REVEALED SOCIAL PREFERENCES

ARTHUR DOLGOPOLOV

George Mason University

MIKHAIL FREER

University of Leuven (KU Leuven)

Abstract. We apply a revealed preference approach to test the consistency of observed behavior with theories of social preferences. In particular, we provide revealed preference criteria for the observed set of choices generated by inequality averse preferences and by other-regarding preferences that exhibit increasing benevolence. We further apply these tests to some experimental data on dictator games. Finally, we show how to apply constructed tests to other games commonly used to study social preferences, including ultimatum, investment and carrot-stick games.

1 Introduction

The revealed preference approach, pioneered by Samuelson (1938), departs from the fact that, although we cannot observe complete preference relation profiles of agents, we can observe their choices over some budget sets. This approach was widely used to develop and apply tests for theories of individual behavior. Starting from Richter (1966) and Afriat (1973) the approach has been applied to construct tests of individual and collective decision making (see Chambers and Echenique, 2016, for comprehensive overview of the results).

Date: February 3, 2018.

In this paper we use a revealed preference approach to construct tests for consistency with different theories of social preferences. We start with a standard choice environment over linear budgets, in which a player decides how to allocate money between herself and another player. Then we provide an empirical applications to illustrate the methodology using experimental data. Finally, we generalize tests for various other games that are used to study social preferences, including ultimatum, trust and carrot-stick games.

We consider three theories: other-regarding preferences, inequality aversion, and increasing benevolence. Other-regarding preference assumes that an agent derives some utility (or disutility) from a payoff received by her counter-part. This assumption has testable implications and is supported by the data (see e.g. Andreoni, 1990; Andreoni and Miller, 2002; Charness and Rabin, 2002; Fisman et al., 2007; Porter and Adams, 2016). Other-regarding preferences have been shown to have implications for efficiency when incorporated into models of general equilibrium (Dufwenberg et al., 2011) and bilateral exchange (Benjamin, 2015). This model of social preferences has also been used with evolutionary games (Szabo and Szolnoki, 2012) and studies of social distance (Hoffman et al., 1996).

A more restrictive assumption is that agents do not like inequality; that is, they encounter some disutility if payoffs are unbalanced. This assumption is usually referred to as *inequality aversion*, and it also explains some experimental data (see e.g. Fehr and Schmidt, 1999; Bolton and Ockenfels, 2000, 2006; Fehr et al., 2006). Inequality aversion is a popular instrument in political science, and is often applied to studies of income redistribution (Fong, 2001; Tyran and Sausgruber, 2006; Höchtl et al., 2012; Durante et al., 2014; Agranov and Palfrey, 2015).

Finally, Cox et al. (2008) used an alternative assumption called *increasing benevolence*. The idea is that a player's willingness to pay for an additional dollar received by another player is an increasing function of the player's own payoff. This is also a natural assumption. It has been used to guarantee efficiency in bilateral exchange (Benjamin, 2015). In addition, (Benjamin, 2015) further showed how this assumption is better suited to explain some common phenomena, e.g., the rotten-kid theorem (Becker, 1974), than monotonicity ("altruistic preferences").

Revealed preference theory has been applied to study social preferences starting with Andreoni and Miller (2002). In particular, Andreoni and Miller (2002); Fisman et al. (2007); Deb et al. (2014); Porter and Adams (2016) all applied revealed preference theory to test consistency of observed choices with the hypothesis of other-regarding preferences.[1] Castillo et al. (2017) showed that both proposers and responders in an ultimatum game are consistent with having other-regarding preferences. Cox et al. (2008) used a revealed preference approach to construct a nonparametric framework for classification of people in terms of their level of altruism and out-of-sample predictions based on observed choices. However, the latter two papers make assumptions that are different from basic assumptions used in Andreoni and Miller (2002). Castillo et al. (2017) uses a weaker assumption; nonetheless, in the context of a dictator game it would have exactly the same empirical content as the one used by Andreoni and Miller (2002).[2] Existing literature concentrates exclusively on the characterization of other-regarding or "altruistic" preferences.

The remainder of this paper is organized as follows. Section 2 presents the general set up and the revealed preference tests. Section 3 provides empirical illustrations. Section 4 demonstrates the extension of constructed revealed preference tests for other games. All proofs are collected in Appendix A.

2 Theoretical Framework

A dictator game is structured as follows: A player decides how to allocate a given amount of money between herself and the recipient, and the chosen allocation is implemented. Hence, this game can be

[1]Becker et al. (2015) consider revealed preferences for notions of distributive justice, which are close to the social preferences, but have to also be symmetric over agents.

[2]The difference between these theories can be tested using responders' choices.

perceived as a decision problem, in which one player chooses a two-dimensional vector allocation: payoff to self and payoff to another player.

Let $X \subseteq \mathbb{R}^2_+$ be the **set of alternatives**. For every $x \in X$ let $x = (x_s, x_o)$, where x_s is the payoff to self and x_o is the payoff to another player. Let $p \in R^2_{++}$ be a price vector. We normalize the income to one at every point, so that the budget set can be defined as follows, $B(p) = \{x \in X : px \leq 1\}$. Let $E = (x^t, p^t)_{t=1}^T$ be an **experiment**, which consists of T choices (x^t) at a given price vector (p^t). Moreover, we assume that chosen points x^t are such that $p^t x^t = 1$.[3] A function $u(x) : X \to \mathbb{R}$ **rationalizes** the consumption experiment E if $\forall y \in B(p^t)$ $u(x^t) \geq u(y)$ for every $t \in \{1, \ldots, T\}$.

2.1 Other-Regarding Preferences. In the case of other-regarding preferences, the dictator cares about her own payoff and the payoff of the recipient. We do not make an explicit assumption of whether the dictator derives utility or disutility from x_o.

Definition 1. *An experiment $E = (x^t, p^t)_{t=1}^T$ is* **rationalizable with other-regarding preferences** *if there is a continuous utility function $u(x_s, x_o)$ monotone in x_s that rationalizes E.*

This would include the utility function that is monotone in both payoffs as a special case.

Definition 2. *An experiment $E = (x^t, p^t)_{t=1}^T$ is consistent with* **Generalized Axiom of Revealed Preference (GARP)** *if and only if we have $p^{t_1} x^{t_n} \leq p^{t_1} x^{t_1}$ for all sequences $x^{t_1}, \ldots, x^{t_n}$, such that $p^{t_{j+1}} x^{t_j} \leq p^{t_{j+1}} x^{t_{j+1}}$, $j \in \{i, \ldots, n-1\}$.*

Proposition 1 (Afriat (1967); Diewert (1973); Varian (1982))**.** *An experiment is rationalizable with other-regarding preferences if and only if it satisfies GARP.*

[3]This is rather a technical assumption dictated by non-satiation of preferences which would be required further. All the further reasoning can be done without this assumption using more complicated notation.

This is accurate, since for GARP to be satisfied the utility function should be at least locally non-satiated, and GARP is sufficient for the existence of a utility function that is monotone in all goods.

2.2 Inequality Aversion A more specific theory is inequality aversion. In this case, the subject cares not solely about her payoff but also about the difference between her and the other player's payoff. Therefore, in order to continue, we need to define *inequality measure.*

Definition 3. *A continuous function $f(x_s, x_o)$ is an inequality measure if:*

- $f(x_s, x_o) \geq 0$, *for every* x_s, x_o;
- $f(x_s, x_o) = 0$ *if and only if* $x_s = x_o$;
- *if* $x_s > x_o$, *then* $f(x_s, x_o)$ *is decreasing in* x_o *and increasing in* x_s;
- *if* $x_s < x_o$, *then* $f(x_s, x_o)$ *is increasing in* x_o *and decreasing in* x_s;
- $f(\max\{x_s, x_o\}, \min\{x_s, x_o\}) \leq f(\min\{x_s, x_o\}, \max\{x_s, x_o\})$.

Definition 3 generalizes measures of inequality frequently used in the literature.[4] It can be easily seen that the following functions are specific cases of Definition 3:

- Inequality in differences $\delta(x_s, x_o) = |x_s - x_o|$ (e.g. Fehr and Schmidt, 1999; Tyran and Sausgruber, 2006; Agranov and Palfrey, 2015):
$$f(x_s, x_o) = \begin{cases} |x_s - x_o| & \text{if } x_s \geq x_o \\ \beta|x_s - x_o| & \text{if } x_o > x_s \end{cases}$$
where $\beta \leq 1$.
- Inequality in shares (e.g., Bolton and Ockenfels, 2000):
$$f(x_s, x_o) = \begin{cases} |\frac{x_s}{x_s + x_o} - \frac{1}{2}| & \text{if } x_s \geq x_o \\ \beta|\frac{x_s}{x_s + x_o} - \frac{1}{2}| & \text{if } x_o > x_s \end{cases}$$

[4] Closely related axiomatization has been used by Fehr et al. (1998) for the gift exchange game in order to show how preferences for the fair outcome can generate the reciprocal choices.

where $\beta \leq 1$.

– Gini Index $f(x_s, x_o) = \frac{|x_s - x_o|}{2(x_s + x_o)}$ (e.g. Durante et al., 2014)

If an agent is inequality averse, then the dictator maximizes a utility function $u(x_s, f(x_s, x_o))$ that is increasing in x_s and decreasing in $f(x_s, x_o)$.

Definition 4. *An experiment is* **rationalizable with inequality averse preferences** *if there are continuous functions $f(x_s, x_o)$ and $u(x_s, f(x_s, x_o))$ such that $f(x_s, x_o)$ is an inequality measure and $u(x_s, f(x_s, x_o))$ is a utility function that is increasing in x_s and decreasing in $f(x_s, x_o)$.*

To obtain rationalizability with inequality averse preferences, we only need to add the condition that every subject must choose to allocate to herself at least as much as to the other player. Necessity of this condition is clear. The reason is that if $x_s < x_o$, then the player could obtain greater utility by increasing x_s at the costs of x_o. More interesting is that this condition together with GARP suffices for rationalizability with inequality averse preferences.

Proposition 2. *An experiment is rationalizable with inequality averse preferences if and only if it satisfies GARP and $x_s^t \geq x_o^t$ for every $t \in \{1, \dots, T\}$.*

Remark 1. *An experiment is rationalizable with inequality averse preferences, if and only if it is rationalizable with any inequality measure.*

The above remark follows directly from the proof of Proposition 2 (see Appendix A). Recall that for rationalization we require that there is a measure of inequality, such that a subject acts as if she is maximizing some utility function increasing in her own payoff and decreasing in the measure of inequality. Remark 1 shows that it does not matter which measure of inequality we choose. If a subject is rationalizable with a particular measure, it would be implied that she is rationalizable with any other. Hence, the test is a comprehensive test of inequality version hypothesis, which does not need require a specific inequality measure.

2.3 Increasing Benevolence The increasing benevolence assumption is another theory nested within the other-regarding preferences. Increasing benevolence means that a subject's willingness to pay for an additional dollar given to the other player is increasing in own payoff x_s.[5]

Denote the demand function for x_o by $D_o(p_s, p_o)$ and the demand for x_s by $D_s(p_s, p_o)$. Since we operate in a two-dimensional case, one demand can be immediately derived from another $D_s(p_s, p_o) = \frac{1-p_o D_o(p_s,p_o)}{p_s}$.

Definition 5. *An experiment is* ***rationalizable with increasing benevolence preferences*** *if there is a rational demand function $D_o(p_s, p_o)$ such that*

- $D_o(p_s^t, p_o^t) = x_o^t$, *and*
- $\frac{p_s}{p_o} \geq \frac{p_s'}{p_o'}$ *and* $\frac{1-p_o' D_o(p_s,p_o)}{p_s'} \geq D_s(p_s, p_o)$ *implies* $D_o(p_s, p_o) \leq D_o(p_s', p_o')$.

This condition is equivalent to normality. A good x_o is said to be normal if higher income implies that more of it is consumed.

Figure 1 illustrates why the definition of increasing benevolence in this case is equivalent to normality. Assume that x^1 is a point chosen from the budget defined by p^1; then, the new budget is such that $\frac{p_s}{p_o} \geq \frac{p_s'}{p_o'}$ (x_s is relatively more expensive in the new budget) and the old bundle is attainable. The dashed line shows the parallel downward shift of the budget defined by p_2. Hence, the choice from the dashed budget should be with at least as much x_o as from p^1 (due to the substitution effect). Furthermore, since dashed and p^2 budgets are different only in income, then normality would guarantee that the choice from p^2 would be "above" the x^1.

[5]This can be defined more formally using the marginal rate of substitution – $WTP = 1/MRS = \frac{u_{x_o}}{u_{x_s}}$ is increasing in x_s. We use the reduced form definition of this, which is necessary but not sufficient. However, it is sufficient to guarantee the empirical implications described by Cox et al. (2008). Moreover, if we define MRS via the ratio of the inverse demand functions (to guarantee the existence of MRS), some sufficiency result can be inferred. Although, one can easily check that if we, for instance, assume that x_s and x_o are substitutes, then the demand conditions would be sufficient for the MRS version.

