrm{R}}^0\, z^j\, \tilde{\mathrm{R}}^0\, z^0$ and $z^\ell\, \tilde{\mathrm{R}}^0\, z^{\ell-1}\, \tilde{\mathrm{R}}^0\, z^{\ell-2}\, \tilde{\mathrm{R}}^0 \ldots \tilde{\mathrm{R}}^0\, z^{j+1}\, \tilde{\mathrm{R}}^0\, z^0$ for some $j \in \{1, \ldots, m\}$. If Ω satisfies Ag2-W-Garp, then $z^1\, \tilde{\mathrm{R}}_J^0\, z^2\, \tilde{\mathrm{R}}_J^0\, z^3\, \tilde{\mathrm{R}}_J^0 \ldots \tilde{\mathrm{R}}_J^0\, z^j\, \tilde{\mathrm{R}}_J^0\, z^0$ and $z^\ell\, \tilde{\mathrm{R}}_J^0\, z^{\ell-1}\, \tilde{\mathrm{R}}_J^0\, z^{\ell-2}\, \tilde{\mathrm{R}}_J^0 \ldots \tilde{\mathrm{R}}_J^0\, z^{j+1}\, \tilde{\mathrm{R}}_J^0\, z^0$ for some $j \in \{1, \ldots, m\}$.*

Proof of Lemma 2 We only proof the statement based on Ag1-W-Garp; the proof for Ag2-W-Garp works analogously. The proof follows closely the steps in the proof of Lemma 1 and we will omit most of it. It might again be helpful to refer to Figure 13 and replace, for example, z^3 with z^0.

By Lemma 1 we know that $z^0 \in \partial C$. Let z^j be the choice in $\{z^i\}_{i=1}^{\ell}$ such that the pairs (z^j, z^0) and (z^0, z^{j+1}) are on the same supporting hyperplane of C, respectively. Then following Steps 2 and 3 in the proof of Lemma 1 shows that $z^1\, \tilde{\mathrm{R}}^0\, z^2\, \tilde{\mathrm{R}}^0 \ldots, \tilde{\mathrm{R}}^0\, z^j\, \tilde{\mathrm{R}}^0\, z^0$. Following Steps 4 and 5 in

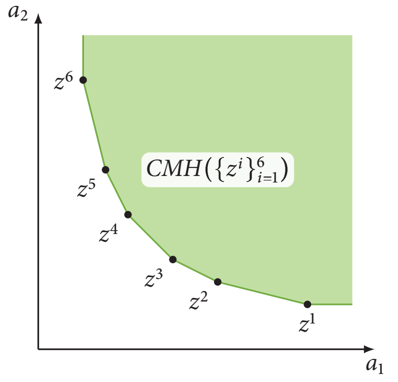


Figure 13: An illustration for the proof of the Lemmata.

the proof of Lemma 1 shows that $z^\ell \, \tilde{R}^0 \, z^{\ell-1} \, \tilde{R}^0 \, \ldots, \tilde{R}^0 \, z^{j+1} \, \tilde{R}^0 \, z^0$. Then the statement in the Lemma already follows.

Note that the Lemma also implies that if some z^i and z^0 are adjacent on C, then $z^i \, \tilde{R}^0 \, z^0$.

Note that $z^j \, \tilde{R}^0 \, z^0$ and $z^{j+1} \, \tilde{R}^0 \, z^0$ are both possible without violating Ag1-W-Garp, because by Lemma 1, $z^0 \in \partial C$. Also, $z^j \, \tilde{P}^0 \, z^0$ and $z^{j+1} \, \tilde{P}^0 \, z^0$ are both possible without violating Ag1-W-Garp if z^0 is a vertex on C (i.e., a corner point). ■

We are now ready to proof the first part of the Proposition, which states that in the two-dimensional case, the following conditions are equivalent:

(1) Ag1-Garp,
(2) Ag1-W-Garp,
(3) Ag2-Garp,
(4) Ag2-W-Garp.

