Journal of Behavioral and Experimental Finance 22 (2019) 214–222

Contents lists available at ScienceDirect

Journal of Behavioral and Experimental Finance

journal homepage: www.elsevier.com/locate/jbef

oTree: The Equality Equivalence Test☆

F. Holzmeister[a,*], R. Kerschbamer[b]

[a] University of Innsbruck, Department of Banking and Finance, Universitätsstraße 15, Innsbruck, Austria
[b] University of Innsbruck, Department of Economics, Universitätsstraße 15, Innsbruck, Austria

ARTICLE INFO

Article history:
Received 27 November 2018
Received in revised form 10 March 2019
Accepted 4 April 2019
Available online 16 April 2019

Keywords:
oTree
Equality Equivalence Test
Distributional preferences

ABSTRACT

Based on a small set of assumptions on preferences, Kerschbamer (2015) introduces a geometric delineation of distributional preferences and a parsimonious, non-parametric identification procedure — the *Equality Equivalence Test* (EET). The assumptions of the test result in a mutually exclusive taxonomy of social preference archetypes, nesting all empirically relevant types identified in the literature. This article presents a ready-to-use software module for use with *oTree* (Chen et al., (2016), which facilitates the implementation of the EET in the laboratory, the field, or online. The app can be straightforwardly configured and parametrized using a single file and can be seamlessly integrated into existing projects. Furthermore, the app features predefined evaluations of subjects' responses and provides a real-time report of the results in the experimenter's dashboard. By this means, the module offers a comprehensive, flexible, and time-saving tool for implementing and conducting the EET in a myriad of configurations determined by the user.

1. Introduction

Large parts of the economic literature are based on the assumption that material self-interest is the only motivation of rational decision makers. However, everyday experience as well as behavioral evidence suggest that concerns for the well-being of others might well affect people's behavior in many economic and social interactions. This, in turn, has lead to a large body of literature on other-regarding preferences, where arguments beyond material self-interest enter the decision makers' utility function. One class among these theories – commonly referred to as distributional or social preferences – assumes that the decision maker's utility does not only depend on the own material payoff, but may also be a function of the material well-being of others.[1] Distributional preferences have been shown to be behaviorally and economically relevant in many different environments—see, e.g., Sobel (2005) and Fehr and Schmidt (2006) for comprehensive surveys.[2]

During the last decades, several methodologies to identify and classify distributional preferences have evolved—see Kerschbamer (2013, 2015) for a thorough review and discussion of the related literature. However, there has neither been reached a consensus on the basic motivations behind social preferences nor on the question how to delineate distributional archetypes. While several studies start out with a given set of social preference types and utilize experimental designs allowing for discriminating between them (see, e.g., Charness and Rabin, 2002; Engelmann and Stobel, 2004; Cabrales et al., 2010; Iriberri and Rey-Biel, 2013), other studies employ test designs that only allow for identification of certain archetypes (see, e.g., Andreoni and Miller, 2002; Fisman et al., 2007). Moreover, many identification procedures

☆ The software application presented in this article is free of charge and licensed under an adapted MIT open source license with a citation requirement. Any use of the software – whether as a whole or in parts – implies the acceptance of the license agreement available in the folder for download. Demo versions of several task variants are available at https://demo-eet.herokuapp.com. Financial Support of the Austrian Science Fund (SFB F63) is greatfully acknowledged. The app is available for download at the website of Felix Holzmeister (www.holzmeister.biz) and the journal's GitHub repository (https://github.com/JBEF/).

* Corresponding author.
E-mail address: felix.holzmeister@uibk.ac.at (F. Holzmeister).
URL: http://www.holzmeister.biz (F. Holzmeister).

[1] Other arguments that may enter a decision maker's utility function are other people's behavior (as in reciprocity models; e.g., Charness and Rabin, 2002; Dufwenberg and Kirchsteiger, 2004), other people's payoff expectations (as in guilt aversion models; e.g., Charness and Dufwenberg, 2006; Battigalli and Dufwenberg, 2007), or other people's concerns for others' well-being (as in type-based models; e.g., Levine, 1998; Andreoni and Bernheim, 2009).

[2] Important distributional preference types discussed in the literature include concerns for relative income (Duesenberry, 1949), altruism (Becker, 1974; Andreoni and Miller, 2002), envy (Bolton, 1991; Kirchsteiger, 1994; Mui, 1995), spitefulness (Levine, 1998), inequity aversion (Fehr and Schmidt, 1999; Bolton and Ockenfels, 2000; Charness and Rabin, 2002), equity aversion (Charness and Rabin, 2002; Fershtman et al., 2012), Rawlsian preferences (Charness and Rabin, 2002), Leontief preferences (Andreoni and Miller, 2002; Fisman et al., 2007), maximin preferences (Engelmann and Stobel, 2004), surplus maximization (Engelmann and Stobel, 2004), egalitarian motives (Dawes et al., 2007; Fehr et al., 2008).

https://doi.org/10.1016/j.jbef.2019.04.001
2214-6350/

rely on a set of strong structural assumptions with respect to a decision maker's utility or motivational function. For example, following up on the pioneering model introduced by Fehr and Schmidt (1999), several studies rely on the assumption of piecewise linearity (e.g., Charness and Rabin, 2002; Cabrales et al., 2010; Iriberri and Rey-Biel, 2013), whereas social value orientation measures – such as the ring test (Liebrand, 1984) and the slider task (Murphy et al., 2011) – presume linear preferences.

Kerschbamer (2015) illustrates that a set of four rather weak and primitive assumptions on the decision maker's preferences results in a mutually exclusive and comprehensive delineation of nine distributional preference types. Furthermore, the same set of assumptions gives rise to an elicitation method – the *Equality Equivalence Test* (EET) – that discriminates between social preference archetypes based on core features of the decision maker's preferences, rather than properties of identification procedures or structural assumptions on the utility function. As a by-product, the test yields a two-dimensional non-metric index of preference intensity.

This article presents a ready-made software implementation of the EET for use with *oTree* (Chen et al., 2016). As an open-source, object-oriented web framework, *oTree* provides a platform-independent environment deployable on any device, including smartphones and tablets, facilitating the implementation of experiments in the laboratory, the field, or online. As a ready-made application, the EET app taps the full potential of *oTree*, implying straightforward setup and usage, seamless integration into existing projects, responsive graphical design, multi-language support, and automated testing. Relevant parameters, properties of the experimental protocol, and the graphical display are configured by specifying a set of predefined variables in a single file. In addition, the app features predefined evaluations of experimental subjects' responses and provides a real-time report of the results in the experimenter's dashboard. By that means, the module offers a comprehensive and time-saving tool for conducting the EET with any arbitrary parametrization in different experimental environments.