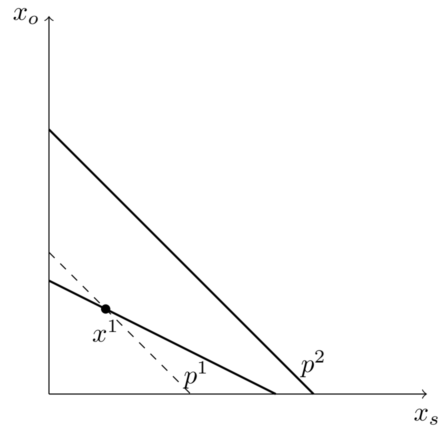


FIGURE 1. Increasing Benevolence and Normality

Definition 6. *An experiment $E = (x^t, p^t)_{t=1}^T$ is consistent with* **Normality Axiom of Revealed Preference (NARP)** *if and only if for all observations $t, v \in \{1, \dots, T\}$ if $p_o^t/p_s^t \leq p_o^v/p_s^v$ and $x_s^v \leq \frac{1-p_o^t x_o^v}{p_s^t}$, then $x_o^v \leq x_o^t$.*

Due to the equivalence between increasing benevolence and normality of demand in x_o we can directly employ the result from Cherchye et al. (2018) to test for increasing benevolence.[6]

Proposition 3 (Cherchye et al. (2018))**.** *An experiment is rationalizable with increasing benevolence preferences if and only if it satisfies NARP.*

2.4 Distance from rationality All tests presented so far are binary: they give the answer as to whether the experiment is consistent with a theory or not. However, it has been consistently shown that people make mistakes in decision making. Therefore, it is necessary to account for those mistakes. For this purpose, we use the **Houtman-Maks index (HMI)**.[7] HMI is the maximum fraction of data that can

[6]If a reader is not convinced by the equivalence argument above, please see p.375 in Cherchye et al. (2018) where it is directly proven that the increasing benevolence property is satisfied.

[7]We use the HMI due to the fact that it is the only index which can be applied to the test of Inequality Aversion. Critical Cost Efficiency Index introduced by

be rationalized by a given theory. That is, if in a total of T observations, the maximum subset which is consistent with GARP is τ, then $HMI = \tau/T$ for this experiment. For the technical details regarding the implementation of the HMI index, see Appendix B.

3 Empirical Illustration

We use data from two studies of dictator games. In both, subjects repeatedly played a dictator game with different relative prices and endowments as in Andreoni and Miller (2002). In every period subjects were asked to allocate tokens between themselves and another person, choosing a point on a linear budget $p_s^t x_s + p_o^t x_o \leq m^t$.

The first study by Fisman et al. (2007) contains results of experiments with 76 undergraduates from UC Berkeley.[8] In this study, every subject faced 50 different budgets with randomly determined prices. The second study by Porter and Adams (2016) contains results of experiments with 89 subjects recruited from the general population from the southeast region of the UK. In this study every subject faced 11 different budgets with predetermined prices.

When applied to data, notions of rationality prove to be very strict at least for the first dataset: no more than 16% of subjects can be rationalized with other-regarding preferences and no more than 11% if we consider nested theories. Therefore, it makes sense to relax the notion of rationality and allow for some probability that people make mistakes. For this purpose we report results for the HMI level of .9, although results are fairly robust to other levels of HMI (see Appendix D). The HMI level of .9 is equivalent to allowing for deviations from rationality in no more than 10% of budgets.

Afriat (1973) would not adequately work in the context of inequality aversion (for more details see Appendix). The money pump index introduced by Echenique et al. (2011) is defined for GARP only. The swaps index proposed by Apesteguia and Ballester (2015) can be applied only in the context of finite choice sets.

[8]Experiment contains two other treatments which we do not consider in our analysis. One of the treatments uses step-shaped budgets and another is a dictator game with two recipients.

In addition, we want to control for false positives, which gives the probability that a random decision making rule would look consistent with the test.

For this purpose we use the following procedures. First, we compute the Bronars (1987) test by generating 1000 pseudo subjects who make decisions uniformly distributed along the budget line. The second is the bootstrap index introduced by Andreoni and Miller (2002) and Harbaugh et al. (2001). This measure controls for possible behavioral rules which can cause false positive results even if people would take decisions at random. To compute the bootstrap power of the test, we calculate the empirical distribution of the shares of income spent on each commodity – in our case the subject's payoff and the other's payoff – and simulate the pseudo subjects who make their choices at random, but distributed according to the e.d.f.

To compare pass rates taking into account the power, we use the **predictive success index (PSI)** introduced by Selten (1991).[9] The predictive success index is defined as the difference between the share of people that satisfies an axiom at the given level of HMI and the probability that random choices will satisfy the axiom at the same level of HMI. This index ranges between -1 and 1, with -1 meaning no subject passes while all random subjects pass and 1 meaning every subject passes while none of the random subjects do.

Table 1 collects results for both datasets. The first column lists the theory for which analysis is performed. The second column reports pass rates. The third and fourth columns present the power computations according to Bronars and the bootstrap methods. By power, we mean the fraction of random subjects who fail to perform consistently with the test. Last two columns present the predictive success index using Bronars and bootstrap powers. Recall that both inequality aversion and increasing benevolence are nested within the other-regarding preferences model. That is, increasing benevolence and inequality aversion

[9]Methodology of using predictive success index in the revealed preference context was introduced by Beatty and Crawford (2011). Statistical interpretation of the index which allows us to construct confidence intervals was proposed by Demuynck (2015).

Fisman et al. (2007) data					
		Power of Test		PSI	
Theory	Pass Rate	Bronars	Bootstrap	Bronars	Bootstrap
Other-regarding	58 (76.32%)	100.00%	99.63%	0.76	0.76
95% conf. interval	(65.18% - 85.32%)	(99.99% - 100.00%)	(99.58% - 99.67%)	(0.67 - 0.86)	(0.66 - 0.86)
Inequality Aversion	32 (55.17%)	100.00%	92.58%	0.55	0.48
95% conf. interval	(41.54% - 68.26%)	(100.00% - 100.00%)	(92.39% - 92.76%)	(0.42 - 0.68)	(0.35 - 0.61)
Increasing Benevolence	12 (20.69%)	100.00%	100.00%	0.21	0.21
95% conf. interval	(11.17% - 33.35%)	(100.00% - 100.00%)	(100.00% - 100.00%)	(0.10 - 0.31)	(0.10 - 0.31)

Porter and Adams (2016) data					
		Power of Test		PSI	
Theory	Pass Rate	Bronars	Bootstrap	Bronars	Bootstrap
Other-regarding	81 (91.01%)	68.72%	80.01%	0.60	0.71
95% conf. interval	(83.05% - 96.04%)	(68.41% - 69.02%)	(79.75% - 80.28%)	(0.54 - 0.66)	(0.65 - 0.77)
Inequality Aversion	51 (62.96%)	98.62%	90.50%	0.62	0.53
95% conf. interval	(51.51% - 73.44%)	(98.55% - 98.70%)	(90.31% - 90.69%)	(0.51 - 0.72)	(0.43 - 0.64)
Increasing Benevolence	69 (85.19%)	85.68%	65.67%	0.71	0.51
95% conf. interval	(75.55% - 92.10%)	(85.44% - 85.90%)	(65.35% - 65.98%)	(0.63 - 0.79)	(0.43 - 0.59)

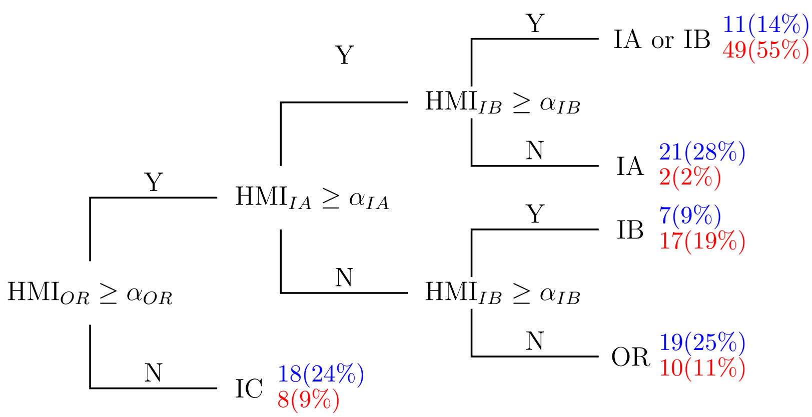
Table 1. Results with rationality presumed at HMI=.9

make stronger assumptions about underlying social preferences. For this reason, we report a nested theory analysis for the two stronger theories; that is, results are conditional on satisfying the other-regarding preference assumption. We estimate the probability of success using the subexperiment, using only subjects who are consistent with other-regarding preferences hypothesis at the given level of HMI.

First, we observe that both inequality aversion and increasing benevolence preferences are significantly restrictive. That is, more than 5% of the population who are consistent with other-regarding preferences, cannot be rationalized by either inequality averse or increasing benevolence preferences. In terms of comparing which of the nested theories is better, we find results that differ between datasets.

Inequality aversion appears to describe data better in the first dataset; at the same time, increasing benevolence performs at least as well as inequality aversion in Porter and Adams (2016) data, while the difference is not statistically significant. This inconsistency between the two datasets provides evidence in line with Fehr et al. (2006), who showed that social preferences may depend on the demographical characteristics of the population. Controlling for the power of the test does not appear to affect these results.

3.1 Mixed Types Analysis Note that none of the nested theories can explain the behavior of the entire sample. At the same time, we see that all of them perform well. In addition, different theories have quite different empirical implications, and we observe that the correlations between pass rates for inequality aversion and increasing benevolence are quite low: .22 (with a confidence interval of $[-.01, .42]$) for Fisman et al. (2007) data and .51 (with a confidence interval of $[.34, .65]$) for Porter and Adams (2016) data.[10] This shows that there is a non-trivial probability that different subjects can be consistent with different notions of rationality. Therefore, further we perform a mixed type analysis.



Top numbers are for Fisman et al. (2007) data,
bottom numbers are for Porter and Adams (2016) data.

Figure 2. Classification Tree

[10]Given that the tests are binary, appropriate statistic is ϕ-coefficient. It is a version of correlation coefficient for two binary variables. Both logit and probit regression coefficients are insignificant for Fisman et al. (2007) data: logit regression coefficient is 1.2 (with 95% confidence interval of $[-0.06, 2.61]$) and probit regression coefficient is 0.75 (with 95% confidence interval of $[-0.04, 1.57]$). That is we can not reject that two nested theories are unrelated for Fisman et al. (2007) data. For Porter and Adams (2016) data the relationship is much stronger: logit coefficient is 3.09 (with 95% confidence interval of $[1.74, 4.99]$), probit coefficient is 1.84 (with 95% confidence interval of $[1.07, 2.72]$). That is, odds of the subject being consistent with inequality aversion are at least $e^{1.74} = 5.7$ times higher if she is also consistent with increasing benevolence than if she is not.

We use the classification presented in Figure 2. We sort subjects according to three sequential binary classification steps. First, if a subject is not consistent with other-regarding preferences at threshold α_{OR}, she is classified as inconsistent with other-regarding preferences (IC). Then, we compare whether she is consistent with inequality aversion or increasing benevolence with thresholds α_{IA} and α_{IB} respectively. If the subject is not consistent with either, she is classified as other-regarding (OR). If the subject is consistent with both, she is classified as inequality averse or increasing benevolent (IA or IB). If the subject is consistent with only one theory, she is classified as inequality averse (IA) or increasing benevolent (IB).

It is still necessary to determine the thresholds for the classification tree. In order to do this, we modify the unsupervised machine learning methodology from Liu et al. (2000). The idea behind this is to maximize the information gain from adding a particular cluster. We base this measure on HMI, but the approach is general (for more detailed explanation see Appendix C). The thresholds obtained are as follows: $\alpha_{OR} = 45/50$; $\alpha_{IA} = \alpha_{IB} = 41/50$ for Fisman et al. (2007) data and $\alpha_{OR} = \alpha_{IA} = \alpha_{IB} = 10/11$ for Porter and Adams (2016) data.

In the first dataset, we can see that a large set of subjects can be described by inequality aversion. This provides additional evidence that inequality aversion organizes the data better than increasing benevolence, but at the individual level. In the second experiment, the distinction is unclear, as both nested theories describe behavior well; however, as a caveat here, note that power of test for increasing benevolence is significantly lower than the one for inequality aversion.[11] The test is particularly conservative, and favors the theory which guarantees the least probability of returning the false positive outcome. As a final

[11]In Porter and Adams (2016) experiment, three participants gave more to their counterparts than they kept for themselves in all 11 budgets. Such altruistic behavior is maximally inconsistent with inequality aversion, implying an HMI of zero (theoretically HMI for other theories starts with 1). We exclude these three subjects from the mixed type analysis. We show that this does not significantly affect results in Appendix C.

remark, there is a share of the population (11-25%) that cannot be explained by either inequality aversion or increasing benevolence, but still has other-regarding preferences. This also provides additional evidence for both assumptions being significantly restrictive.[12]

Several previous papers (see e.g. Andreoni and Miller, 2002; Porter and Adams, 2016) attempted to classify people into different types according to their giving behavior. However, they used a parametric specification to characterize particular types of preferences, while the method we offer is completely non-parametric. Moreover, since we use the HMI to classify people, this allows us to use a subset of data according to which a subject is consistent to construct out-of-sample predictions.

4 Discussion

Next, we discuss the context dependence of social preferences and show how to apply our theoretical results toward ultimatum, investment and carrot-stick games. In addition, we demonstrate how to implement the classification of people according to their level of altruism, which was proposed by Cox et al. (2008) under partial observability and how to identify the subjective notion of fair outcome.

4.1 Revealed Social Preferences Beyond Dictator Games The revealed preference approach can be applied to test theories of social preferences in other games. It is particularly important, given that social preferences are context dependent (see e.g., Engellman and Strobel (2004)) and different motives would cause different theories of social preferences to perform best.[13] Our empirical findings confirm that this is true even if we relax the parametric assumptions. Cox et al. (2008) found supportive evidence for increasing benevolence

[12]The pass rates guarantee us that at least 5% of the population cannot be explained by assumptions other than other-regarding preferences. We borrow the 5% criterion from Polisson et al. (2017) who used it to test that the theory is significantly restrictive. This criterion is tested using the probability of success estimated according to the Clopper-Pearson procedure.