That (1) ⇒ (2) and (3) ⇒ (4) is obvious. We will first prove that (2) ⇒ (1). The proof that (4) ⇒ (3) is practically the same and we omit it. We will then prove that (2) ⇒ (4). The proof that (4) ⇒ (2) is practically the same and we omit it.

(2) ⇒ (1) Suppose that without loss of generality with respect to the indices, we have $z^1 \, \tilde{R}^0 \, z^2 \, \tilde{R}^0 \ldots \tilde{R}^0 \, z^{\ell-1} \, \tilde{R}^0 \, z^\ell$. We will show that $z^\ell \, \tilde{P}^0 \, z^1$ violates Ag1-W-Garp. Let $C = CMH(\{z^i : z^i \, \tilde{R} \, z^\ell\})$. By Lemma 1, $z^\ell \in \partial C$. Suppose $z^\ell \, \tilde{P}^0 \, z^1$. Then if $z^1 >_S z^\ell$, $z^\ell \, \tilde{P}^0 \, z^1$ implies that $z^\ell \, P^0 \, z^1$. But then, as $z^1 \in C$, we must have $z^\ell \, P^0 \, z^j$ where $z^j >_S z^\ell$ and z^j and z^ℓ are adjacent. Then it follows with Lemma 2 that $z^j \, \tilde{R}^0 \, z^\ell$, which violates Ag1-W-Garp. If instead $z^1 <_S z^\ell$, $z^\ell \, \tilde{P}^0 \, z^1$ implies that $z^\ell \, \tilde{P}^0 \, z^k$ where $z^k <_S z^\ell$ and z^k and z^ℓ are adjacent. Again it follows with Lemma 2 that $z^k \, \tilde{R}^0 \, z^\ell$, which violates Ag1-W-Garp.

(2) ⇒ (4) Suppose that Ag1-W-Garp is violated such that $z^i \, \tilde{R}^0 \, z^j \, \tilde{P}^0 \, z^i$. We will show that this implies that Ag2-W-Garp is violated. We either have (i) $z^i \, R^0 \, z^j \, P^0 \, z^i$, or (ii) $z^i [R^0_J \cap >_S] z^j \, P^0 \, z^i$, or

(iii) $z^i \, \mathrm{R}^0 \, z^j [\mathrm{P}^0_\mathrm{J} \cap \succ_\mathrm{S}] z^i$, or (iv) $z^i [\mathrm{R}^0_\mathrm{J} \cap \succ_\mathrm{S}] z^j [\mathrm{P}^0_\mathrm{J} \cap \succ_\mathrm{S}] z^i$. Case (i) violates W-Garp and therefore Ag2-W-Garp. Case (iv) is impossible as we cannot have $z^i \succ_\mathrm{S} z^j \succ_\mathrm{S} z^i$.

Consider case (ii). We have $z^j \prec_\mathrm{S} z^i$ but also $z^j \, \mathrm{P}^0 \, z^i$, which implies $z^j \, \tilde{\mathrm{P}}^0_\mathrm{J} \, z^i$. But $z^i \, \mathrm{R}^0_\mathrm{J} \, z^j$, which violates Ag2-W-Garp. Consider case (iii). We have $z^i \prec_\mathrm{S} z^j$ but also $z^i \, \mathrm{R}^0 \, z^j$, which implies $z^i \, \tilde{\mathrm{R}}^0_\mathrm{J} \, z^j$. But $z^j \, \mathrm{P}^0_\mathrm{J} \, z^i$, which violates Ag2-W-Garp.

B.6 *Proof of Theorem 2*

Necessity of Ag-W-Garp is straightforward to show with similar arguments as in the proof of Afriat's Theorem; see for example Varian's (1982) proof. Sufficiency will be shown by explicitly constructing virtual price vectors. The proof for Ag-Sy-W-Garp is somewhat more involved, but as in two dimensions symmetry means that budgets and observations are simply "mirrored". It works analogously and we omit it.