The paper is organized as follows: Section 2 briefly summarizes the key assumptions of the EET and the resulting classification of distributional preference archetypes. Section 3 introduces the ready-made EET app for use with *oTree*, detailing how to set up, parameterize, and utilize its features. Section 4 outlines the predefined analysis and data stored. Section 5 illustrates the ready-made implementation of the "admin report", graphically summarizing experimental subjects' responses in real time.

2. The "Equality Equivalence Test"

The *Equality Equivalence Test* (EET), introduced by Kerschbamer (2013, 2015), constitutes an experimental procedure to elicit individual-level distributional preferences and their intensities based on a multiple choice list format. Similar to the certainty equivalent method, frequently applied to elicit individual-level attitudes towards risk (see, e.g., Cohen et al., 1987; Abdellaoui et al., 2011), the EET requires experimental subjects to indicate their preferences in a menu of binary choices, where one of the two options is held constant across the set of decision-making problems. The methodology of the EET stems from a small set of primitive assumptions on the decision maker's preferences, resulting in a well delineated and comprehensive, mutually exclusive distinction between different archetypes of distributional concerns. By this means, the EET serves as a simple but parsimonious tool for eliciting and characterizing experimental subjects' individual-level social preferences in a two-person context.

2.1. Assumptions

Closely following the definitions in Kerschbamer (2015), let the set of feasible income allocations A be the non-negative quadrant of $\mathbb{R}^2$ and let $a = (m, o)$ denote an income allocation that gives a material payoff of m (for 'my') to the decision maker and a material payoff o (for 'other') to the other person. Let $\succcurlyeq$ denote the decision maker's preference relation over income allocations in A.[3]

Assumption 1. The decision maker's preference relation over income allocations is complete and transitive. That is, it holds for $\succcurlyeq$ that (i) for every allocation pair $a, a' \in A$, either $a \succcurlyeq a'$ or $a' \succcurlyeq a$ (or both) and (ii) for every triple $a, a', a'' \in A$, if $a \succcurlyeq a'$ and $a' \succcurlyeq a''$, then $a \succcurlyeq a''$.[4]

Assumption 2. The decision maker's preference relation over income allocations satisfies strict monotonicity in the own material payoff m. That is, when comparing two arbitrary allocations (m, o) and (m', o) in A with o being equal, it must hold that $(m, o) \succ (m', o)$ if and only if $m > m'$ and $(m, o) \sim (m', o)$ if and only if $m = m'$.

Assumption 3. The decision maker's preference relation over income allocations satisfies piecewise monotonicity in the other person's material payoff o. That is, when comparing two arbitrary allocations (m, o) and (m, o') in A with m being equal and $o < o'$, it must hold that the decision maker's preference relation between (m, o) and (m, o') – i.e., whether $\succ$, $\prec$, or $\sim$ holds – is constant whenever $o > m$ (disadvantageous inequality) and constant whenever $m > o'$ (advantageous inequality). Put differently, the decision maker's preference relation over any two income allocations with the same material payoff m but $o \neq o'$ only depends on whether the decision maker is behind or ahead.

Assumption 4. The decision maker's preference relation over income allocations satisfies strict monotonicity in both payoffs along the ray $m = o$. That is, when comparing two arbitrary allocations (m, o) and (m', o') in A with $m = o$ and $m' = o'$, it must hold that $(m, o) \succ (m', o')$ if and only if $m > m'$ and $(m, o) \sim (m', o')$ if and only if $m = m'$.

2.2. Delineation of archetypes

Given Assumptions 1–4, the decision maker's social preference type can be determined by identifying the location of two sections of the indifference curve through some symmetric reference allocation $r = (e, e)$ in the (m, o)-space: (i) the section that passes the domain of *disadvantageous inequality* (the area north-west of the 45° line through $r = (e, e)$, i.e., where $m < o$) and (ii) the section that passes the domain of *advantageous inequality* (the area south-east of the 45° line, i.e., where $m > o$).

In particular, the choice space is divided into the subsets $\{x_1, x_2, x_3\}$ and $\{y_1, y_2, y_3\}$ as depicted in Fig. 1: Given Assumptions 1–4, a decision maker's indifference curve through the

[3] Following standard conventions, $a \succcurlyeq a^*$ is to be read as "allocation a is weakly preferred over allocation a^*". The asymmetric part of $\succcurlyeq$, i.e., "allocation a is strictly preferred over allocation a^*" is defined as $a \succ a^* \iff a \succcurlyeq a^*$ but not $a^* \succcurlyeq a$. The symmetric part of $\succcurlyeq$, i.e., "the decision maker is indifferent between allocation a and allocation a^*" is defined as $a \sim a^* \iff a \succcurlyeq a^*$ and $a^* \succcurlyeq a$.

[4] Note that Assumption 1 in Kerschbamer (2015) includes continuity, which simplifies the representation of assumptions and features of distributional preference archetypes but is not required for the identification of preferences per se (as the boundaries of the indifference sets are derived from the revealed bounds of upper and lower contour sets).

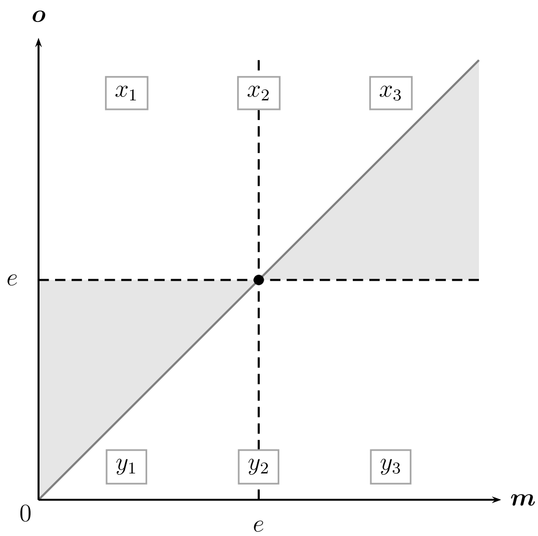


Fig. 1. Delineation of social preference archetypes (based on Kerschbamer, 2015). Assumptions 2 (strict own-payoff-monotonicity) and 4 (strict equal-payoff-monotonicity) imply that a decision maker's indifference curve through the reference allocation $r = (e, e)$ cannot pass either of the two shaded areas.

Table 1
Definition of archetypes of distributional preferences as proposed by Kerschbamer (2013, 2015). $\{x_1, x_2, x_3\}$ and $\{y_1, y_2, y_3\}$ denote the subsets of the choice space in the domain of disadvantages and advantageous inequality, respectively, as illustrated in Fig. 1. $m < o$ indicates allocations where the decision maker is behind (disadvantageous inequality), whereas $m > o$ indicates allocation where the decision maker is ahead of the other person (advantageous inequality). $\partial u/\partial o$ denotes the partial derivative of the decision maker's utility function $u(m, o)$ with respect to the material well-being of the other person. *IC* denotes "indifference curve".