[13]Context dependence is closely related to the fact that different motivations are triggered in different games.

preferences in their data, while we find the opposite. However, the results are not contradictory, as Cox et al. (2008) analyzed games with reciprocity motives. This could may imply better performance of increasing benevolence. In particular, they used data from Cox (2004), in which the budget sets are determined according to the actions of another player. In our experiment, however, budgets are determined independently of actions of another player. Further, we provide revealed preference tests of the theories of social preferences for other games. Following Cox et al. (2008), we consider second-movers in two-stage dynamic games.[14]

4.1.1 Ultimatum Game. First, a proposer is given an endowment m^t and asked to allocate it between herself and a responder, given that $p_p^t x_p + p_r^t x_r = m^t$, where x_p denotes the proposer's earnings and x_r denotes the responder's earnings. In the second stage, the responder decides whether to accept or reject the allocation. If the allocation is accepted, it is implemented; otherwise, both players get zero.

The case of other-regarding preferences is already considered in Castillo et al. (2017). Increasing benevolence is not well-defined for the binary choices, since it requires budgets to be linear.

Further, we state the test for inequality aversion.

The experiment, on the side of the responder, is a sequence of binary decisions between proposed allocations and zero payoffs. Figure 3 shows the decision problem, as well as acceptance (A^t) and rejection (R^t) regions in $(f(x_r, x_p), x_r)$ coordinates. To be more precise, we can define the acceptance and rejection regions as follows:

$$A^t = \{(f(x_r, x_p), x_r) : x_r \geq x_r^t \text{ and } f(x_r, x_p) \leq f(x_r^t, x_p^t)\}$$

and

$$R^t = \{(f(x_r, x_p), x_r) : x_r \leq x_r^t \text{ and } f(x_r, x_p) \geq f(x_r^t, x_p^t)\}$$

If point $(f(x_r^t, x_p^t), x_r^t)$ was accepted, then every allocation from its acceptance area should be accepted as well. Otherwise, every allocation

[14]For the behavior of first-movers, beliefs play a significant role. This significantly complicates the framework and limits the empirical content of the theory. For the revealed preference analysis of the first-movers, see Castillo et al. (2017).

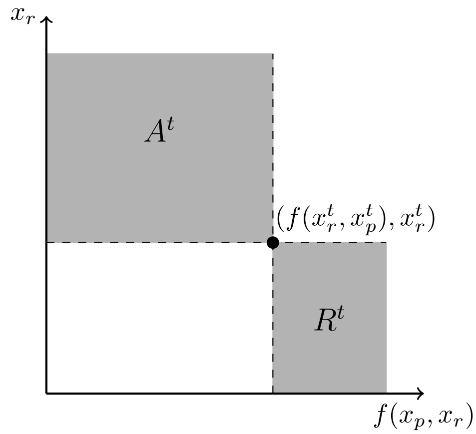


FIGURE 3. Acceptance and Rejection area in Ultimatum Game

from its rejection area should be rejected. Denote by A_x the set of all allocations being accepted and by R_x the set of all allocations being rejected.

Corollary 1. *Let $f(x_f, x_s)$ be a measure of inequality. An experiment is rationalizable with inequality averse preferences if and only if*

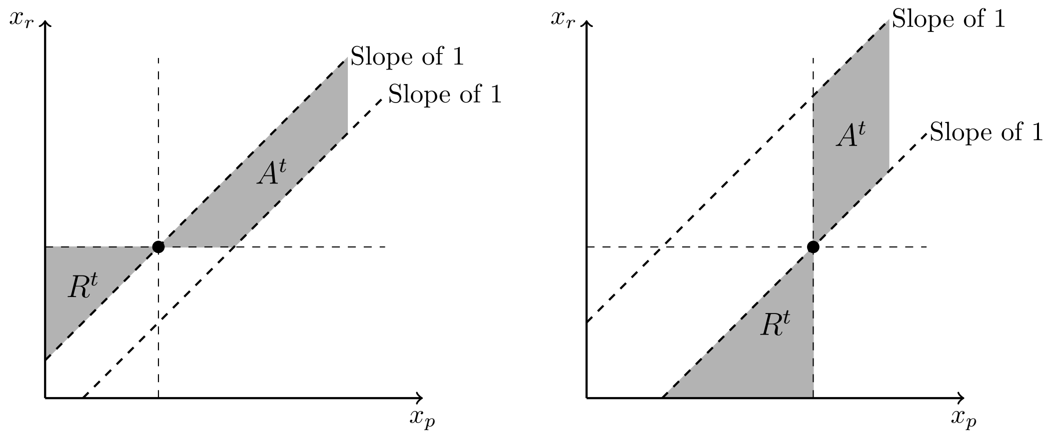
$$R_x \cap \left(\bigcup_{x^t \in A_x} A^t \right) = \emptyset$$

and

$$A_x \cap \left(\bigcup_{x^t \in R_x} R^t \right) = \emptyset$$

The *decision space* is defined in terms of proposer and responder payoffs. Hence, the tests of the theory would be more illustrative in the decision space. Interestingly, this plot would translate differently for different measures of inequality. This allows one to compare the performance of different measures of inequality. We consider examples of inequality aversion in differences ($f(x_r, x_p) = |x_r - x_p|$) and inequality aversion in shares ($f(x_r, x_p) = \left|\frac{x_r}{x_r + x_p} - \frac{1}{2}\right|$).

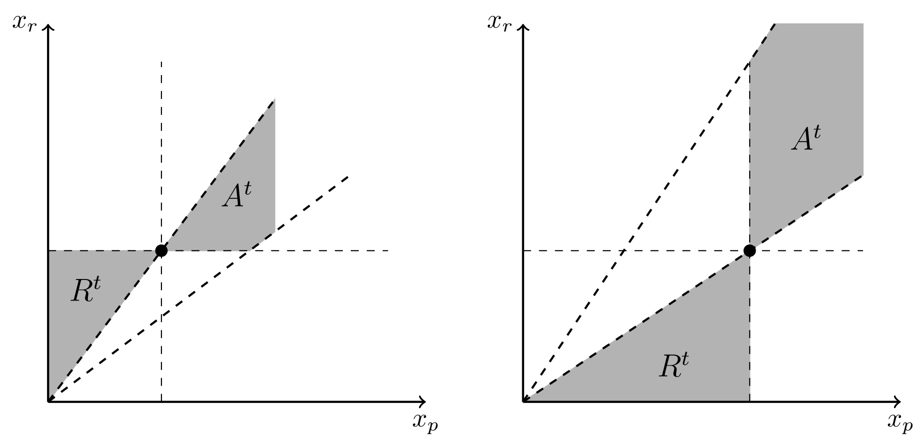
Figure 4 shows the acceptance and rejection regions for the inequality aversion in difference. In this case, every point that lies on the line of slope one which goes through x has the same inequality level as x. The acceptance and rejection regions differ if different players get higher



(a) $x_r > x_p$ (b) $x_r \leq x_p$

FIGURE 4. Acceptance and Rejection regions in Ultimatum Games for Inequality Aversion in differences

payoffs; the reason is that the direction in which inequality increases differs. Rejection regions (triangles) give the responder lower payoff and increase inequality. Acceptance regions are represented by the stripes. This is caused by the fact that the increased payoff should deliver the measure of inequality, which does not exceed the original level.



(a) $x_r > x_p$ (b) $x_r \leq x_p$

FIGURE 5. Acceptance and Rejection regions in Ultimatum Games for Inequality Aversion in shares

Figure 5 shows acceptance and rejection regions for inequality aversion in shares. In this case, every point that lies on the line that goes through zero and x is has the same inequality level as x. Comparing rejection and acceptance regions from Figures 4 and 5, we can see that they are different. Therefore, these measures have different testable implications.

4.1.2 Investment Game. Players start with an endowment of I. The first-mover sends an amount $s \in [0, I]$ to the second-mover who receives ks. Then the second-mover returns an amount of $r \in [0, ks]$, and the first-mover receives pr. Then, the final payoffs are $x_f = I - s + pr$ and $x_s = I + ks - r$. Hence, considering family of investment games with different p would generate sufficient price variation to apply revealed preference tests.

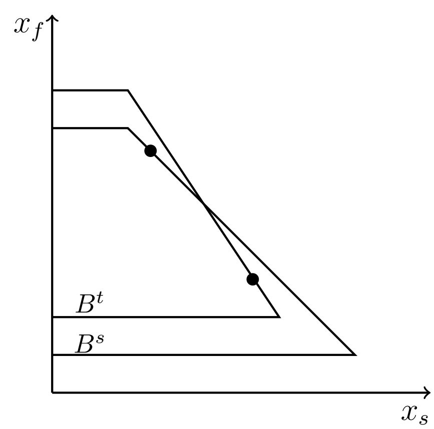


FIGURE 6. Second-mover's budget set in the investment game.

Figure 6 presents the budget set of the second-mover. Choices on the horizontal segments are not feasible. Therefore, Proposition 1 can be directly used to test for other-regarding preferences. Figure 6 shows possible violation of GARP, hence, there is empirical content beyond other-regarding preferences. Moreover, if all choices are such that $(x_f, x_s) \in B^s$ and $x_s \geq x^t$ for every $t, s \in \{1, \dots, T\}$, then Proposition 3 can be applied to test for increasing benevolence. If we design the parameters such that $x_f = x_s$ outcome is available at every budget, then Proposition 2 can be applied to test for inequality aversion.

Therefore, in the investment game (as in dictator), rationalization with one measure of inequality implies rationalization with every measure of inequality.

4.1.3 Carrot-Stick Game. Players have an endowment of I. The first-mover chooses the amount to be sent $s \in [0, I]$. Then, the second-mover can "return" the amount $r \in [-s, s]$ and the first-mover receives pr. This results in payoffs $x_f = I - s + pr$ and $x_s = I + s - |r|$.

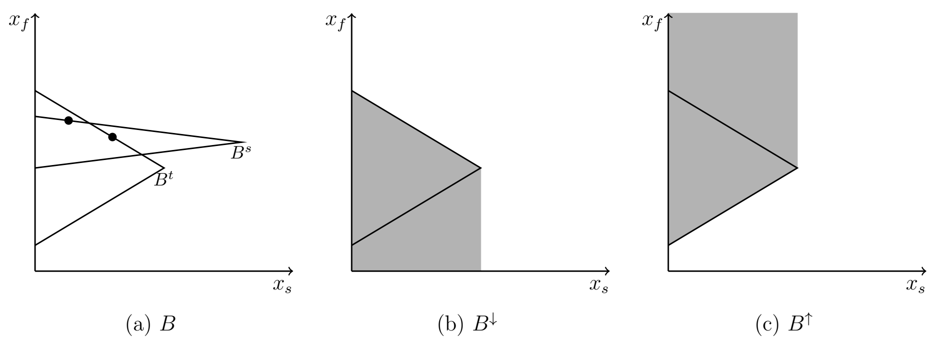


FIGURE 7. Second-mover's budget set in the carrot-stick game.

Figure 7(a) presents the budget sets that the second-mover faces. In this game budgets have both "upper" and "lower" borders. Therefore, it is possible to test whether a player has "altruistic" or "envious" preferences. Preferences are said to be *altruistic* if utility is increasing in both x_s and x_f. Preferences are said to be *envious* if utility is increasing in x_s and decreasing in x_f. Denote by B^t the budget (as on Figure 7) based on which the subject makes a choice. Denote by $B^{\downarrow}(^{\uparrow}) = \{(x'_s, x'_f) : (x_s, x_f) \in B \text{ such that } x_s \geq x'_s \text{ and } x'_f \geq (\leq) x_f\}$. Denote by $B^{\downarrow\downarrow}(^{\uparrow\uparrow}) = \{(x'_s, x'_f) : (x_s, x_f) \in B \text{ such that } x_s \geq x'_s \text{ and } x'_f \geq (\leq) x_f\}$ with at least one inequality being strict. Figures 7(b) and 7(c) illustrate the construction of $B^{\downarrow}$ and $B^{\uparrow}$ respectively. The shaded areas show the part of the space added by taking the closure of the budget.

Corollary 2. *An experiment $E = (x^t, B^t)_{t=1}^T$ is rationalizable with altruistic [envious] preferences if and only if there is no sequence $x^{t_1}, \ldots, x^{t_n}$, such that $x^{t_j} \in (B^{t_{j+1}})^{\downarrow}[^{\uparrow}]$ for every $j \in \{i, \ldots, n-1\}$ and $x^{t_n} \in (B^{t_1})^{\downarrow}[^{\uparrow}]$, implies $x^{t_j} \notin (B^{t_{j+1}})^{\downarrow\downarrow}[^{\uparrow\uparrow}]$ and $x^{t_n} \notin (B^{t_1})^{\downarrow\downarrow}[^{\uparrow\uparrow}]$.*

Figure 7(a) shows the violation of GARP (equivalent to the condition in Corollary 2). Hence, altruism has empirical content. A symmetric example can be constructed to show that envious preferences also have empirical content. Moreover, if a subject is altruistic, she would never use a stick, while an envious subject would never use a carrot. Increasing benevolence is nested within altruistic preferences. Therefore, the stick is not consistent with increasing benevolent preferences. Moreover, if all choices are such that $(x_f, x_s) \in B^s$ and $x_s \geq x^t$ for every $t, s \in \{1, \dots, T\}$, then Proposition 3 can be applied to test for increasing benevolence.