Suppose Ag-W-Garp is satisfied. Let $C = CMH(\{z^i : z^i \, \tilde{\mathrm{R}} \, z^m\})$. By Lemma 1, $z^m \in \partial C$. By Lemma 2, if z^m and z^i are adjacent on ∂C, then $z^i \, \tilde{\mathrm{R}}^0 \, z^m$ and by Ag-Garp, [not $z^m \, \tilde{\mathrm{P}}^0 \, z^i$]. Then $B(r^m) \cap \mathrm{int} C = \varnothing$, which implies $\mathrm{int}\mathcal{L}(\tilde{\mathrm{R}}^0, z^m) \cap C = \varnothing$. As both C and $\mathcal{L}(\tilde{\mathrm{R}}^0, z^m)$ are convex, there exists, by the supporting hyperplane theorem, a hyperplane which supports both C and $\mathcal{L}(\tilde{\mathrm{R}}^0, z^m)$. This hyperplane has to be of the form $\{a \in \mathbb{A} : (r^m_1 + \theta^m, r^m_2)[a - z^m] = 0\}$ with $\theta^m \geq 0$.

If z^m is a choice from a d-experiment, we can let $\theta^m = 0$, as $\mathcal{L}(\tilde{\mathrm{R}}^0, z^m) = B(r^m)$. So suppose z^m is a choice from a p- or v-experiment. There are two cases: (i) there exists no z^j such that $z^j [\tilde{\mathrm{R}} \cap \succ_\mathrm{S}] z^m$ and $r^m[z^m - z^j] \geq 0$, or (ii) there does. In case (i), we let $\theta^m = 0$. In case (ii), let z^k be the observation that is on C, adjacent to z^k, with $z^k \succ_\mathrm{S} z^m$. Then by Lemma 2, $z^k [\tilde{\mathrm{R}}^0 \cap \succ_\mathrm{S}] z^m$ and $r^m[z^m - z^k] > 0$. We let

$$\check{\theta}^m = \frac{r^m[z^m - z^k]}{z^k_1 - z^m_1}.$$

and $\theta^m = \check{\theta}^m + \varepsilon$ with $\varepsilon \geq 0$. For all the possible cases, the virtual price vector for z^m is then $\bar{r}^m = (r^m_1 + \theta^m, r^m_2)$. In case (ii), for $\varepsilon = 0$, we have $\bar{r}^m[z^m - z^k] = 0$; see Figure 14.(a) for an illustration.

We now distinguish two further subcases: (ii.a) there exists no z^ℓ such that $z^\ell [\tilde{\mathrm{R}} \cap \prec_\mathrm{S}] z^m$ and $\bar{r}^m[z^m - z^\ell] \geq 0$, or (ii.b) there does. In case (ii.a) we can choose $\varepsilon > 0$ so small that $\bar{r}^m[z^m - z^k] < 0$ and $\bar{r}^m[z^m - z^i] < 0$ for all $z^i [\tilde{\mathrm{R}} \cap \prec_\mathrm{S}] z^m$. In case (ii.b), suppose that for $\varepsilon = 0$, $\bar{r}^m[z^m - z^\ell] = 0$; this is illustrated in Figure 14.(b). Then by Lemma 2, we must have $z^\ell \, \mathrm{R}^0 \, z^m$ and $z^k \, \tilde{\mathrm{R}}^0 \, z^m$, and therefore $z^\ell \, \mathrm{R}^0 \, z^k$ and $z^k \, \tilde{\mathrm{R}}^0 \, z^\ell$. Then $z^\ell \, \tilde{\mathrm{P}}^0 \, z^m$ (which implies $z^\ell \, \tilde{\mathrm{P}}^0 \, z^k$) or $z^k \, \tilde{\mathrm{P}}^0 \, z^m$ (which implies $z^k \, \tilde{\mathrm{P}}^0 \, z^\ell$) would violate Ag-Garp. Then we can let $\varepsilon = 0$, and $\bar{r}^m[z^m - z^k] = \bar{r}^m[z^m - z^\ell] = 0$; then if R was based on $\bar{r}^m$, Garp would not be violated. Suppose that instead, for $\varepsilon = 0$, $\bar{r}^m[z^m - z^\ell] > 0$; this is illustrated in Figure 14.(c). Then again we must have $z^\ell \, \tilde{\mathrm{R}}^0 \, z^j$ and $z^k \, \tilde{\mathrm{R}}^0 \, z^j$. But this implies $z^\ell \, \tilde{\mathrm{P}}^0 \, z^k$ and $z^k \, \tilde{\mathrm{P}}^0 \, z^\ell$, which violates Ag-Garp.