Preference Type	$m < o$	$m > o$	*IC passes...*	
Spiteful (competitive)	$\partial u/\partial o < 0$	$\partial u/\partial o < 0$	x_3	y_1
Kick-down (bully the underlying)	$\partial u/\partial o = 0$	$\partial u/\partial o < 0$	x_2	y_1
Equality averse (anti-egalitarian)	$\partial u/\partial o > 0$	$\partial u/\partial o < 0$	x_1	y_1
Envious (grudging)	$\partial u/\partial o < 0$	$\partial u/\partial o = 0$	x_3	y_2
Selfish (own money maximizing)	$\partial u/\partial o = 0$	$\partial u/\partial o = 0$	x_2	y_2
Kiss-up (crawl to the bigwigs)	$\partial u/\partial o > 0$	$\partial u/\partial o = 0$	x_1	y_2
Inequality averse (egalitarian)	$\partial u/\partial o < 0$	$\partial u/\partial o > 0$	x_3	y_3
Maximin (Rawlsian, Leontief)	$\partial u/\partial o = 0$	$\partial u/\partial o > 0$	x_2	y_3
Altruistic (efficiency loving)	$\partial u/\partial o > 0$	$\partial u/\partial o > 0$	x_1	y_3

reference point $r = (e, e)$ necessarily passes through one (and only one) of the x-subsets and one (and only one) of the y-subsets. Thus, apparently, the set of assumptions results in nine possible constellations of indifference curves, defining nine mutually exclusive archetypes of social preferences. As a byproduct, the EET gives rise to a two-dimensional ordinal index, the (x, y)-score, characterizing both a subject's archetype and the preference intensity (see Section 4). The nine social preference types are listed in Table 1; Fig. 2 showcases typical (piecewise linear) indifference curves associated with the particular types.[5] For a discussion of the core features of the implied distributional preference types and their relation to other definitions of social concerns, see Kerschbamer (2015).

3. Setup and usage of the app

Kerschbamer (2015) marks the start of his paper on the *Equality Equivalence Test* (EET) with a familiar quote:

"Everything should be made as simple as possible, but not one bit simpler". — attributed to Albert Einstein (1879–1955)

The ready-to-use software module for implementing the EET in *oTree* (Chen et al., 2016), illustrated in this article, acts upon the same principle. Similar to the apps presented in Holzmeister and Pfurtscheller (2016) and Holzmeister (2017), offering tools for implementing frequently used risk elicitation methods, the EET app is self-contained and, thus, can be utilized and seamlessly integrated in any experiment conducted with the *oTree* framework. The app facilitates the implementation of the EET as well as a variety of modifications discussed in Kerschbamer (2015). In a user-friendly and straightforward manner, thoroughly documented variables are specified in a single file (`config.py`) at the root of the app's directory. The file `config.py` consists of *oTree*'s `Constants` class and specifies several variables to set parameters and different features of the test. As the app is programmed as a standard *oTree* application, pre-implemented configurations can be modified or extended by custom-designed features without a hitch.

The app can be easily translated into arbitrary languages using Django's i18n internationalization routines. That is, all texts displayed to subjects are tagged in the scripts and templates such that translations are disentangled from the source code. Instructions on how to utilize the translation features are included in the download packages. Similarly, all numbers displayed in monetary units are flagged with currency field tags such that real world or experimental currency denotations can be globally defined in `settings.py`. The segregation of numbers and units facilitates the app's usage in any arbitrary currency denomination without adapting the source code. Moreover, the app features "bots" (via `tests.py`) allowing for automated testing of the particular configuration and parametrization of the test by simulating participant behavior. Further information on how to run tests using command line and browser bots are provided in the download package.

3.1. Parametrization

The parametrization of the EET app utilizes the same denotation as introduced by Kerschbamer (2015). In particular, the menu of binary choices, subjects are exposed to (in the symmetric version of the test), is characterized by four parameters: e, g, s, and t. The parametrization is illustrated in Fig. 3.

(i) The parameter $e \in \mathbb{R}^+$ determines the locus of the reference allocation, i.e. the *equal payoff* allocation $(m, o) = (e, e)$. **(ii)** $g \in \mathbb{R}^+$ constitutes a 'gap' variable, characterizing the *vertical distance* to (e, e), i.e., the 'other' payoff in option 'Left' is equal to $(e + g)$ in the x-list and $(e - g)$ in the y-list. Note that g should be restricted to be strictly smaller than e to rule out negative or zero monetary payoffs. **(iii)** The parameter $s \in \mathbb{R}^+$ is a 'step size' variable, characterizing the *horizontal distance* between two adjacent (m, o)-allocations in both the x- and the y-list; i.e., the 'my' payoff varies in steps of size s around the locus payoff e. **(iv)** $t \in \mathbb{N}^+$ acts as a 'test size' variable, determining the *number of steps* (of size s) to be made to the left and to the right starting from the allocation $(m, o) = (e, e + g)$ in the x-list and the allocation $(m, o) = (e, e - g)$ in the y-list, respectively. To preserve disadvantageous and advantageous inequality within the x- and the y-list, t is restricted to be smaller or equal to g/s.

The parametrization outlined above gives rise to $2t + 1$ binary choice problems in the domain of disadvantageous inequality (x-list) and advantageous inequality (y-list), respectively, i.e., $4t + 2$

[5] It is important to note that the EET does not assume piecewise linearity. The assumption is only used in the figure to simplify and standardize the graphical representation of indifference curves.