Further we consider inequality-averse preferences. Similarly to the Corollary 1 we need to map budget to the $(f(x_s, x_f), x_f))$ space. Redefine $B^{\downarrow} = \{(f(x'_s, x'_f), x'_f)) : (x_s, x_o) \in B$ such that $x'_f \geq x_f$ and $f(x'_s, x'_f) \leq f(x_s, x_o)\}$ and by $B^{\downarrow\downarrow}$ the strict interior.

Corollary 3. *Let $f(x_s, x_f)$ be an inequality measure. An experiment $E = (x^t, B^t)_{t=1}^T$ is rationalizable with inequality averse preferences if and only if there is no sequence $x^{t_1}, \dots, x^{t_n}$, such that $x^{t_j} \in (B^{t_{j+1}})^{\downarrow}$ for every $j \in \{i, \dots, n-1\}$ and $x^{t_n} \in (B^{t_1})^{\downarrow}$, implies $x^{t_j} \notin (B^{t_{j+1}})^{\downarrow\downarrow}$ and $x^{t_n} \notin (B^{t_1})^{\downarrow\downarrow}$.*

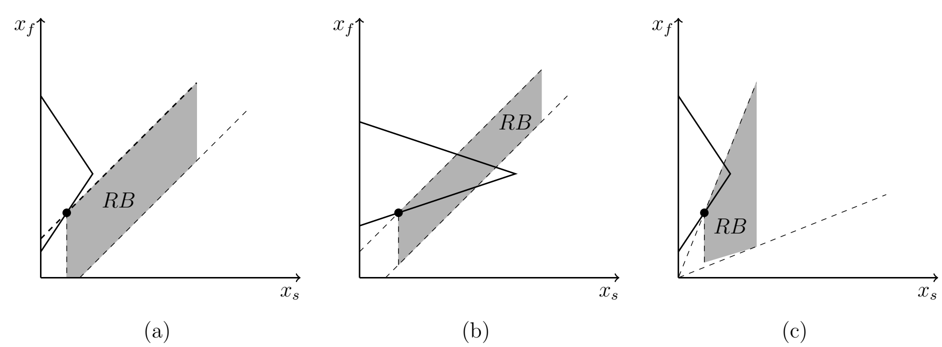


FIGURE 8. Using the stick with inequality averse preferences.

Figure 8(a) presents the case with an inequality averse (in differences) subject using the stick. The shaded area (RB) presents the set of points better than the chosen action. None of the points that dominate the chosen action are in the budget. Figure 8(b) shows that if p is low enough, then using stick is no longer rational. Figure 8(c) shows that testable implications in the carrot-stick game would depend on the particular measure of inequality. In the same budget as Figure 8(a), the choice is consistent with inequality aversion in differences, but not in shares. In addition, this illustrates that implications for the carrot-stick game are different for different measures of inequality. Finally, let us note that inequality aversion is the only of the above-mentioned theories that can rationalize using carrot and stick by the same subject.

4.2 Revealing Altruism and Fairness Cox et al. (2008) provide the revealed preference-based approach to classify people in terms of altruism. A demand function $D(p_s, p_o)$ is **more altruistic than** $\tilde{D}(p_s, p_o)$ if $D_o(p_s, p_o) \geq \tilde{D}_o(p_s, p_o)$ for every $p \in \mathbb{R}_{++}$.

Corollary 4. *Consider consumption experiments E and $\tilde{E}$ in which subjects faced the same prices. Moreover, assume that both experiments satisfy NARP. Experiments are rationalizable with increasing benevolence preferences such that $D(p_s, p_o)$ is more altruistic than $\tilde{D}(p_s, p_o)$ if and only if $\tilde{x}^t \geq x^s$ for all $p^s = \tilde{p}^t$.*

The necessity of this condition is quite obvious since we consider experiments with similar prices; it is also sufficient. That is, we can reconstruct demand functions such that the first experiment will be more altruistic than the second one.

An alternative method is to classify people based on their notion of fairness. Inequality aversion considers equal allocations as fair outcomes; however, the notion of fair outcome can be different. If we were to allow for a subjective notion of fair outcome, the theory would not have any additional empirical content besides GARP (having other-regarding preferences). However, the revealed preference approach can allow us to bound the fair outcome. Moreover, we need to secure that a notion of fair outcome is independent from the menu. Assume that $\chi^* < 1$ is the *fair ratio of payoffs*, that is, the outcome is fair if and only if $x_s/x_o = \chi^*$. We assume that $\chi^* < 1$, as otherwise the subject can be rationalized as inequality averse.

Likewise, we need to slightly change the definition of the inequality measure. We assume that $f(x_s, x_o) = 0$ if and only if $x_s/x_o = \chi^*$. Moreover, $f(x_s, x_o)$ is increasing in x_o and decreasing in x_s if $x_s/x_o <$

χ^*; and $f(x_s, x_o)$ is decreasing in x_o and increasing in x_s if $x_s/x_o > \chi^*$. The last property of an inequality measure needs to be restated as follows: if $x_s/x_o = \chi < \chi^*$, then there is $x'_s \leq x_s$, such that $f(x_s, x_o) \geq f(x_o, x'_s)$.

Corollary 5. *Let χ^* determine the notion of fair outcome. An experiment is rationalizable with inequality averse preferences if and only if an experiment satisfies GARP and $x^t_s \geq \chi^* x_o$ for every $t \in \{1, \dots, T\}$.*

Although the identification in Corollary 5 is deterministic, the constraints obtained can be used as the moment inequalities. This would allow for stochastic bounds that take into account decision making or measurement errors.[15]

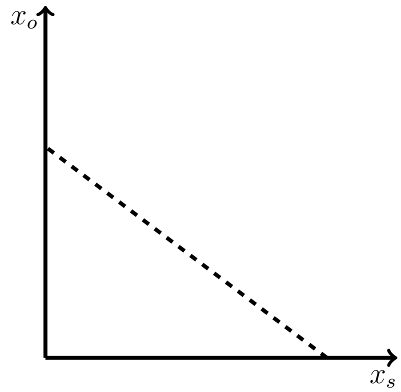
Appendix A: Proofs

Proof of Proposition 2 Before we begin the proof, let us introduce some additional notation. Let $IA = (Y, \geq_{IA})$ be a *partially ordered space*, where $Y \subseteq \mathbb{R}^2_+$, with $y = (x_s, f(x_s, x_o))$ for every $y \in Y$ and $y \geq_{IA} y'$ if $x_s \geq x'_s$ and $f(x_s, x_o) \leq f(x'_s, x'_o)$. Denote by $>_{IA}$ the strict part of $\geq_{IA}$. Note that $f(x_s, x_o)$ defines the injective mapping from X to Y.

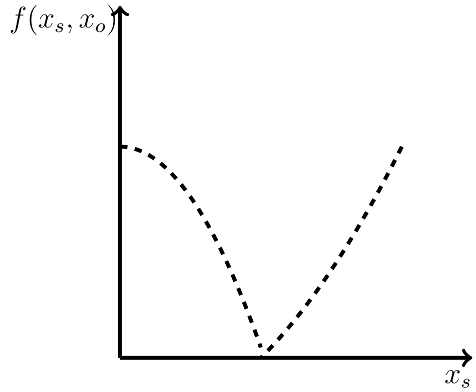
First, let us translate the budget lines to the IA space. Mapped budgets are not necessarily linear. It can be easily seen that the budget set in the new space is compact subset (in natural topology) of IA, since we require the measure of inequality to be a continuous function. Figure 9 illustrates the mapping of the budget line from X to Y.

Denote by B the budget line in IA obtained from mapping a linear budget $px = 1$. Denote by $B^{\downarrow} = \{y :$ there is $y' \in B$ such that $y \geq_{IA} y'\}$ the **downward closure** of budget B. Denote by $B^{\downarrow\downarrow} = \{y :$ there is $y' \in B$ such that $y' >_{IA} y$ the **interior** of budget set B. Denote by $\partial B = B^{\downarrow} \setminus B^{\downarrow\downarrow}$ the **boundary** of $B^{\downarrow}$. Then, an experiment is rationalizable with inequality averse preferences if and only if there

[15] See Chernozhukov et al. (2007) for general results on partial identification and Aguiar and Kashaev (2017) for particular applications to revealed preferences with measurement error. Moreover, methodology from the latter paper directly applies to the results we state.



(a) Budget line in decision space



(b) "Budget line" in Y_{IA} space

FIGURE 9. Mapping of Budget from $(X, \geq)$ to $(Y, \geq_{IA})$.

is a continuous and monotone (with respect to $\geq_{IA}$) utility function $u(x_s, f(x_s, x_o))$, such that observed choices $x^t \in \underset{x \in (B^t)^{\downarrow} \cap \mathbb{R}^2_+}{\operatorname{argmax}} \{u(x_s, f(x_s, x_o))\}$.

Definition 7. *A consumption experiment* $E = (x^t, B^t)_{t=1}^T$ *satisfies* **generalized cyclical consistency (GCC)** *if and only if there is no sequence* $x^{t_1}, \ldots, x^{t_n}$*, such that* $x^{t_j} \in (B^{t_{j+1}})^{\downarrow}$ *for every* $j \in \{i, \ldots, n-1\}$ *and* $x^{t_n} \in (B^{t_1})^{\downarrow}$*, implies* $x^{t_j} \in \partial B^{t_{j+1}}$ *and* $x^{t_n} \in \partial B^{t_1}$*.*

The first lemma can be derived from the results of Nishimura et al. (2017).[16]

Lemma 1. *An experiment satisfies generalized cyclical consistency if and only if there is a continuous, monotone utility function that rationalizes it.*

Figure 10 illustrates the downward closure of the budget in $(Y, \geq_{IA})$ generated by a linear budget in X. It shows that every point for which

[16]One can easily check that conditions on the space are satisfied. That is $\geq_{IA}$ is a continuous order, since we consider a natural topology of $\geq_{IA}$ and order is always continuous in its natural topology. Finally, IA space is Hausdorff and locally compact. Moreover, the mapped budget is compact since it is a result of continuous mapping of compact set.

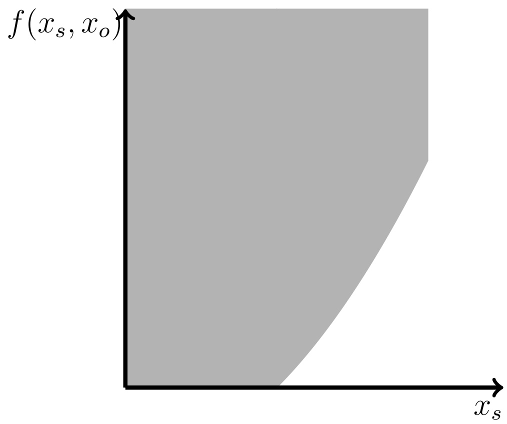


FIGURE 10. Downward closure of the budget in $(Y, \geq_{IA})$.

$x_s < x_o$ lies in the interior of the budget B.[17] Therefore, such a choice could not be generated by inequality averse utility. Hence, further proof proceeds by showing that if we consider $x_s \geq x_o$, then there is a violation of GARP (in X) if and only if there is a violation of generalized cyclical consistency (in IA). Recall that not all points from $B^{\downarrow}$ can be chosen.

To prove that on a half-space such that $x_s \geq x_o$, GARP is equivalent to GCC we show that points are in the linear budget if and only if they are in the $B^{\downarrow}$ and boundary of $B^{\downarrow}$ contains only points such that $px = 1$. We start by characterizing the relation between linear budgets and the budget in IA space.

Lemma 2. *If $x_s \geq x_o$ and $(x'_s, f(x'_s, x'_o)) >_{IA} (\geq_{IA})(x_s, f(x_s, x_o))$, then $(x'_s, x'_o) > (\geq)(x_s, x_o)$.*

Proof. We only prove the case for strict inequalities. Weak inequalities can be proven in a similar fashion. Consider the following cases.

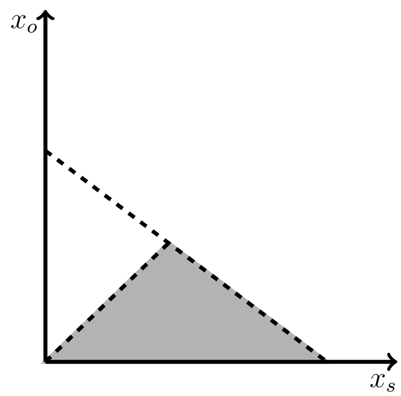
Case 1: $x'_s \geq x'_o$. Then $(x'_s, f(x'_s, x'_o)) >_{IA} (x_s, f(x_s, x_o))$ implies that $x'_s \geq x_s$ and $f(x'_s, x'_o) \leq f(x_s, x_o)$ with at least one inequality being strict. Recall that f is increasing in x_s and decreasing in x_o, hence

[17] Otherwise we can just permute x_s and x_o. Then, the definition of inequality measure would imply that $f(x_o, x_s) \leq f(x_s, x_o)$. Hence, the permuted point is in the interior of the budget.

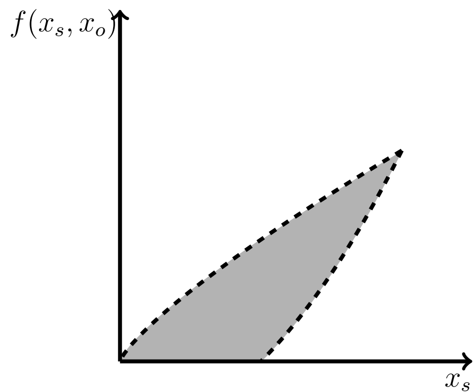
$x'_o > x_o$. This implies that $(x'_s, x'_o) > (x_s, x_o)$.