We can therefore set $\bar{r}^m = (r^m_1 + \theta^m, r^m_2)$ with either $\varepsilon = 0$ or some small $\varepsilon > 0$ for all possible cases and base R on $\bar{r}^m$ without violating Garp. Then all that is left to verify is that there does not exist a z^k such that $\bar{r}^m[z^m - z^k] \geq 0$ and $\bar{r}^k[z^k - z^m] \geq 0$ with one or both inequalities holding strictly. If $\bar{r}^k[z^k - z^m] \geq 0$, then there must exist a $z^\ell \succ_\mathrm{S} z^k$ with $z^\ell \, \tilde{\mathrm{R}}^0 \, z^k$ such that $(r^k_1 + \check{\theta}^k, r^k_2)[z^k - z^\ell] = 0$, which is illustrated in Figure 14.(d). But then $z^\ell[\tilde{\mathrm{R}}^0 \cap \gtrsim_{\mathrm{S}}]z^m$, and we can apply the same arguments as before to show that Ag-Garp is violated.

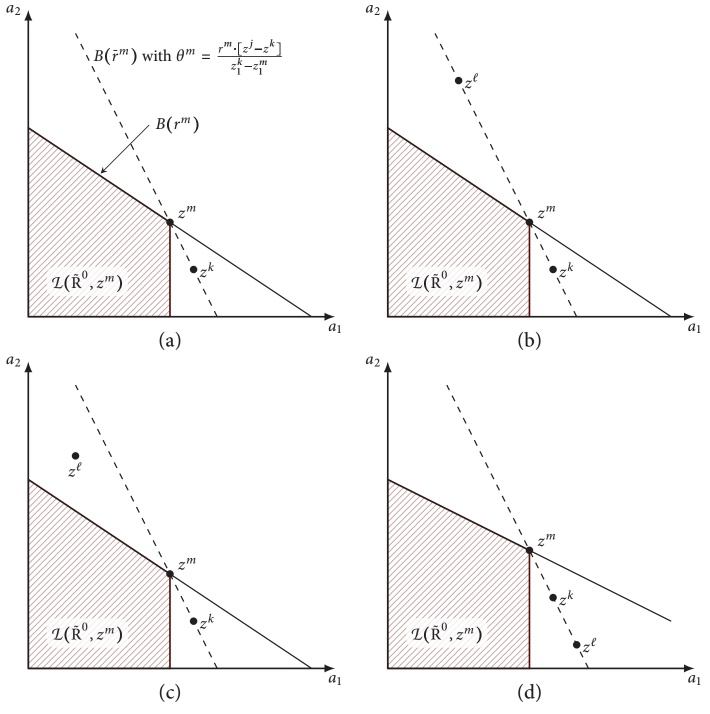


Figure 14: Illustrations for the proof of Theorem 2: Construction of virtual price vectors.