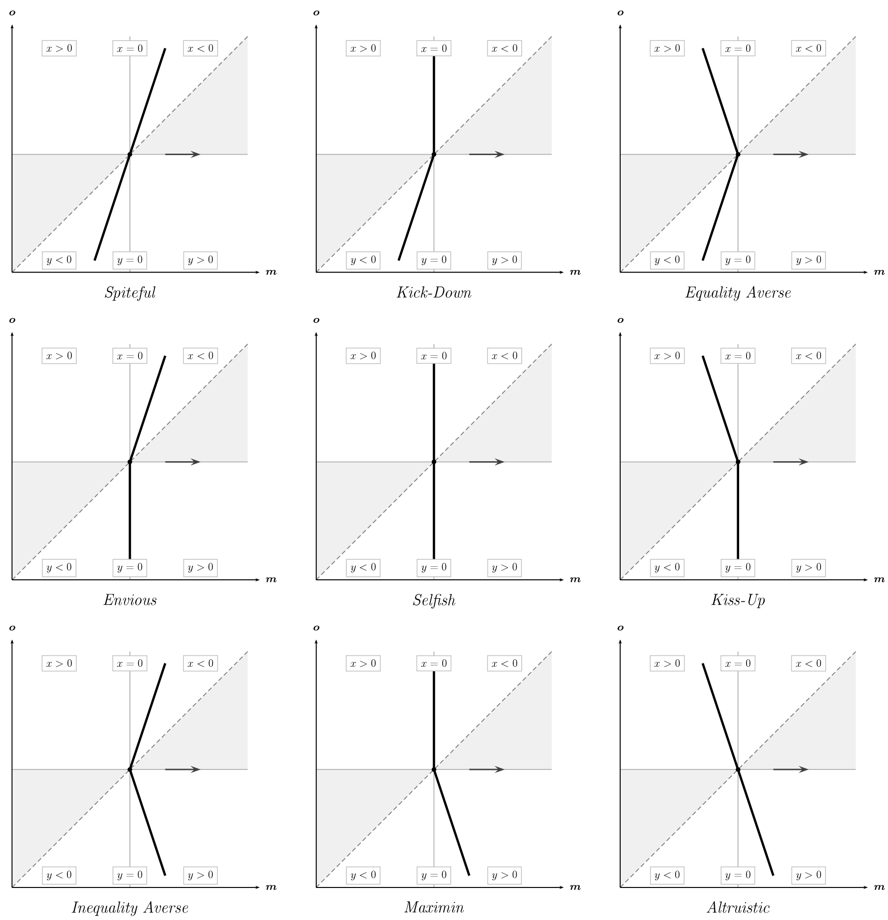


Fig. 2. Typical indifference curves of the nine distributional preference types identified by the EET (based on Kerschbamer, 2015). $\{x_1, x_2, x_3\}$ and $\{y_1, y_2, y_3\}$ denote the subsets of the choice space in the domain of disadvantages and advantageous inequality, respectively. Assumptions 2 (strict own-payoff-monotonicity) and 4 (strict equal-payoff-monotonicity) imply that a decision maker's indifference curve through the reference allocation $r = (e, e)$ cannot pass either of the two shaded areas. Arrows $\rightarrow$ indicate the locus of upper contour sets.

decision in total.[6] For each decision problem, the subject is asked to choose between two alternatives, 'Left' and 'Right', each corresponding to an allocation (m, o)—i.e., one monetary payoff for the decision maker and one for a randomly matched, anonymous subject in the experimental session. The payoff allocations for alternative 'Left' are constructed as depicted in Fig. 3, implying that the monetary payoff of the 'other' person is held constant within the x- and the y-list, whereas the monetary payoff of the decision maker increases monotonically in steps of s. Alternative 'Right' is held constant across all $4t + 2$ choices and offers the equal payoff allocation $(m, o) = (e, e)$.[7] By construction of the test, the assumption of strict m-monotonicity (Assumption 2) is sufficient to ensure that a decision maker would change at most once from alternative 'Right' to alternative 'Left' in either of the two lists, giving rise to the geometric delineation of the nine social preference archetypes outlined in Table 1 and Fig. 2.

Note that the EET, by construction, allows for discrimination between the different types at any arbitrary precision. In particular, it is up to the researcher – by utilizing a suitable parametrization – to define when a decision maker should be considered being individualistic in either of the two domains. Suppose, we define a decision maker to be individualistic if unwilling to give up €k in order to increase the other player's payoff by €1; the parameters in the EET would then need to be set such that $k = s/g \iff s = kg$, whereas e and t can be freely specified.

3.2. Extensions, refinements, and modifications

Besides the symmetric basic implementation of the EET, the app features the extensions and refinements discussed in Kerschbamer (2013, 2015), which might be of interest to address specific research questions and/or gather more fine-grained data to increase the power to discriminate between distributional concerns. In particular, the EET app features implementations of the test with (i) *asymmetric step sizes*, (ii) *asymmetric test sizes*, and (iii) *multiple x- and y-lists*. For a more thorough discussion of the proposed modifications, refer to Kerschbamer (2013).

While the basic version of the EET implies symmetric step sizes of size s, the asymmetric-step-size version of the EET `asymmetric_s = True` modifies the horizontal distances between

6 Note that the minimum test size $t = 1$ implies six binary decisions, three in the x-list and three in the y-list. However, as shown in Kerschbamer (2015), four binary choice problems are actually sufficient to determine the decision maker's social preference type. The EET app features a reduced form version with only four binary choices, implementable by setting `reduced_form = True` in `config.py`. While the parameter t gets obsolete in this configuration, monetary allocations – excluding $(e, e+g)$ and $(e, e-g)$ – are still constructed based on the specified parameter e, g, and s.

7 The app applies a random lottery incentive procedure to avoid potential 'wealth' or 'portfolio effects' (see, e.g., Cubitt et al., 1998; Harrison and Ruström, 2008). As theoretically proven by Azrieli et al. (2012), choosing one out of several decision problems at random is the only incentive compatible mechanism assuming statewise monotonicity of revealed preferences. Accordingly, one of the binary choices made by an 'active' player is randomly picked (with equal probability) at the end of the task and played out according to the player's decision. In case of the 'double role' assignment (see Section 3.3), one of each player's decisions is chosen by an independent random draw to determine participants' payments.