Case 2: $x'_s < x'_o$. Then $(x'_s, f(x'_s, x'_o)) >_{IA} (x_s, f(x_s, x_o))$ implies that $x'_s \geq x_s$ and $f(x'_s, x'_o) \leq f(x_s, x_o)$ with at least one inequality being strict. At the same time $x'_o > x'_s \geq x'_s \geq x'_o$, hence $(x'_s, x'_o) > (x_s, x_o)$. $\square$

For the simplicity of the further argument, we allow $(x_s, x_o) \in \mathbb{R}^2$; that is, we relax the assumption of non-negativity of the points. This allows us to establish a one-to-one relation between allocations in linear budgets and those in $B^{\downarrow}$. If this holds for all the space $\mathbb{R}^2$, the same would hold for the corresponding subspace.



(a) Budget line in decision space



(b) "Budget line" in Y_{IA} space

FIGURE 11. Mapping of budgets in half-space $x_s \geq x_o$

Further, we restrict our attention to the half-space of $x_s \geq x_o$. Figure 11 illustrates the mapping of the linear budget to the $B^{\downarrow}$. The south east boundary of the budget from (b) corresponds to $px = 1$ from (a) and the north west boundary from (b) corresponds to the bisector from (a). Moreover, the shaded area (interior) from (a) corresponds to the shaded area (interior) from (b). We show this in formal terms.

Lemma 3. *Let $x_s \geq x_o$, then $(x_s, f(x_s, x_o)) \in B^{\downarrow}$ if and only if $px \leq 1$.*

Proof. **($\Rightarrow$)** Take $(x_s, f(x_s, x_o)) \in B^{\downarrow}$. By construction of $B^{\downarrow}$ there is $x' = (x'_s, x'_o)$ such that $px' = 1$ and $(x'_s, f(x'_s, x'_o) \geq_{IA} (x_s, f(x_s, x_o))$, then $x' \geq x$ (see Lemma 2). This implies that $px \leq px' = 1$, which is a contradiction.

(⇐) Take $x = (x_s, x_o)$ such that $px \leq 1$. If $x_s > x_o$, then there is $x'_o \geq x_o$ and $x'_s = x_s$, such that $px' = 1$.[18] This implies that $f(x'_s, x'_o) \leq f(x_s, x_o)$. Therefore, $(x_s, f(x_s, x_o)) \leq_{IA} (x'_s, f(x'_s, x'_o)$, and $px' = 1$, hence, $(x_s, f(x_s, x_o)) \in B^{\downarrow}$. If $x_s = x_o$, let $\varepsilon = \frac{1-px}{p_s+p_o} \geq 0$ and $(x'_s, x'_o) = (x_s + \varepsilon, x_o + \varepsilon)$. Then, $(x_s, f(x_s, x_o)) \leq_{IA} (x'_s, f(x'_s, x'_o)$ (since both of them equal to zero), and $px' = 1$, hence, $(x_s, f(x_s, x_o)) \in B^{\downarrow}$. □

Further we show that only points from the boundary of the linear budget can be on the boundary of $B^{\downarrow}$and vice versa.

Lemma 4. *$(x_s, f(x_s, x_o)) \in \partial B$ if and only if $x_s \geq x_o$ and $px = 1$.*

Proof. (⇒) If $(x_s, f(x_s, x_o)) \in \partial B$ then $x_s \geq x_o$ is trivially implied. Then, we can just assign $x'_s = x_o$ and $x'_o = x_s$ and obtain a contradiction since $(x_s, f(x_s, x_o)) <_{IA} (x'_s, f(x'_s, x'_o))$. Therefore, we are left to show that $px = 1$. On the contrary, assume that $px < 1$. Then, we can apply the construction as in the proof of Lemma 3 to get a bundle x' such that $px' = 1$ and $(x_s, f(x_s, x_o)) <_{IA} (x'_s, f(x'_s, x'_o))$. This implies a contradiction

(⇐) On the contrary, assume that $(x_s, f(x_s, x_o)) \in B^{\downarrow\downarrow}$. Then, there is $x' = (x'_s, x'_o)$ on the boundary of $B^{\downarrow}$ such that $(x_s, f(x_s, x_o)) <_{IA} (x'_s, f(x'_s, x'_o))$. Hence, Lemma 2 implies that $1 = px < px'$. This immediately implies a contradiction. □

These two results imply the following corollary.

Corollary 6. *Let $x_s \geq x_o$, then $(x_s, f(x_s, x_o)) \in B^{\downarrow\downarrow}$ if and only if $px < 1$.*

The proof is omitted, as it is similar to the previous ones. To prove that every point from the interior of $B^{\downarrow}$ is strictly inside of a linear budget we assume the contrary and obtain a contradiction from Lemma

[18]This is due to the fact that $f(x_s, x_o)$ is decreasing in x_o for $x_s \geq x_o$. If for that case $x'_o > x_s$, then set $x'_o = x_s$ and apply construction used for the case of $x_s = x_o$.

4. To prove the reverse, we assume that point is in the border and again obtain a contradiction using Lemma 4.

These preliminary results allow us to finalize the proof.

Corollary 7. *Let $x_s^t \geq x_o^t$ for all $t \in \{1, \ldots, T\}$. An experiment satisfies GARP if and only if it satisfies generalized cyclical consistency.*

Proof. ($\Rightarrow$) On the contrary, assume that there is a violation of generalized cyclical consistency. Then, there is a sequence $x^{t_1}, \ldots, x^{t_n}$, such that $x^{t_j} \in (B^{t_{j+1}})^{\downarrow}$ for every $j \in \{i, \ldots, n-1\}$ and $x^{t_n} \in (B^{t_1})^{\downarrow\downarrow}$. Lemma 3 implies that $p^{t_{j+1}} x^{t_j} \leq 1$ and Corollary 6 implies that $p^{t_1} x^{t_n} < 1$ that is a violation of GARP.

($\Leftarrow$) On the contrary, assume that there is a violation of GARP. Then, there is a sequence $x^{t_1}, \ldots, x^{t_n}$, such that $p^{t_{j+1}} x^{t_j} \leq 1$ for every $j \in \{i, \ldots, n-1\}$ and $p^{t_1} x^{t_n} < 1$. Lemma 3 implies that $x^{t_n} \in (B^{t_1})^{\downarrow\downarrow}$ Corollary 6 implies that $x^{t_n} \in (B^{t_1})^{\downarrow\downarrow}$. Hence, there is a violation of generalized cyclical consistency. □

Lemma 7 together with the observation that choices such as $x_s < x_o$ cannot be rationalized with inequality averse preferences would guarantee that GARP and $x_s^t \geq x_o^s$ for every $t, s \in \{1, \ldots, T\}$ are necessary and sufficient conditions for rationalization with inequality averse preferences. As Lemma 1 shows, the generalized cyclical consistency is equivalent to rationalizability with inequality averse preferences.

Proof of Corollary 1 The budget in this case is $B^t = \{(x_s^t, f(x_f^t, x_s^t)), (0,0)\}$. The comprehensive closure of the budget is $(B^t)^{\downarrow} = B^t \cup R^t$. The interior of the comprehensive closure of the budget is $(B^t)^{\downarrow\downarrow} = R^t$. Moreover, one can easily see that $x^s \in A^t$ if and only if $x^t \in R^t$. Hence, the proof can be concluded by applying Lemma 1.

Proof of Corollaries 2 and 3 Proofs of both Corollaries directly follow from Lemma 1 with the corresponding definition of orders.

Proof of Corollary 4 Let us introduce some additional notation. We rely on the result from Cherchye et al. (2018). Let $\omega = \frac{p_o}{p_s}$ and $m = \frac{1}{p_s}$.

Therefore, we can refer to the demands as functions of only ω and m. Moreover, we reenumerate observations s.t. $\tilde{p}^t = p^t$.

Define $\alpha, \beta > 0$ such that

$$1+\beta < \min\left\{\min_{t,s}\left\{\frac{x_o^t}{x_o^s} : x_o^t > x_o^s\right\}, \min_{t,s}\left\{\frac{\tilde{x}_o^t}{\tilde{x}_o^s} : \tilde{x}_o^t > \tilde{x}_o^s\right\}\right\}$$

and

$$a(1+\beta) < \min\left\{\min_{t,s}\left\{\frac{x_o^t}{x_o^s}\right\}, \min_{t,s}\left\{\frac{\tilde{x}_o^t}{\tilde{x}_o^s}\right\}\right\}$$

Let $\delta_{t,v} = \max\{|w^t - w^v|, |m^t - m^v + (w^t - w^v)x_o^t|, |m^t - m^v + (w^t - w^v)\tilde{x}_o^t|\}$ and let $\varepsilon < \min \delta_{t,v}$. Consider the function

$$g(z) = \begin{cases} \alpha & \text{for } z \leq -\varepsilon \\ 1 + \frac{1-\alpha}{\varepsilon} z & \text{for } -\varepsilon \leq z \leq 0 \\ 1 & \text{for } z \geq 0 \end{cases}$$

In addition, consider the function

$$h(z) = \begin{cases} \alpha \frac{1}{z+\varepsilon-1} & \text{for } z < -\varepsilon \\ 1 + \frac{1-\alpha}{\varepsilon} z & \text{for } -\varepsilon \leq z \leq 0 \\ 1 + \beta \frac{z}{z+1} & \text{for } z \geq 0 \end{cases}$$

Then, according to Cherchye et al. (2018), the rationalization of the demand can be obtained as a solution of the following program.

$$\begin{aligned} D_o(w,m) = &\max_r r \\ &\text{s.t. } g(w-w^t)h(m^t + (w-w^t)r - m)r \leq x_o^t \ \forall t \in \{1,\dots,T\} \\ &wr \leq x \end{aligned}$$

For both experiments, the functions $g(z)$ and $h(z)$ are the same. Moreover, since both experiments are composed of the same set of prices, every left hand side of the constraints would be the same for both experiments. In addition, if the left hand side is decreasing in r, then constraint is not binding, hence, we need to concentrate only on increasing left hand sides. At the same time every $\tilde{x}_o^t \geq x_o^t$, hence there is larger r and, consequently, a larger demand for x_o at given prices. Therefore, the $\tilde{D}_o(w,m) \geq D_o(w,m)$, i.e. $\tilde{E}$ is more altruistic than E.

Proof of Corollary 5 The proof of this remark is same as the proof of Proposition 2, and is therefore omitted. Note that the last property of the redefined inequality measure still guarantees that if $x_s/x_o < \chi$, then it would be in the strict interior of the budget in IA.

Appendix B: Computing Houtman-Maks index (HMI)

Multiple approaches to the problem of calculating HMI have been taken in the literature (Choi et al. (2014), Heufer and Hjertstrand (2015)). Choi et al. (2014) and Dean and Martin (2016) used a set cover problem approach to calculate HMI, and we follow in a similar vein. The program below calculates HMI, but does not rely on the linearity of budgets.

We first discuss the calculation of HMI for GARP, which allows us to test for other-regarding preferences and inequality aversion. We will then discuss HMI for NARP, necessary for testing increasing benevolence and other normality-related theories. We require two constants to implement the method. We use the big-M method to selectively activate constraints for a subset of data. Let $M > 1$ be the big-M. Moreover, since strict constraints do not make sense for practical optimization, we introduce an infinitesimal tolerance term $\epsilon < \frac{1}{n}$.

We take as given preference relation $\geq$ with strict part $>$ on space X with $|X| = n$. The (mixed integer) linear program is then to find such $u_i \in [0,1]$ and $\delta \in \{0,1\}$ for all elements in X, indexed by $i \leq n$ that minimize $\sum_{i=1}^{n} \delta_i$, so that

$$u_i \geq u_j + \epsilon - M(\delta_i + \delta_j) \text{ for } \{(x_i, x_j) \in X^2 : x_i > x_j\} \tag{1a}$$

$$u_i \geq u_j - M(\delta_i + \delta_j) \text{ for } \{(x_i, x_j) \in X^2 : x_i \geq x_j\} \tag{1b}$$

Binary variables δ in the linear program make constraints active only for the chosen subset of observations, and we thus ensure that this subset is minimal. Then we can simply calculate HMI as $1 - \frac{\sum_{i=1}^{n} \delta_i}{n}$.

To calculate HMI for NARP, we only need to adjust the constraints in the program to Definition 6. We take all pairs of observations $t, v \in \{1, \ldots, T\}$ for which $p_o^t/p_s^t \leq p_o^v/p_s^v$ and $x_s^v \leq \frac{1 - p_o^t x_o^v}{p_s^t}$. These are the observations for which NARP has implications. The (mixed integer) linear program is then to find such $\delta \in \{0,1\}$ for all t, v above that minimize $\sum_{i=1}^{n} \delta_i$, so that

$$x_o^v \leq x_o^t + M(\delta_t + \delta_v) \tag{2}$$

HMI for JNARP, required for rationalization with both x_s and x_o normal, follows in the same way, but with conditions replaced with those from Definition 8. We therefore omit it here.

Appendix C: Classification

This section discusses our classification methodology and results in more detail. We use the clustering technique from Liu et al. (2000), who developed a method to apply decision trees to the problem of organizing unlabeled data.[19] Decision trees are appropriate for our data due to the nested nature of theories.