Therefore, we can construct a virtual price vector $\bar{r}^i$ for all z^i such that $\{(z^i, \bar{r}^i)\}_{i=1}^N$ satisfies Garp. As these price vectors are constructed such that $\mathcal{L}(\tilde{\mathrm{R}}^0, z^i) \subseteq \{a \in \mathbb{A} : \bar{r}^i a \geq \bar{r}^i z^i\}$, any utility function which rationalises $\{(z^i, \bar{r}^i)\}_{i=1}^N$ also Ag1-rationalises the original data Ω.

The proof for the existence of a Ag2-rationalising justice function works analogously and we omit it.

What remains to be shown is that the utility and justice functions, u and v, are such that for all $a, a' \in \mathbb{A}$ with $a \succ_{\mathrm{S}} a'$, $u(a) \leq u(a')$ implies $v(a) \leq v(a')$ and $v(a) \geq v(a')$ implies $u(a) \geq u(a')$. Suppose Ag-Garp is satisfied. Then Ag1- and Ag2-rationalising utility and justice functions exist. Suppose u Ag1-rationalises the data. Let $C = CMH(\{z^i : z^i \tilde{\mathrm{R}}_{\mathrm{J}}\, a'\})$. Suppose $a \succ_{\mathrm{S}} a'$ and $u(a) \leq u(a')$, but for every rationalising justice function v, we have $v(a) > v(a')$. Then $a \in \mathrm{int}C$ by Proposition 3 and $a' \in \partial C$ by Lemma 1. Let $\rho \in \mathbb{R}^2_{++}$ be such that $\{a, a'\} \subset \partial B(\rho)$. Then there must exist either a z^i, $z^i \tilde{\mathrm{R}}_{\mathrm{J}}\, a'$, such that $z^i \succ_{\mathrm{S}} a'$, $z^i \in \mathrm{int}B(\rho)$, and $z^i_2 < a_2$. Or there must exist z^i, z^j, with $z^i \tilde{\mathrm{R}}_{\mathrm{J}}\, a'$, $z^j \tilde{\mathrm{R}}_{\mathrm{J}}\, a'$, such that $z^i \in \mathrm{int}B(\rho)$, $z^j \notin \mathrm{int}B(\rho)$, and $z^i \succ_{\mathrm{S}} a'$, $z^j \succ_{\mathrm{S}} a$. In both cases, $z^i \tilde{\mathrm{R}}_{\mathrm{J}} a'$ and $z^i \succ_{\mathrm{S}} a'$, and therefore $z^i \tilde{\mathrm{R}}\, a'$. But $u(a) \leq u(a')$, and concavity and monotonicity of u imply $u(z^i) < u(a')$, a contradiction. See Figure 15 for an illustration.

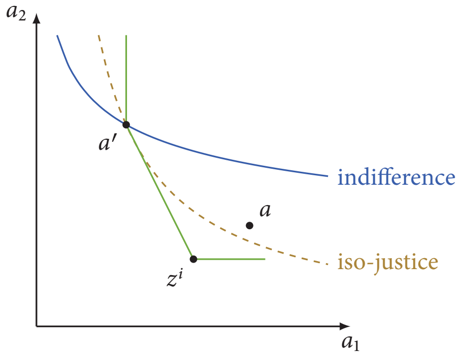


Figure 15: Illustrations for the proof of Theorem 2: If $a \succ_S a'$ and $u(a) \le u(a')$, then $v(a) \le v(a')$ is possible given AgGarp.

Thus, for all $a, a' \in \mathbb{A}$ with $a \succ_S a'$, $u(a) \le u(a')$ implies $v(a) \le v(a')$. But then, with v and u, we have representations of a complete preference and a complete justice relation. Then from Fact 1 it follows that for all $a, a' \in \mathbb{A}$ with $a \succ_S a'$, $v(a) \ge v(a')$ implies $u(a) \ge u(a')$.

B.7 *Proof of Proposition 3*

The first part of the proposition was already proven by Varian (1982, Fact 12) and Knoblauch (1992, Proposition 1). The other parts of the proposition follow from the same arguments as in Knoblauch (1992) and are straightforward to show given the results derived in the proof of Theorem 2. We omit the full proof here. For $L > 2$ (not included in the proposition), we can base the extended revealed preference and justice relations on virtual price vectors described in A.2 and obtain a similar result.