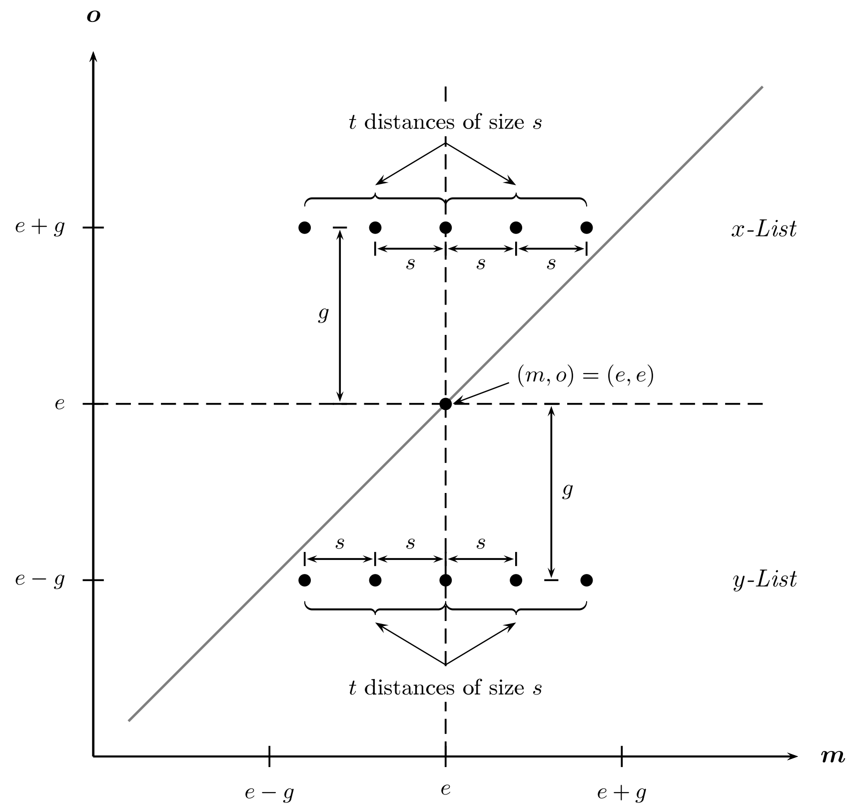


Fig. 3. The geometry of the *Equality Equivalence Test* (based on Kerschbamer, 2015). The x-list refers to binary decisions in the domain of disadvantages inequality ($m < o$); the y-list refers to binary decisions in the domain of advantageous inequality ($m > o$). e determines the locus of the equal material payoff allocation; g is a 'gap' variable characterizing the vertical distance to (e, e); s determines the 'step size' around the locus payoff e; t is a 'test size' variable determining the number of steps (of size s). The EET requires a subject to indicate for each allocation (m, o) in the x- and y-list whether or not he/she prefers it to the equal material payoff allocation $r = (e, e)$.

adjacent allocations in the (m, o)-space in such a way that the step size is small between allocations close to the center but gets larger when moving towards the left and right. By this means, the test's power to discriminate between selfish and non-selfish preferences is increased, without introducing additional binary choices by specifying a higher test size t.[8]

The asymmetric-test-size version of the EET(`asymmetric_t = True`) allows for examining whether a decision maker puts more weight on the material payoff of the other person than on the own material payoff by extending the x-list to the left and the y-list to the right. Implementing this version introduces an additional parameter, a, determining how many choices are added to each of the two lists (starting from the symmetric version based on e, g, s, and t). Note that the x- and the y-list of the asymmetric-test-size version consists of $2t + a + 1$ binary decision problems each, i.e., $4t + 2a + 2$ (rather than $4t + 2$) decision problems in total. As some research questions might call for more precise estimates of the shape of indifference curves in the (m, o)-space, the EET app also allows for implementing a multi-list version of the test. Rather than requiring experimental subjects to complete only one x- and one y-list as in the basic version, `multiple_lists = True` renders an arbitrary number of lists varying in the gap variable g.[9]

3.3. Role assignment and group matching

The EET implies a two-player context with two different roles: the role of an 'active' player (i.e., the decision maker) and the role of an 'inactive' player. Kerschbamer (2015) proposes three different experimental protocols in terms of role assignment: (i) the *fixed role* assignment, (ii) the *role uncertainty* procedure, and (iii) the *double role* assignment. For a thorough discussion of advantages and drawbacks of these assignment protocols, refer to Kerschbamer (2013).

In the fixed role assignment (see, e.g. Cox et al., 2008; Cox and Sadiraj, 2012), roles are assigned *ex ante* (conditional on the configuration of the group matching; see below) and only the 'active' players act as decision makers in the EET's binary choice problems, whereas the 'inactive' players do nothing but receive a payment (based on the matched 'active' player's decision). With role uncertainty (see, e.g. Engelmann and Stobel, 2004; Blanco et al., 2011), both players in a group decide in the role of the decision maker, but are only informed *ex post* which player's decision is relevant for payments (determined by a random draw). Similar to the role uncertainty procedure, both players act as decision makers in the double role assignment (see, e.g. Andreoni

8 By default, setting `asymmetric_s = True` in `config.py` will result in asymmetric steps by determining the decision maker's payoffs as follows: $m_i = e + sgn(i-t) \cdot s \cdot t \cdot (2^{|i-t|}-1)/(2^t-1)$ for $i \in \{0, 1, \ldots, 2t\}$. This functional implies that the 'my' payoff of the outer left, the outer right, and the equal payoff allocation remain identical to the material payoffs of the symmetric basic version of the test, whereas the differences between 'my' payoffs in allocations in between double when moving to adjacent points on the left and right starting from the locus allocation. For example, the setting $e = 10$, $s = 1.5$, and $t = 4$ would result in the following 'my' payoffs: {4.00, 7.20, 8.80, 9.60, 10.00, 10.40, 11.20, 12.80, 16.00}. However, the default implementation can be replaced by an arbitrary function generating a list of 'my' payoffs in `models.py`.

9 To enhance flexibility, an additional variable, `mutliple_g`, takes an arbitrary list as its argument. The elements of this list determine the different gap size parameters to be implemented for the multi-list version of the EET. The number of x- and y-lists in the test, thus, is defined by the number of elements in `mutliple_g`.

and Miller, 2002; Fisman et al., 2007); however, at the same time, both players act as 'passive' players too. Thus, each subject receives two payoffs – one as the 'active', and one as the 'passive' player – in the double role assignment.[10]

The EET app features three different matching protocols: pairs of players may be formed based on either a fixed or a random procedure, or the matching may depend on participants' arrival time on the server. The latter might be particularly useful for online implementations of the EET, for which players do not access the software simultaneously.

3.4. Overall settings and appearance

Typically, the menu of binary decision problems in the EET is presented to subjects in ordered lists, similar to multiple price list formats commonly used in risk preference elicitation methods. In the EET app, the boolean variable `one_page` controls whether the x- and the y-lists are rendered on a single screen or a separate screen each, respectively; the boolean variable `counterbalance` determines whether the ordering of the two lists is identical for all participants, with the x-list being displayed before the y-list, or whether the two lists are randomly counterbalanced.

A typical concern associated with multiple price list formats is the 'compromise effect' (see, e.g., Harrison and Ruström, 2008; Beauchamp et al., 2016): subjects' tendency to anchor towards the middle of a list. Presenting the menu of binary choice problems one-at-a-time on separate screens (`one_choice_per_page`) rather than in form of ordered lists may serve as a remedy for anchoring effects. Another means to mitigate the compromise effect is to randomize the order in which decision problems are displayed (`shuffle_lists`), which might be used with either the tabular form display or the one-at-a-time presentation of choices. A potential drawback of shuffled and/or one-choice-per-page versions of the test, however, is that they might induce a higher frequency of choices violating the monotonicity assumption (see, e.g., Chakravarti et al., 2002). To mitigate this problem, subjects might be presented with an ordered list of the choices they have made earlier in the shuffled and/or the one-choice-per-page version of the test (`revise_decision`), offering them an opportunity to rethink and revise the decisions previously made (as, e.g., in Hedegaard et al., 2018).