To select HMI thresholds for the classification rule we use the information gain purity function from Liu et al. (2000). Unlike usual distance-based measures (e.g. sum of squares) it can potentially be interpreted regardless of the environment. Information gain is the difference in the expected information needed to identify observed data points against uniform distribution before and after the test. In other words, we select the test that minimizes the weighted entropy for all classes of subjects. The intuition behind this approach is the following. We apply sequential binary tests for different theories, checking if each data point passes the test at a given threshold level or not. All of these binary tests convey one bit of information about each data point: whether it passes or not. If the performance of the test is indistinguishable from testing uniformly distributed data, then the test conveys no information. If, however, the performance on real data is clearly different from the random data, we would suspect that the test conveys some information about the population, and we would like to maximize this information. Information theory suggests that information in bits conveyed by such tests can be measured as the negative

[19]Classification trees are widely applied, although mostly as supervised learning technique. That is, it requires having "training" dataset which is already categorized.

logarithm to the base 2 of the number of possible outcomes described by the test.

To calculate the information gain for a group of n data points we introduce n additional fictional points that have uniformly distributed HMI. We then calculate the expected amount of information needed to classify real points against these uniformly distributed points before and after every binary test. Formally, the information gain from clustering with some threshold α is:

$$1 - \frac{1}{2}\left(\left(\alpha + \frac{N_F}{n}\right) E_F + \left((1-\alpha) + \frac{N_P}{n}\right) E_P\right), \tag{3}$$

where n - number of subjects, N_F (N_P) - number of subjects who fail (pass) the test at $HMI \geq \alpha$, E_F (E_P) - entropy for the class of data that passes (fails) the test. Then $(1-\alpha)$ and α are the fractions of points with uniformly distributed HMI that would respectively pass and fail the test.

Information required to identify a data point against a uniform random draw before clustering is 1 bit. After clustering this information is the weighted sum of entropy in each cluster, which in turn is calculated for the points failing the test as

$$E_F = -D_F^R log_2(D_F^R) - D_F^D log_2(D_F^D),$$

where $D_F^R = \frac{\alpha n}{\alpha n + N_F}$ and $D_F^D = \frac{N_F}{\alpha n + N_F}$. These are the fractions of random and real data points in the cluster that fails the test. The entropy for the cluster that passes the test is calculated in the same manner with fractions $D_P^R = \frac{(1-\alpha)n}{(1-\alpha)n + N_P}$ and $D_P^D = \frac{N_P}{(1-\alpha)n + N_P}$.

By substituting expressions for E_F and E_P in (3), we obtain a simplified expression:

$$1 - \frac{1}{2}\left(\alpha I_P^R + (1-\alpha) I_F^R + \frac{N_P}{n} I_P^D + \frac{N_F}{n} I_F^D\right),$$

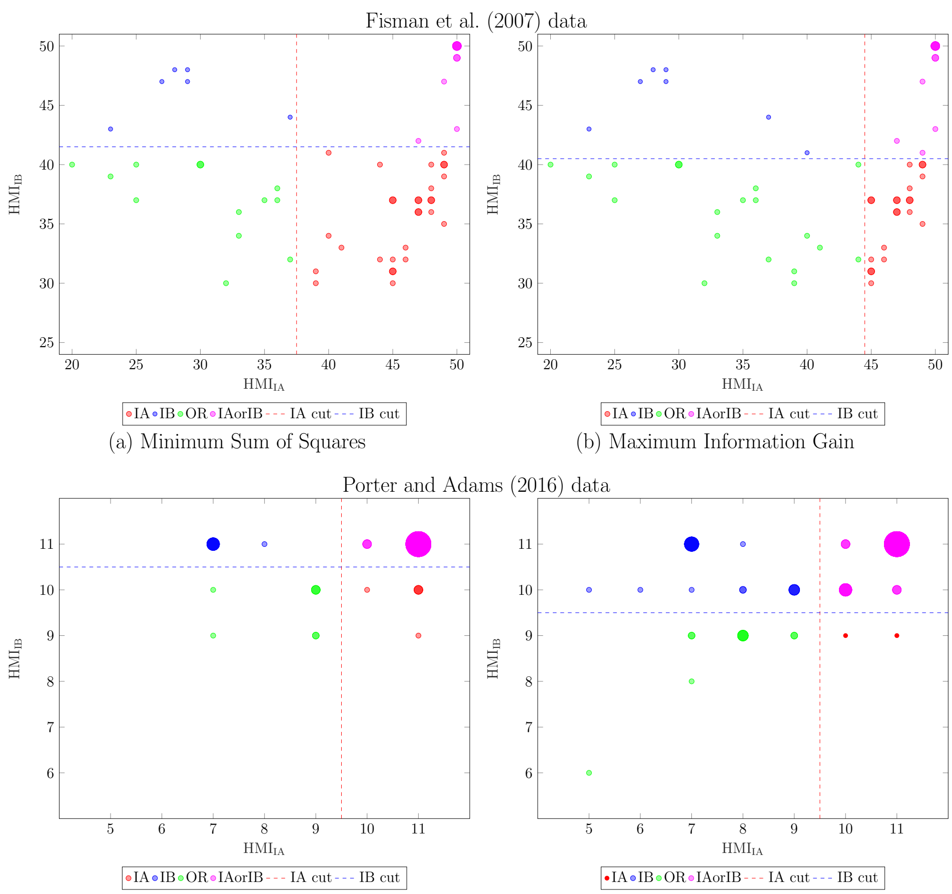
where

$$I_P^R = -log_2\left(D_P^R\right),\ I_F^R = -log_2\left(D_F^R\right),$$

$$I_P^D = -log_2\left(D_P^D\right),\ I_F^D = -log_2\left(D_F^D\right),$$

These four terms represent the information from identifying a point as a data point (D) or as a random point (R) for points that pass the test (P) and fail the test (F). Recall that entropy is the expected amount of information required to decide if some point is an observed data-point or a uniformly-generated random point given the result of a binary test.

We apply this procedure sequentially, first separating the inconsistent cluster by applying the test for other-regarding preferences and then clustering the remaining data according to nested theories. However, we omit the inconsistent points in the figure below.



(a) Minimum Sum of Squares (b) Maximum Information Gain

(c) Minimum Sum of Squares (d) Maximum Information Gain

FIGURE 12. Mixed Types Analysis

Classification thresholds for nested theories and classified points are presented in Figure 12 along with the alternative classification based on minimizing within-cluster sums of square distances from cluster means.

Fisman et al. (2007) data				
IA	IB	IA or IB	OR	IC
29 (38%)	6 (8%)	10 (13%)	13 (17%)	18 (24%)
Porter and Adams (2016) data				
IA	IB	IA or IB	OR	IC
5 (6%)	8 (9%)	40 (45%)	7 (8%)	26 (29%)

Table 2. Classification of Subjects (Minimal sum of squares)

The latter largely agrees with our information gain measure, as can be seen from comparing classification results in Figure 2 and Table 2. The classification is fairly robust to other approaches: 4-means clustering of nested theories agrees with our classification only for half of the data, but qualitatively the results are similar.

Appendix D: Detailed Empirical Results

D.1: Additional Theories As an alternative assumption to the increasing benevolence Cox et al. (2008) offered a similar condition, but for one's own payoff. Moreover, it would also be a legitimate assumption to claim normality of both goods. These are two additional theories we are going to test.

Normality of keeping (x_s) may organize the data better, as has been partially shown by Cherchye et al. (2018). While the normality of x_s can be checked by applying the permuted version of NARP, the normality in both giving and keeping needs a different test called Joint Normality Axiom of Revealed Preferences.

Definition 8. *An experiment* $E = (x^t, p^t)_{t=1}^T$ *is consistent with* **Joint Normality Axiom of Revealed Preference (JNARP)** *if and only if for all observations* $t, v \in \{1, \dots, T\}$ *if* $p_o^t/p_s^t \leq p_o^v/p_s^v$ *and* $x_o^t < x_o^v$, *then* $x_s^t \leq x_s^v$.

Both x_s and x_o are normal goods if and only if an experiment satisfies JNARP. Proof of this fact uses the same logic as proof of Proposition 3 and Theorem 2 in Cherchye et al. (2018) and is omitted.

D.2: Fisman et al. (2007) Data.

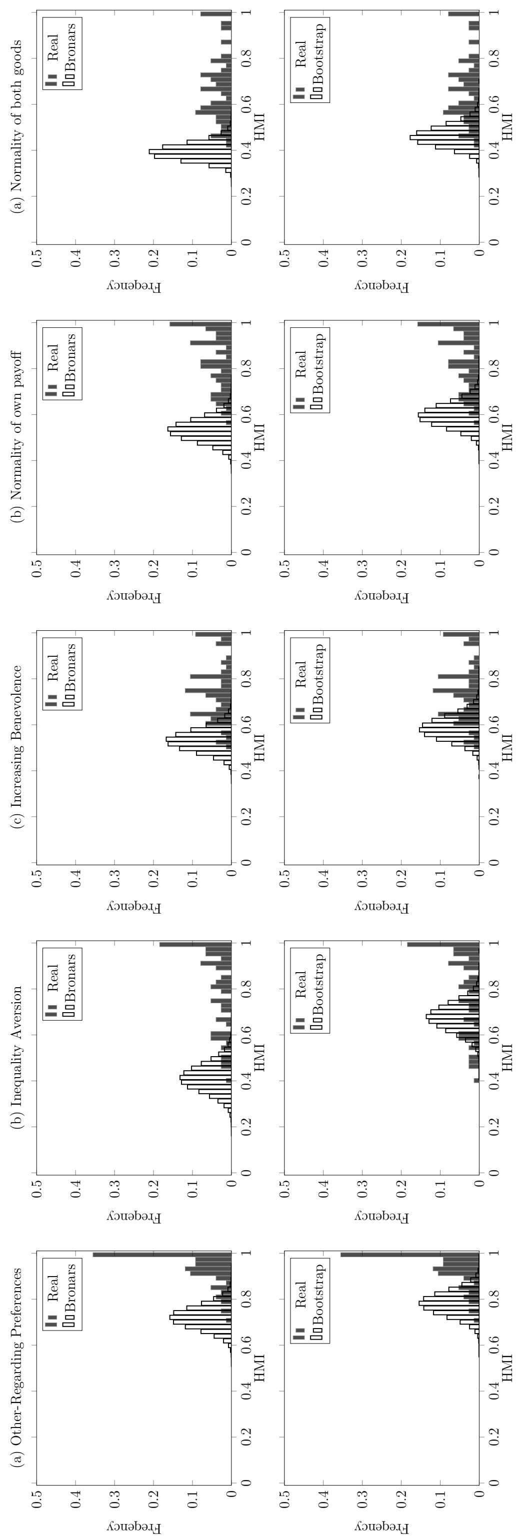


FIGURE 13. Distributions of HMI for Fisman et al. (2007) Data

Figure 13 presents the distributions of HMI indexes for the theories that we are testing. The first row demonstrates the distribution of the HMI for the real subjects in comparison to the distribution of the HMI for the Bronars' power test. The second row presents the distribution of the HMI for the real subjects in comparison to the distribution of the HMI for the bootstrap power test. In order to test the theory, we compare the distribution of its HMI to the distribution of powers of the test. For all five theories, we see that the real subjects pass both the Bronars' and the bootstrap tests; that is, they perform better than random subjects.[20] We confirm that all five theories have empirical support, and at most a quarter of the data needs to be dropped to rationalize an average subject.

Further, we present comparisons of the theories. For this part of the analysis, we restrict our attention to the HMI levels of .8, .9, .95 and 1 (no deviations).

Figure 14 presents the pass rates with a confidence interval for every theory.[21] We can see that other-regarding preferences demonstrate higher pass rates for all given levels of HMI.

[20]For other-regarding preferences the mean HMI for real subjects is .92; for Bronars subjects it is .7; for bootstrap subjects it is .76 (p-values$<$ .001 for both comparisons using Wilcoxon test and t-tests). For inequality aversion the mean HMI for real subjects is .79; for Bronars subjects it is .4; for bootstrap subjects it is .66 (p-values$<$ .001 for both comparisons using Wilcoxon test and t-tests). For increasing benevolence preferences the mean HMI for real subjects is .74; for Bronars subjects it is .52; for bootstrap subjects it is .57 (p-values$<$ .001 for both comparisons using Wilcoxon test and t-tests). For normality of x_s the mean HMI for real subjects is .83; for Bronars subjects it is .52; for bootstrap subjects it is .58 (p-values$<$ .001 for both comparisons using Wilcoxon test and t-tests). For normality of both x_s and x_o the mean HMI for real subjects is .67; for Bronars subjects it is .38; for bootstrap subjects it is .45 (p-values$<$ .001 for both comparisons using Wilcoxon test and t-tests).

[21]Confidence intervals are computed using the Clopper-Pearson procedure, since the pass rate can be perceived as a binomial variable.