REFERENCES

Afriat, S. N. (1967): "The Construction of Utility Functions From Expenditure Data," *International Economic Review* 8(1):67–77.

Andersson, F. and Lyttkens, C. H. (1999): "Preferences for Equity in Health behind a Veil of Ignorance," *Health Economics* 8(5):369–378.

Andreoni, J. and Miller, J. (2002): "Giving According to GARP: An Experimental Test of the Consistency of Preferences for Altruism," *Econometrica* 70(2):737–753.

Banerjee, S. and Murphy, J. H. (2006): "A Simplified Test for Preference Rationality of Two-Commodity Choice," *Experimental Economics* 9(9):67–75.

Becker, A. and Miller, L. M. (2009): "Promoting Justice by Treating People Unequally: An Experimental Study," *Experimental Economics* 12(4):437–449.

Becker, N., Häger, K., and Heufer, J. (2013): "Revealed Notions of Distributive Justice II: Experimental Analysis," Working paper.

Bolton, G. E. and Ockenfels, A. (2000): "ERC: A Theory of Equity, Reciprocity, and Competition," *American Economic Journal* 90(1):166–193.

Brøndsted, A. (1983): *An Introduction to Convex Polytopes*, *Graduate Texts in Mathematics*, volume 90, Springer-Verlag New York.

Cappelen, A. W., Drange Hole, A., Sørensen, E. Ø., and Tungodden, B. (2007): "The Pluralism of Fairness Ideals: An Experimental Approach," *American Economic Review* 97(3):818–827.

Cappelen, A. W., Konow, J., Sørensen, E. Ø., and Tungodden, B. (2010): "Just Luck: An Experimental Study of Risk Taking and Fairness," Working Paper.

Choi, S., Fisman, R., Gale, D., and Kariv, S. (2007): "Revealing Preferences Graphically: An Old Method Gets a New Tool Kit," *American Economic Review* 97(2):153–158.

Cox, J. C., Friedman, D., and Sadiraj, V. (2008): "Revealed Altruism," *Econometrica* 76(1):31–69.

Dickinson, D. L. and Tiefenthaler, J. (2002): "What is Fair? Experimental Evidence," *Southern Economic Journal* 69(2):414–428.

Diewert, W. E. (1973): "Afriat and Revealed Preference Theory," *Review of Economic Studies* 40(3):419–425.

Engelmann, D. and Strobel, M. (2004): "Inequality Aversion, Efficiency, and Maximin Preferences in Simple Distribution Experiments," *American Economic Review* 94(4):857–869.

——— (2007): "Preferences over Income Distributions: Experimental Evidence," *Public Finance Review* 35(2):285–310.

Fehr, E. and Schmidt, K. M. (1999): "A Theory of Fairness, Competition, and Cooperation," *The Quarterly Journal of Economics* 114(3):817–868.

Fisman, R., Kariv, S., and Markovits, D. (2007): "Individual Preferences for Giving," *American Economic Review* 97(5):1858–1876.

Fleissig, A. R. and Whitney, G. (2011): "A Revealed Preference Test of Rationing," *Economics Letters* 113(3):234–236.

Forges, F. and Minelli, E. (2009): "Afriat's Theorem for General Budget Sets," *Journal of Economic Theory* 144(1):135–145.

Forsythe, R., Horowitz, J. L., and Savin, N. (1994): "Fairness in Simple Bargaining Experiments," *Games and Economic Behavior* 6(3):347–369.

Fostel, A., Scarf, H. E., and Todd, M. J. (2004): "Two New Proofs of Afriat's Theorem," *Economic Theory* 24(1):211–219.