Violations of monotonicity also occur in the ordered-lists version of the test—see Kerschbamer (2015). A technical means to preclude reversals in revealed preferences, as suggested by Andersen et al. (2006), is to enforce at most one switching point in the menu of binary decision problems in the x- and the y-list, respectively. Setting `enforce_consistency = True` implies that all options 'Left' below a selected option 'Left' and all options 'Right' above some selected option 'Right' are checked automatically, imposing strict monotonicity of revealed preferences. In this version of the test subjects practically make a single decision on each list: they choose at which decision problem they want to switch from 'Right' to 'Left' (see Kerschbamer and Müller, 2017, for an application)

In addition to the task-specific configurations outlined above, two boolean variables (`instructions` and `results`) control whether or not to display a separate `html`-template for instructions and a summary of the results, respectively.[11] The result screen resembles the decision a subject made in the choice which has been randomly drawn for payment and explains how the final payoff of both players is derived.

4. Data output

Despite providing a comprehensive tool for implementing the EET in *oTree*, our app offers three different kinds of evaluations of subjects' distributional concerns, to be used for further analysis: (i) (x, y)-scores, (ii) parameter intervals in the piecewise linear utility model, and (iii) willingness to pay intervals. Note that the derivation of all three measures is strictly dependent on subject's consistency. Thus, participants violating the assumption of m-monotonicity, i.e., switching between option 'Right' and option 'Left' more than once, are excluded from the analysis.[12] For all subjects not violating monotonicity, all three measures are automatically determined by the app and stored in the database once all decisions in the x- and the y-list have been submitted.

4.1. (x, y)-Scores

Based on a subject's choices in the menu of binary decision problems, Kerschbamer (2015) introduces an ordinal measure to characterize both the archetype and the intensity of distributional concerns. This is done by translating the switching point in the x- and the y-list into a two-dimensional index: the (x, y)-score. While the x-score characterizes the subject's behavior (and preferences) in the domain of disadvantageous inequality, i.e., the decision problems in the x-list, the y-score summarizes a subject's behavior in the realm of advantageous inequality, i.e., the choices in the y-list. In particular, the x- and the y-score are defined by

$$x = (t + 0.5) - \sum_{i}^{2t+1} \mathbb{I}_{R_i} \tag{1}$$

$$y = -(t + 0.5) + \sum_{i}^{2t+1} \mathbb{I}_{L_i} \tag{2}$$

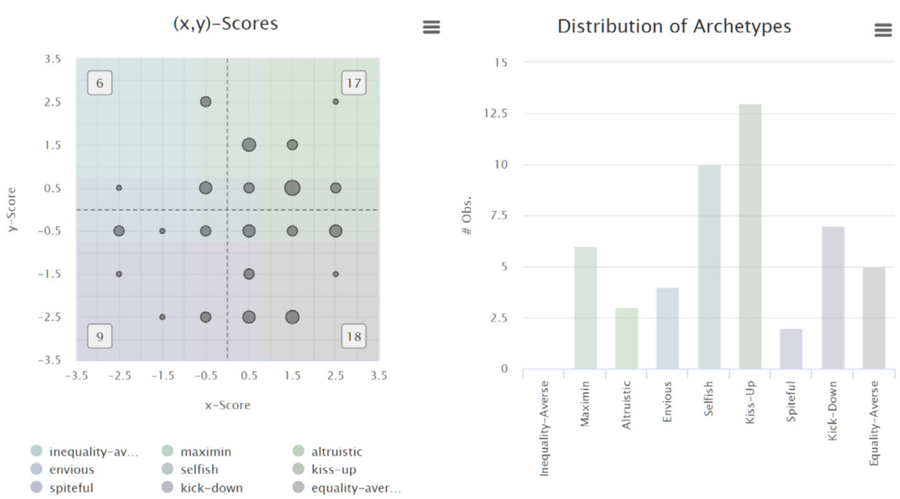
where $\mathbb{I}_{R_i}$ ($\mathbb{I}_{L_i}$) is an indicator variable taking the value 1 if the participant revealed to prefer option 'Right' ('Left') over 'Left' ('Right') for some decision problem $i \in \{1, 2, \ldots, 2t + 1\}$. Note that the definitions above imply that either of the scores can take on $2(t+1)$ different values. By construction, a positive (negative) x-score corresponds to benevolent (malevolent) behavior in the realm of disadvantageous inequality, whereas a positive (negative) y-score corresponds to benevolent (malevolent) behavior in the domain of advantageous inequality. In addition, the magnitude of each of the two scores serves as an ordinal measure of the intensity of distributional concerns in the corresponding inequality domain.[13]

10 Note that other role assignment protocols can easily be implemented by adapting and/or extending the source code in `models.py` and/or `pages.py`. For instance, the app could be modified to a 'single-player' version where all players are assigned the role of the 'active' player while the 'inactive' role is assigned to participants who are absent from the experimental session.

11 Note that the instructions included in the `html`-files in the app's template directory only serve as examples. They do only refer to the default settings in `config.py` and need to be adapted to the particular set of configurations chosen.

12 Note that the problem how to deal with multiple switching behavior in the analysis is the same as for risk preference elicitation methods based on multiple price list formats. Thus, the expedients applied in the realm of risk preference elicitation time and again might be considered for the EET as well: (i) inconsistent choices may be excluded from the data analysis, (ii) the overall number of 'Right' choices may be used as an preference indicator, irrespective of multiple switching behavior, or (iii) the mean of the switching points may serve as a proxy for distributional concerns. Alternatively – as discussed previously – a single switching point might be enforced to rule out violations of the monotonicity assumption at the outset.

13 Note that the (x, y)-score can only be reasonably determined for single-list versions of the EET, but not multi-list versions. For multi-list versions, each list would give rise to a separate (x, y)-score which cannot be straightforwardly aggregated as they are based on an ordinal measurement.



Downloads ▾ Data ▾ Show/Hide ▾ Search:

ID	σ	γ	wtp^a	wtp^d	x-score	y-score	Archetype
1	None	-0.2500	-0.2000	None	2.50	-0.50	Kiss-Up
2	-0.2500	None	None	-0.2000	-0.50	2.50	Maximin
3	-0.2500	0.1250	0.1429	-0.2000	-0.50	0.50	Selfish
4	-0.2500	None	None	-0.2000	-0.50	-2.50	Kick-Down
5	None	-0.2500	-0.2000	None	2.50	-0.50	Kiss-Up
6	0.1250	-0.2500	-0.2000	0.1429	0.50	-0.50	Selfish
7	-0.2500	-0.2500	-0.2000	-0.2000	-0.50	-0.50	Selfish
8	0.3250	-0.2500	-0.2000	0.4815	1.50	-0.50	Kiss-Up
9	None	-0.2500	-0.2000	None	2.50	-0.50	Kiss-Up
10	-0.2500	None	None	-0.2000	-0.50	2.50	Maximin

Showing 1 to 10 of 50 entries

Previous 1 2 3 4 5 Next

Fig. 4. Screenshots of the bubble chart of (x, y)-scores (top left), the histogram of archetypes (top right), and the summary table containing information on midpoints of piecewise linear utility model parameter intervals, midpoints of willingness to pay intervals, and (x, y)-scores (bottom panel) in the admin report of *oTree*'s admin interface. The "Data" dropdown menu allows to select whether to display lower and upper bounds or midpoints of parameter and willingness to pay intervals, respectively; the "Show/Hide" dropdown menu allows for selecting an arbitrary set of columns in the table. Data points are based on browser bots simulation ($n = 50$) of the default game implementation.