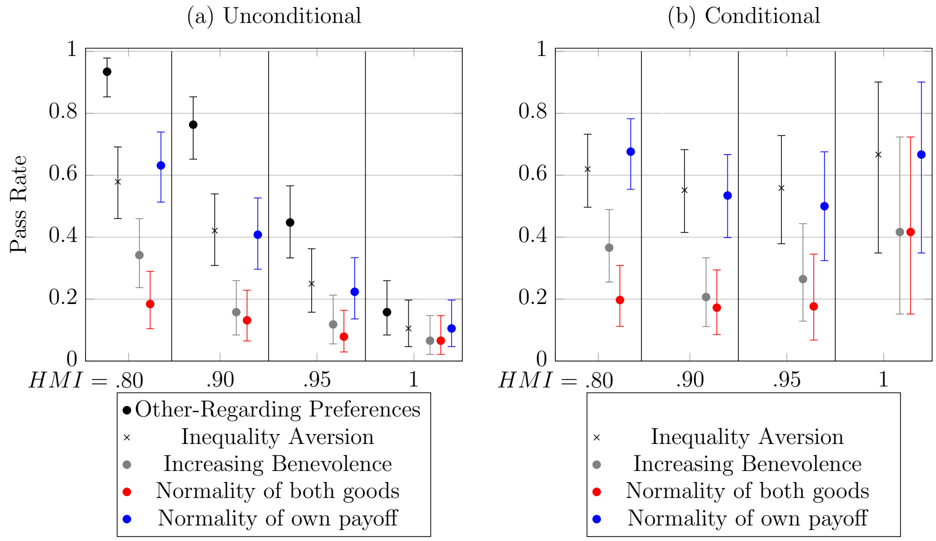


FIGURE 14. Pass Rates for Fisman et al. (2007) Data

Figure 14(a) shows that increasing benevolence preferences perform significantly worse than other-regarding preferences. Inequality aversion does better than increasing benevolence (the difference is significant for the HMI levels of .8, .9) and worse than other-regarding preferences (the difference is significant for the HMI levels of .8, .9). Figure 14(b) presents the conditional pass rates. In particular, it shows that inequality aversion still outperforms the increasing benevolence preferences (the difference is significant for the HMI levels of .8, .9). Moreover, if we assume the inequality aversion preferences, then at least 20% of population who have other-regarding preferences are not behaving as if they have inequality aversion preferences. Assuming increasing benevolence preferences would cost us about 50% of population. This shows that all nested theories are significantly restrictive. Furthermore, the normality of x_s organizes data better than increasing benevolence and joint normality. This result confirms the findings of Cherchye et al. (2018) who applied these tests to the Andreoni and Miller (2002) data. Finally, the normality in the own payoff performs as successfully as inequality aversion in this data.

The top row in Figure 15 shows the predictive success levels with confidence intervals with both Bronars and bootstrap as the control.

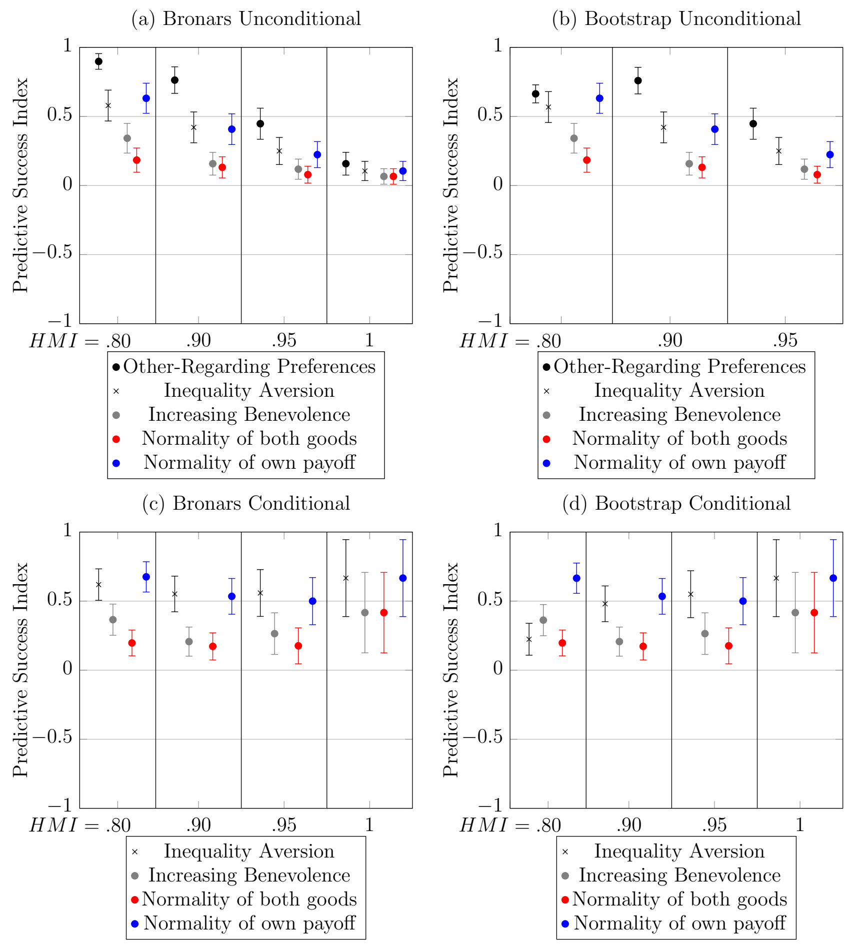


FIGURE 15. Predictive Success Index for Fisman et al. (2007) Data

First of all, we see that the lower bounds of the confidence intervals for the predictive success of all the theories are above zero. Therefore, all of the presented theories predicts the observed behavior better than random (Bronars or bootstrap) decision making. Comparing the predictive success of other-regarding and inequality aversion preferences we see mixed evidence. While the predictive success is always higher for other-regarding preferences, the difference is not always significant. We also see that other-regarding preferences outperform the increasing benevolence preferences at every level of HMI for both Bronars and bootstrap random controls.

In order to compare the predictive success of inequality aversion and increasing benevolence preferences, we take into account the fact that theories are nested. This requires us to not only use the subset of subjects who are consistent with the other-regarding preferences at the given HMI, but apart from that, to use random subjects who are also consistent with other-regarding preferences rationalization at the given level of HMI. In particular, if we consider the conditional predictive success for HMI=90%, we take only the subsample of subjects who are consistent with other-regarding preferences rationalization with the HMI of 90%. We then generate for every set of budgets a 1000 random pseudo-subjects who are consistent with other-regarding preferences rationalization with the HMI of 90%. We then compute their HMIs for inequality aversion and increasing benevolence preferences to use those as *random pass rates* for corresponding predictive success indexes. In order to generate random choices that are consistent with other-regarding preferences rationalization at a given HMI, we use the following procedure. We take the random subsample of the experiment that contains 80%, 90% or 95% of budgets and generate the random choices which are consistent with other-regarding preferences rationalization using the Heufer (2014) procedure. We unconditionally place the random choices for the remaining budgets. We calculate this both for the Bronars' and bootstrap tests. The first case follows Heufer (2014) exactly, generating choices that approximate a uniform distribution on each budget, while satisfying GARP. With bootstrap we only need to truncate the empirical distribution at each step to the admissible region, and draw choice points from the resulting distribution. This step is trivial for the Bronars case, since conditional distribution is also a uniform distribution.

The bottom row in Figure 15 presents the conditional predictive success index. Under the Bronars test, inequality averse preferences theory outperforms increasing benevolence preferences at every level of HMI, and difference is significant at $HMI = .8$ and $HMI = .9$. Under the bootstrap test, we see that for higher levels of HMI, inequality aversion performs better than increasing benevolence, while difference is only significant at $HMI = .9$.

D.3: Porter and Adams (2016) Data. In this experiment every subject has played two sets of budgets in a row. One of them giving to strangers and the other one giving to parents, while the order differs between treatments. We only consider giving to strangers since we want to remain consistent with the analysis we conducted for Fisman et al. (2007) data.[22] We only consider treatments in which players started with a dictator game with strangers to guarantee comparability with the other dataset. Moreover, the second part was a surprise for subjects (it was not announced before the end of the first part), so we can consider the games as comparable.

However, there are some differences in design. The first difference is the population. Fisman et al. (2007) conducted an experiment with undergraduates from UC Berkley, while Porter and Adams (2016) used a sample of the adults from the southeast region of the UK. Another important difference is that Fisman et al. (2007) used a constant exchange rate of tokens (experimental currency) to dollars, while Porter and Adams (2016) had a changing exchange rate, through which the price variation was implemented. Let us illustrate this with an example of a decision problem. The subject is given 40 tokens and can decide how much to pass and to hold, while every token she holds converts into 10 pence and every token she passes converts into 30 pence. If one wants to get an equal allocation of tokens (20 pass and 20 hold), then the allocation of real world currency will not be equal (6 pounds pass and 2 pounds hold). One the other hand, if one wants to keep an equal allocation of real world currency (3 pounds pass and 3 pounds hold), then the allocation of tokens will not be equal (10 tokens pass and 30 tokens hold).

This is important, as the inequality measure for Fisman et al. (2007) data would be similar whether we think about endowments in tokens or dollars. For Porter and Adams (2016), it would be different. Therefore, we look at both inequality aversion in real currency and in experimental currency. Remark 5 allows us to test for inequality aversion in experimental currency in the same manner as for inequality aversion

[22]Porter and Adams (2016) show that preferences of giving to strangers and parents are significantly different.

in real currency. We simply set the equal allocation of tokens as the fair outcome x^*, while the inequality measure is still in terms of real currency payoffs.

We present an analysis similar to one we conducted for Fisman et al. (2007) data with the only difference that we have two separate versions of inequality aversion. We also report tests for normality of own payoff and normality of both goods. This brings the total number of theories to six. In addition, since this experiment has fewer budgets, we use the following levels of HMI$= 9/11, 10/11, 11/11$.

Figure 16 presents the distribution of the HMI for all theories. As before, we use the Bronars and bootstrap tests to estimate power. Figure 16 consists of six panels: (a) for other-regarding preferences; (b) for inequality aversion in real currency; (c) for inequality aversion in experimental currency; (d) for increasing benevolence preferences; (e) for normality of own payoff and (f) for normality of both goods. Most theories outperform random decision making for this data as well.[23] The exception is inequality aversion in experimental currency. Although it performs better under the Bronars test, it shows almost the same levels of HMI as bootstrap subjects.[24]

[23]For other-regarding preferences, the mean HMI for real subjects is .96; for Bronars subjects it is .82; for bootstrap subjects it is .78 (p-values$< .001$ for both comparisons using Wilcoxon test and t-tests). For inequality aversion in real currency, the mean HMI for real subjects is .84; for Bronars subjects it is .46; for bootstrap subjects it is .59 (p-values$< .001$ for both comparisons using Wilcoxon test and t-tests). For increasing benevolence preferences, the mean HMI for real subjects is .93; for Bronars subjects it is .68; for bootstrap subjects it is .7 (p-values$< .001$ for both comparisons using Wilcoxon test and t-tests). For normality of x_s, the mean HMI for real subjects is .95; for Bronars subjects it is .68; for bootstrap subjects it is .68 (p-values$< .001$ for both comparisons using Wilcoxon test and t-tests). For normality of both x_s and x_o, the mean HMI for real subjects is .92; for Bronars subjects it is .57; for bootstrap subjects it is .6 (p-values$< .001$ for both comparisons using Wilcoxon test and t-tests).

[24]For inequality aversion in experimental currency the mean HMI for real subjects is .77; for Bronars subjects it is .49; for bootstrap subjects it is .68 (p-values$< .001$ for both comparisons using Wilcoxon test and t-tests).

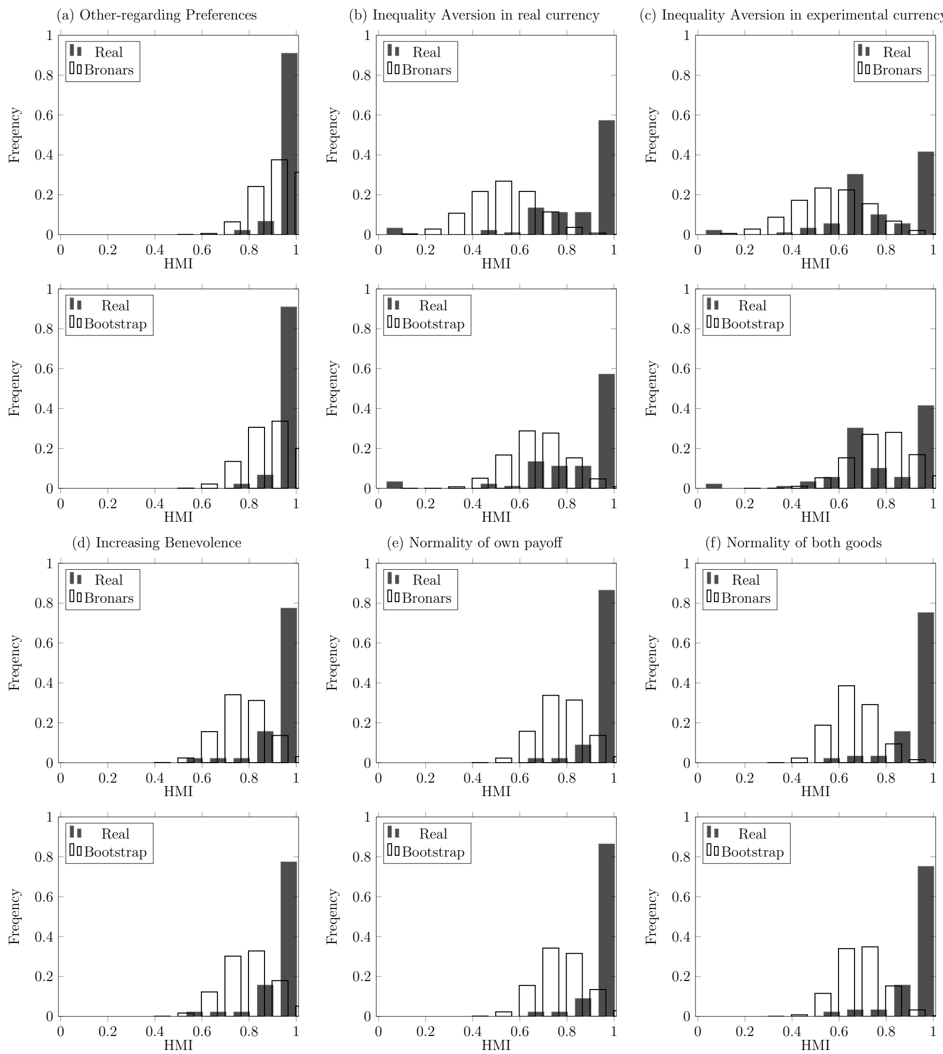


Figure 16. Distributions of HMI for Porter and Adams (2016) Data

Figure 13 shows that regardless of framing, subjects are more prone to be inequality averse in real currency than in experimental one. Moreover, power for increasing benevolence is lower in this experiment. Therefore, further comparison should be done based on the predictive success index.