Grünbaum, B. (2003): *Convex Polytopes*, *Graduate Texts in Mathematics*, volume 221, Springer-Verlag New York.

Güth, W., Schmidt, C., and Sutter, M. (2003): "Fairness in the Mail and Opportunism in the Internet: A Newspaper Experiment on Ultimatum Bargaining," *German Economic Review* 4(2):243–265.

Güth, W. and Tietz, R. (1990): "Ultimatum Bargaining Behavior: A Survey and Comparison of Experimental Results," *Journal of Economic Psychology* 11(3):417–449.

Harsanyi, J. C. (1953): "Cardinal Utility in Welfare Economics and in the Theory of Risk-Taking," *Journal of Political Economy* 61(5):434–435.

——— (1955): "Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility," *Journal of Political Economy* 63(4):309–321.

Heufer, J. (2011): "Stochastic Dominance and Nonparametric Comparative Revealed Risk Aversion," *Ruhr Economic Papers* #289, TU Dortmund University, Discussion Paper.

——— (2012a): "Afriat's Theorem, General Budgets, and Concavity: A Geometric Approach," Working paper.

——— (2012b): "Revealed Preference and Nonparametric Analysis: Continuous Extensions and Recoverability," *Ruhr Economic Papers* #324, TU Dortmund University, Discussion Paper.

——— (2013a): "A Geometric Approach to Revealed Preference via Hamiltonian Cycles," *Theory and Decision* forthcoming.

——— (2013b): "Quasiconcave Preferences on the Probability Simplex: A Nonparametric Analysis," *Mathematical Social Sciences* 65(1):21–30.

Karni, E. and Safra, Z. (2002a): "Individual Sense of Justice: A Utility Representation," *Econometrica* 70(1):263–284.

——— (2002b): "Intensity of the Sense of Fairness: Measurement and Behavioral Characterization," *Journal of Economic Theory* 105(2):318–337.

Knoblauch, V. (1992): "A Tight Upper Bound on the Money Metric Utility Function," *American Economic Review* 82(3):660–663.

Krawczyk, M. (2010): "A Glimpse Through the Veil of Ignorance: Equality of Opportunity and Support for Redistribution," *Journal of Public Economics* 94(1-2):131–141.

Matzkin, R. L. (1991): "Axioms of Revealed Preference for Nonlinear Choice Sets," *Econometrica* 59(6):1779–1786.

Nguema, J. F. (2003): "Sense of Impartiality," *Economics Bulletin* 4(31):1–7.

Rawls, J. (1971): *A Theory of Justice*, Cambridge: Harvard University Press.

Rose, H. (1958): "Consistency of Preference: The Two-Commodity Case," *Review of Economic Studies* 25(2):124–125.

Samuelson, P. A. (1974): "Complementarity: An Essay on the 40th Anniversary of the Hicks-Allen Revolution in Demand Theory," *Journal of Economic Literature* 12(4):1255–1289.

Schildberg-Hörisch, H. (2010): "Is the Veil of Ignorance only a Concept about Risk? An Experiment," *Journal of Public Economics* 94(11-12):1062–1066.

Tan, J. H. W. and Bolle, F. (2006): "On the Relative Strength of Altruism and Fairness," *Theory and Decision* 60:35–67.

Traub, S., Seidl, C., and Schmidt, U. (2009): "An Experimental Study on Individual Choice, Social Welfare, and Social Preferences," *European Economic Review* 53(4):385–400.

Varian, H. R. (1982): "The Nonparametric Approach to Demand Analysis," *Econometrica* 50(4):945–972.

——— (1983): "Non-parametric Tests of Consumer Behaviour," *Review of Economic Studies* 50(1):99–110.

Yaari, M. E. and Bar-Hillel, M. (1984): "On Dividing Justly," *Social Choice and Welfare* 1(1):1–24.

Yatchew, A. J. (1985): "A Note on Non-Parametric Tests of Consumer Behaviour," *Economics Letters* 18(1):45–48.