4.2. Parameters in a piecewise linear model

A potential shortcoming of the (x, y)-score as a non-metric measure of preference intensity is the implied lack of comparability of results of studies utilizing different sets of test parameters (as the score is not normalized to e, g, s, and t). As a remedy, the (x, y)-score can be directly translated into lower and upper bounds of parameter intervals in a structured model. The most commonly used functional form in the realm of distributional preference modeling is the piecewise linear model introduced by Fehr and Schmidt (1999) to describe self-centered inequality aversion, and its extension by Charness and Rabin (2002) to allow for other types of distributional concerns. The piecewise linear utility function reads

$$u_{\gamma,\sigma}(m, o) = \begin{cases} (1-\sigma)m + \sigma o & \text{if } m \le o \\ (1-\gamma)m + \gamma o & \text{if } m > o, \end{cases} \quad (3)$$

where the two parameters γ and σ are both assumed to be strictly smaller than 1 (to preserve monotonicity). Thus, a decision maker's utility is described as a linear combination of the own and the other player's material payoff, where the weight put on the other's payoff might depend on whether the decision maker is behind or ahead. Apparently, $\sigma = 0$, $\sigma > 0$, and $\sigma < 0$ corresponds to individualistic, benevolent, and malevolent behavior in the domain of disadvantages inequality, whereas, correspondingly, $\gamma = 0$, $\gamma > 0$, and $\gamma < 0$ corresponds to individualistic, benevolent, and malevolent behavior in the domain of advantages inequality. Note, however, that the elicitation procedure does only allow for obtaining estimates for the lower and upper bounds of σ and γ, respectively, but not point estimates. For a thorough discussion on how the choices of a subject with preferences characterized by (3) translate into parameter intervals of the model and the implied relationship to (x, y)-scores, refer to Kerschbamer (2015).[14]

14 For multi-list versions of the test, estimates for the lower and upper bound of σ and γ are obtained by regressing the decision maker's payoffs on the other person's payoff in the allocations determining the decision maker's point of indifference, forcing the regression line through the equal payoff allocation $(m, o) = (e, e)$. In particular, the lower (upper) bound estimate of σ and γ are the coefficients of a ordinary least squares regression (without a constant) of the

4.3. Willingness to pay

As an alternative to parameters in the piecewise linear model, distributional preference intensities can be expressed in terms of the decision maker's willingness to pay for an increase or decrease of the other person's material payoff (wtp) in both the domain of disadvantageous inequality (wtp_d) and the domain of advantageous inequality (wtp_a).[15] The piecewise linear model characterized by (3) implies that the decision maker's willingness to pay, defined as

$$wtp = \frac{\partial_o u(m,o)}{\partial_m u(m,o)} \tag{4}$$

—where $\partial_o u(m,o)$ and $\partial_m u(m,o)$ denote the partial derivatives of $u_{\gamma,\sigma}(m,o)$ with respect to o and m, respectively—is piecewise constant. If $\sigma \geq 0$ ($\gamma \geq 0$), $wtp_d = \sigma(1-\sigma)^{-1}$ ($wtp_a = \gamma(1-\gamma)^{-1}$) gives the amount in terms of the 'own' material payoff the decision maker is willing to give up in the domain of disadvantageous inequality (advantageous inequality) in order to *increase* the other player's material payoff by one unit; symmetrically, if $\sigma < 0$ ($\gamma < 0$), $wtp_d = -\sigma(1-\sigma)^{-1}$ ($wtp_a = -\gamma(1-\gamma)^{-1}$) gives the amount in terms of the 'own' material payoff the decision maker is willing to give up in the domain of disadvantageous inequality (advantageous inequality) in order to *decrease* the other person's material payoff by one unit. As for the parameter estimates in the piecewise linear model, the decision maker's willingness to pay can only be approximated by means of the interval boundaries, but not point estimates.

5. Admin report

To further enhance usability, the EET app utilizes *oTree*'s 'admin report' to provide real-time graphical and tabular information and analysis on the data gathered during the experimental session. The admin report can be launched via *oTree*'s administrator interface and features a histogram of distributional archetypes, a graphical representation of the distribution of (x, y)-scores, and a table summarizing participants' (x, y)-scores, the parameter intervals in a piecewise-linear utility model, the corresponding willingness to pay intervals, and the associated archetype in a paginated format.

In particular, the admin report features a bubble chart of (x, y)-scores (see top left panel in Fig. 4), a frequency chart of the nine social preference archetypes (see top right panel in Fig. 4), and a summary table of all relevant information (see bottom panel in Fig. 4). Both figures can be downloaded in publishable quality in different file formats (.png, .jpg, .pdf, or .svg) directly using the 'hamburger' button in the top right corner. The summary table allows for filtering by any characters using the 'search' bar, selecting or deselecting any table columns for display, and sorting the data based on any column. The table contents can be downloaded directly in different file formats (.xlsx, .csv, or .pdf), copied to the clipboard, or printed immediately.

m-payoffs in the allocations presented in option 'Left' of the decision maker's first 'Left' (last 'Right') choice on the corresponding o-payoffs for both the x- and y-list separately. However, since multi-list versions of the EET are apparently way more likely to provoke violations of the test's assumptions, they might require more involved econometric models (such as, e.g., finite mixture models) to produce meaningful results. The estimates produced by the software should thus be interpreted carefully for multi-list versions.

15 Note that using the wtp's as measures of distributional concerns in the domain of advantageous and disadvantageous inequality might be more convenient for econometric analyses as they – unlike the parameters from the piecewise linear model – are symmetrically scaled around zero.