Figure 17 presents the pass rates for all theories. Since the latter five theories are nested within the other-regarding preferences, we also present the pass rates conditional on a subject being consistent with

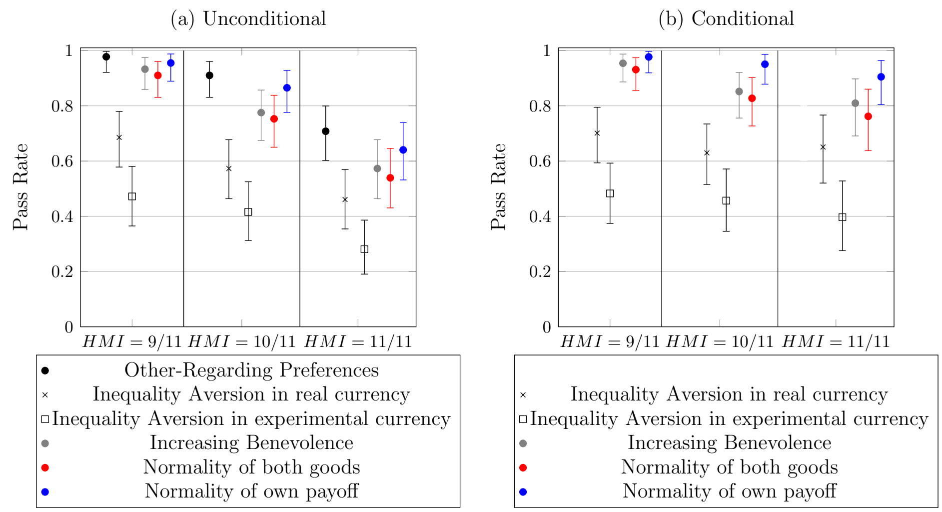


FIGURE 17. Pass Rates for Porter and Adams (2016) Data

other-regarding preferences at a given level of HMI. All nested theories are significantly restrictive, moreover, we observe the ordering of nested theories which is reverse from the one obtained using Fisman et al. (2007) data. Increasing benevolence preferences tend to be more consistent with the data than inequality averse preferences. Moreover, the difference is statistically significant for $HMI = 9/11$ and $HMI = 10/11$. Let us move on to the predictive success in order to further investigate this, while controlling for the power of the test.

Figure 18 presents the value of predictive success indexes. As before, we use the Bronars and bootstrap tests to control for both conditional and unconditional predictive success. Ordering of the theories is preserved controlling for power of the test if we restrict HMI for a high enough level ($\geq 10/11$). Observation that the subject's own payoff appears to act as a normal good carries on to this dataset as well. Normality of x_s organizes data better than normality of x_o. Moreover the difference is statistically significant for the bootstrap test and $HMI \leq 10/11$. Additionally, due to the small amount of budgets for the low levels of HMI, normality in both goods looks rather favorable, since it has the higher power. However, this effect disappears at high enough levels of HMI.

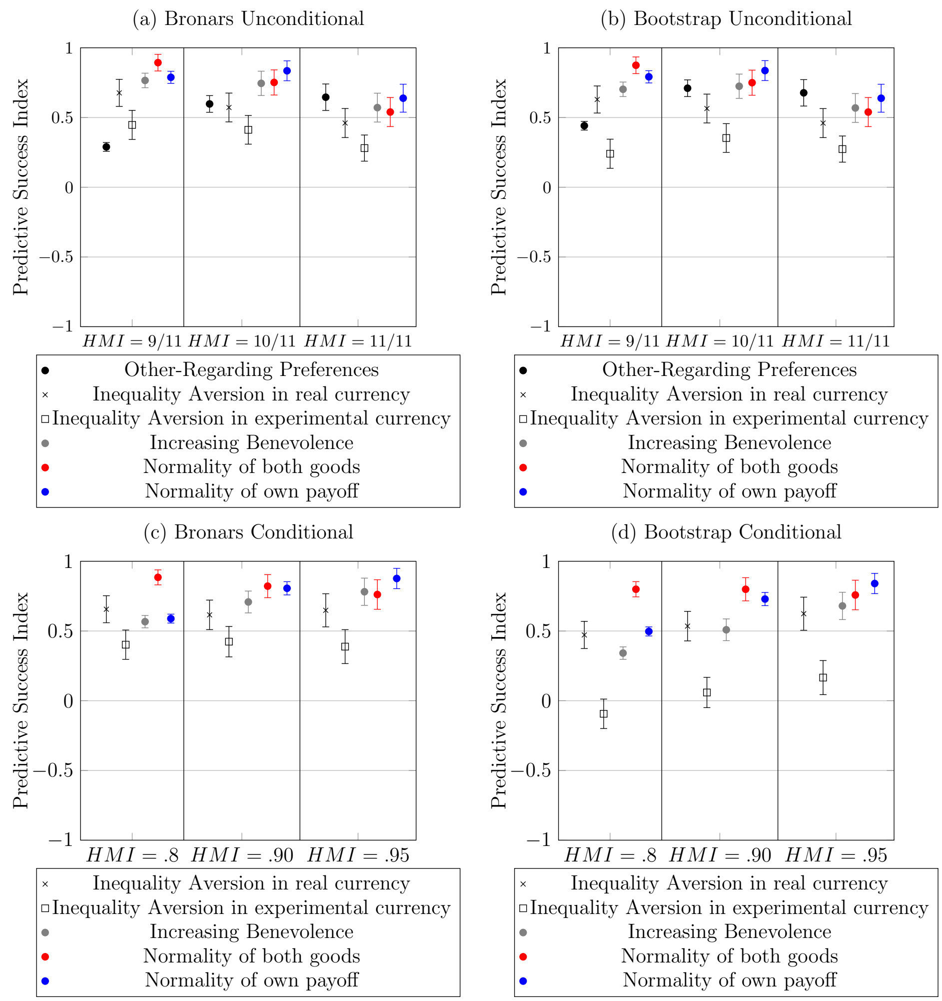


FIGURE 18. Predictive Success Index for Porter and Adams (2016) Data

REFERENCES

AFRIAT, S. N. (1967): "The construction of utility functions from expenditure data," *International economic review*, 8, 67–77.

——— (1973): "On a system of inequalities in demand analysis: an extension of the classical method," *International economic review*, 460–472.

AGRANOV, M. AND T. R. PALFREY (2015): "Equilibrium tax rates and income redistribution: A laboratory study," *Journal of Public*

Economics, 130, 45–58.

Aguiar, V. and N. Kashaev (2017): "Stochastic Revealed Preferences with Measurement Error: Testing for Exponential Discounting in Survey Data," *Working Paper*.

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

Andreoni, J. and J. Miller (2002): "Giving according to GARP: An experimental test of the consistency of preferences for altruism," *Econometrica*, 70, 737–753.

Apesteguia, J. and M. A. Ballester (2015): "A measure of rationality and welfare," *Journal of Political Economy*, 123, 1278–1310.

Beatty, T. K. and I. A. Crawford (2011): "How demanding is the revealed preference approach to demand?" *The American Economic Review*, 101, 2782–2795.

Becker, G. S. (1974): "A theory of social interactions," *Journal of political economy*, 82, 1063–1093.

Becker, N., K. Häger, and J. Heufer (2015): "That's justnot fair: Gender differences in notions of justice," *Ruhr Economic Papers, No. 578*.

Benjamin, D. J. (2015): "Distributional preferences, reciprocity-like behavior, and efficiency in bilateral exchange," *American Economic Journal: Microeconomics*, 7, 70–98.

Bolton, G. E. and A. Ockenfels (2000): "ERC: A theory of equity, reciprocity, and competition," *American economic review*, 166–193.

——— (2006): "Inequality aversion, efficiency, and maximin preferences in simple distribution experiments: comment," *The American economic review*, 96, 1906–1911.

Bronars, S. G. (1987): "The power of nonparametric tests of preference maximization," *Econometrica: Journal of the Econometric Society*, 693–698.

Castillo, M. E., P. J. Cross, and M. Freer (2017): "Nonparametric utility theory in strategic settings: Revealing preferences and beliefs from games of proposal and response," *Working Paper*.

Chambers, C. P. and F. Echenique (2016): *Revealed preference theory*, vol. 56, Cambridge University Press.

Charness, G. and M. Rabin (2002): "Understanding social preferences with simple tests," *The Quarterly Journal of Economics*, 117, 817–869.

Cherchye, L., T. Demuynck, and B. De Rock (2018): "Normality of demand in a two-goods setting," *Journal of Economic Theory*, 173, 361–382.

Chernozhukov, V., H. Hong, and E. Tamer (2007): "Estimation and confidence regions for parameter sets in econometric models," *Econometrica*, 75, 1243–1284.

Choi, S., S. Kariv, W. Müller, and D. Silverman (2014): "Who is (more) rational?" *The American Economic Review*, 104, 1518–1550.

Cox, J. C. (2004): "How to identify trust and reciprocity," *Games and Economic Behavior*, 46, 260–281.

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

Dean, M. and D. Martin (2016): "Measuring rationality with the minimum cost of revealed preference violations," *Review of Economics and Statistics*, 98, 524–534.

Deb, R., R. S. Gazzale, and M. J. Kotchen (2014): "Testing motives for charitable giving: A revealed-preference methodology with experimental evidence," *Journal of Public Economics*, 120, 181–192.

Demuynck, T. (2015): "Statistical inference for measures of predictive success," *Theory and Decision*, 79, 689–699.

Diewert, W. E. (1973): "Afriat and revealed preference theory," *The Review of Economic Studies*, 40, 419–425.

Dufwenberg, M., P. Heidhues, G. Kirchsteiger, F. Riedel, and J. Sobel (2011): "Other-regarding preferences in general equilibrium," *The Review of Economic Studies*, 78, 613–639.

Durante, R., L. Putterman, and J. Van der Weele (2014): "Preferences for redistribution and perception of fairness: An experimental study," *Journal of the European Economic Association*, 12, 1059–1086.

Echenique, F., S. Lee, and M. Shum (2011): "The money pump as a measure of revealed preference violations," *Journal of Political Economy*, 119, 1201–1223.

Engellman, D. and M. Strobel (2004): "Inequality aversion, efficiency, and maximin preferences in simple distribution experiments," *The American Economic Review*, 94, 857–869.

Fehr, E., G. Kirchsteiger, and A. Riedl (1998): "Gift exchange and reciprocity in competitive experimental markets," *European Economic Review*, 42, 1–34.

Fehr, E., M. Naef, and K. M. Schmidt (2006): "Inequality aversion, efficiency, and maximin preferences in simple distribution experiments: Comment," *The American economic review*, 96, 1912–1917.

Fehr, E. and K. M. Schmidt (1999): "A theory of fairness, competition, and cooperation," *The quarterly journal of economics*, 114, 817–868.

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

Fong, C. (2001): "Social preferences, self-interest, and the demand for redistribution," *Journal of Public economics*, 82, 225–246.

Harbaugh, W. T., K. Krause, and T. R. Berry (2001): "GARP for kids: On the development of rational choice behavior," *The American Economic Review*, 91, 1539–1545.

Heufer, J. (2014): "Generating random optimising choices," *Computational Economics*, 44, 295–305.

Heufer, J. and P. Hjertstrand (2015): "Consistent subsets: Computationally feasible methods to compute the Houtman–Maks-index," *Economics Letters*, 128, 87–89.

Höchtl, W., R. Sausgruber, and J.-R. Tyran (2012): "Inequality aversion and voting on redistribution," *European economic review*, 56, 1406–1421.

Hoffman, E., K. McCabe, and V. L. Smith (1996): "Social distance and other-regarding behavior in dictator games," *The American Economic Review*, 86, 653–660.

Liu, B., Y. Xia, and P. S. Yu (2000): "Clustering through decision tree construction," in *Proceedings of the ninth international conference on Information and knowledge management*, ACM, 20–29.

Nishimura, H., E. A. Ok, and J. K.-H. Quah (2017): "A Comprehensive Approach to Revealed Preference Theory," *The American Economic Review*, 107, 1239–1263.

Polisson, M., J. K.-H. Quah, and L. Renou (2017): "Revealed preferences over risk and uncertainty," .

Porter, M. and A. Adams (2016): "For love or reward? Characterising preferences for giving to parents in an experimental setting," *The Economic Journal*, 126, 2424–2445.

Richter, M. K. (1966): "Revealed preference theory," *Econometrica: Journal of the Econometric Society*, 635–645.

Samuelson, P. A. (1938): "A note on the pure theory of consumer's behaviour," *Economica*, 5, 61–71.

Selten, R. (1991): "Properties of a measure of predictive success," *Mathematical Social Sciences*, 21, 153–167.

Szabo, G. and A. Szolnoki (2012): "Selfishness, fraternity, and other-regarding preference in spatial evolutionary games," *Journal of theoretical biology*, 299, 81–87.

Tyran, J.-R. and R. Sausgruber (2006): "A little fairness may induce a lot of redistribution in democracy," *European Economic Review*, 50, 469–485.

Varian, H. R. (1982): "The nonparametric approach to demand analysis," *Econometrica: Journal of the Econometric Society*, 945–973.