References

Abdellaoui, M., Driouchi, A., L'Haridon, O., 2011. Risk aversion elicitation: Reconciling tractability and bias minimization. Theory and Decision 71, 63–80.
Andersen, S., Harrison, G.W., Lau, M.I., Ruström, E.E., 2006. Elicitation using multiple price list formats. Exp. Econ. 9, 383–405.
Andreoni, J., Bernheim, B.D., 2009. Social image and the 50–50 norm: A theoretical and experimental analysis of audience effects. Econometrica 77 (5), 1607–1636.
Andreoni, J., Miller, J., 2002. Giving according to garp: An experimental test of the consistency of preferences for altruism. Econometrica 70 (2), 737–753.
Azrieli, Y., Chambers, C.P., Healy, P.J., 2012. Incentives in experiments: A theoretical analysis. Working Paper..
Battigalli, P., Dufwenberg, M., 2007. Guilt in games. AEA Pap. Proc. 97 (2), 170–176.
Beauchamp, J.P., Benjamin, D.J., Chabris, C.F., Laibson, D.I., 2016. Controlling for the compromise effect debiases estimates of risk preference parameters. Working Paper.
Becker, G.S., 1974. A theory of social interactions. J. Political Economy 82 (6), 1063–1093.
Blanco, M., Engelmann, D., Normann, H.T., 2011. A within-subject analysis of other-regarding preferences. Games Econom. Behav. 72 (2), 321–338.
Bolton, G.E., 1991. A comparative model of bargaining: theory and evidence. Amer. Econ. Rev. 81 (5), 1096–1136.
Bolton, G.E., Ockenfels, A., 2000. A theory of equity, reciprocity, and competition. Am. Econ. Rev. 90 (1), 166–193.
Cabrales, A., Miniaci, R., Piovesan, M., Ponti, G., 2010. Social preferences and strategic uncertainty: an experiment on markets and contracts. Amer. Econ. Rev. 100 (5), 2261–2278.
Chakravarti, D., Krish, R., Paul, P., Srivastava, J., 2002. Partitioned presentation of multicomponent bundle prices: evaluation, choice and underlying processing effects. J. Consum. Psychol. 12 (3), 215–229.
Charness, G., Dufwenberg, M., 2006. Promises and partnership. Econometrica 74 (6), 1579–1601.
Charness, G., Rabin, M., 2002. Understanding social preferences with simple tests. Q. J. Econ. 117 (3), 817–869.
Chen, D.L., Schonger, M., Wickens, C., 2016. Otree. an open-source platform for laboratory, online, and field experiments. J. Behav. Exp. Financ. 9, 88–97.
Cohen, M., Jaffray, J.-Y., Said, A., 1987. Experimental comparison of individual behavior under risk and under uncertainty for gains and for losses. Organ. Behav. Human Decis. Process. 39, 1–22.
Cox, J.C., Sadiraj, V., 2012. Direct tests of individual preferences for efficiency and equity. Econ. Inquiry 50 (4), 920–931.
Cox, J.C., Sadiraj, K., Sadiraj, V., 2008. Implications of trust, fear, and reciprocity for modeling economic behavior. Exp. Econ. 11 (1), 1–24.
Cubitt, R.P., Starmer, C., Sugden, R., 1998. On the validity of the random lottery incentive system. Exp. Econ. 1, 115–131.
Dawes, C.T., Fowler, J.H., Johnson, T., McElreath, R., Smirnv, O., 2007. Egalitarian motives in humans. Nature 446, 794–796.
Duesenberry, J.S., 1949. Income, Saving, and the Theory of Consumer Behavior. Harvard University Press, Cambridge, MA.
Dufwenberg, M., Kirchsteiger, G., 2004. A theory of sequential reciprocity. Games Econom. Behav. 47 (2), 268–298.
Engelmann, D., Stobel, M., 2004. Inequality aversion, efficiency, and maximin preferences in simple distribution experiments. Amer. Econ. Rev. 94 (4), 857–869.
Fehr, E., Bernhard, H., Rockenbach, B., 2008. Egalitarianism in young children. Nature 454, 1079–1083.
Fehr, E., Schmidt, K.M., 1999. A theory of fairness, competition, and cooperation. Q. J. Econ. 114, 817–868.
Fehr, E., Schmidt, K.M., 2006. The economics of fairness, reciprocity and altruism – experimental evidence and new theories. In: 1 of Handbook of the Economics of Giving. Altruism and Reciprocity, pp. 615–691.
Fershtman, C., Gneezy, U., List, J.A., 2012. Equity aversion: Social norms and the desire to be ahead. Am. Econ. J. Microecon. 4 (4), 131–144.
Fisman, R., Kariv, S., Markovits, D., 2007. Individual preferences for giving. Amer. Econ. Rev. 97 (5), 1858–1876.
Harrison, G.W., Ruström, E.E., 2008. Risk aversion in the laboratory. In: Research in Experimental Economics 12. Emerald, Bingley, UK, pp. 41–196.
Hedegaard, M., Kerschbamer, R., Müller, D., Tyran, J.-R., 2018. Distributional Preferences Explain Individual Behavior Across Games and Time. Department of Economics, University of Innsbruck, mimeo.
Holzmeister, F., 2017. Otree: ready-made apps for risk preference elicitation methods. J. Behav. Exp. Financ. 16, 33–38.
Holzmeister, F., Pfurtscheller, A., 2016. Otree: the 'bomb' risk elicitation task. J. Behav. Exp. Financ. 10, 105–108.
Iriberri, N., Rey-Biel, P., 2013. Elicited beliefs and social information in modified dictator games: What do dictators believe other dictators do?. Quantitive Econ. 4 (3), 515–547.

Kerschbamer, R., 2013. The geometry of distributional preferences and a non-parametric identification approach. Working Pap. Econ. Stat. 2013–2025.
Kerschbamer, R., 2015. The geometry of distributional preferences and a non-parametric identification approach: the equality equivalence test. Eur. Econ. Rev. 76, 85–103.
Kerschbamer, R., Müller, D., 2017. Social preferences and political attitudes: an online experiment on a large heterogeneous sample. Working Pap. Econ. Stat. 201, 7–16.
Kirchsteiger, G., 1994. The tole of envy in ultimatum games. J. Econ. Behav. Organ. 25 (3), 373–389.
Levine, D.K., 1998. Modeling altruism and spitefulness in experiments. Rev. Econ. Dyn. 1, 593–622.
Liebrand, W.B.G., 1984. The effect of social motives, communication and group size on behaviour in an n-person multi-stage mixedmotive game. Eur. J. Soc. Psychol. 14 (3), 239–264.
Mui, V.-L., 1995. The economics of envy. J. Econ. Behav. Organ. 26 (3), 311–336.
Murphy, R.O., Ackermann, K.A., Handgraaf, M.J.J., 2011. Measuring social value orientation. Judgment Decis. Mak. 6 (8), 771–781.
Sobel, J., 2005. Interdependent preferences and reciprocity. J. Econ. Lit. 43 (2), 392–436.