Discussion Papers
Department of Economics
University of Copenhagen

No. 19-06

Distributional Preferences Explain Individual Behavior Across Games and Time

Morten Hedegaard, Rudolf Kerschbamer, Daniel Müller and Jean-Robert Tyran

Øster Farimagsgade 5, Building 26, DK-1353 Copenhagen K., Denmark
Tel.: +45 35 32 30 01 – Fax: +45 35 32 30 00
http://www.econ.ku.dk

ISSN: 1601-2461 (E)

Distributional Preferences Explain Individual Behavior Across Games and Time

Morten Hedegaard, Rudolf Kerschbamer, Daniel Müller and Jean-Robert Tyran*

May 9, 2019

Abstract

We use a large and heterogeneous sample of the Danish population to investigate the importance of distributional preferences for behavior in a public good game and a trust game. We find robust evidence for the significant explanatory power of distributional preferences. In fact, compared to twenty-one covariates, distributional preferences turn out to be the single most important predictor of behavior. Specifically, subjects who reveal benevolence in the domain of advantageous inequality contribute more to the public good and are more likely to pick the trustworthy action in the trust game than other subjects. Since the experiments were spread out more than one year, our results suggest that there is a component of distributional preferences that is stable across games and over time.

Keywords: Distributional preferences, social preferences, Equality-Equivalence Test, representative online experiment, trust game, public goods game, dictator game.
JEL classification: C72, C91, D64.

*Hedegaard: University of Copenhagen, e-mail: morten.hedegaard@gmail.com. Kerschbamer: University of Innsbruck, e-mail: rudolf.kerschbamer@uibk.ac.at. Müller: University of Innsbruck, e-mail: daniel.mueller@uibk.ac.at. Tyran: University of Vienna and University of Copenhagen, e-mail: jean-robert.tyran@univie.ac.at. We benefited from discussion with Ingvild Almås, Björn Bartling, Yves Breitmoser, Adrian Bruhin, Anna Dreber Almenberg, Ernst Fehr, Ben Greiner, Michael Pfaffermayr, Georg Sator and Roman Sheremeta as well as from comments at the ESA in Berlin and the 2019 Thurgau Experimental Meeting. We thank Erik Wengström and the rest of the iLEE team for providing data from the first wave of experiments in the iLEE as well as Eva Gregersen, Nikolaos Korfiatis and Thomas Alexander Stephens for their support in conducting the experiment. We gratefully acknowledge generous financial support from the Carlsberg Foundation and from the Austrian Science Fund (FWF) through special research area grant SFB F63, as well as through grant numbers P22669, P26901, P27912 and I2027-G16.

1 Introduction

While standard economic theory typically assumes that agents care solely about their own material payoff, there is by now ample evidence that the payoff of other people matters to decision makers as well. This finding has important implications for both economic theory and policy. For example, to evaluate the acceptance of tax policy, distributional preferences have to be taken into account. The emerging empirical evidence led to the development of new models of social preferences that aim at improving the predictive power of standard economic theory.[1] These models have subsequently become highly influential. While in general there is mounting evidence that distributional preferences matter in specific contexts, less is known about their predictive power across games and their stability across time. The current paper sheds new light on this open question.

In this paper, we elicit distributional preferences using the Equality-Equivalence Test (EET; Kerschbamer, 2015) in a large and heterogeneous sample of the Danish population. The experiment is conducted online using the Internet Laboratory for Experimental Economics (iLEE) based at the University of Copenhagen. In this panel, participants take part in several online experiments in four different waves. We exploit this rich source of experimental and survey data to make two contributions to the literature.

First, we investigate the predictive power of distributional preferences for behavior in two games – a linear public goods game (PGG) and a binary trust game (TG). We use the information on distributional preferences (and incentivized beliefs about other participants' contributions), to derive point predictions for individual contributions in the PGG and for second mover behavior in the TG. We then show that i) actual behavior is in line with point predictions and ii) subjects classified as benevolent contribute more in the PGG and are more likely to pick the trustworthy option in the TG, even after controlling for detailed measures of socio-economics, personality, cognitive ability and attitudes. A dominance analysis (Azen and Budescu, 2003) shows that distributional preferences are the single most important predictor of behavior across games. Our results highlight that taking distributional preferences into account improves the predictive power of economic theory.

Second, we provide evidence on the distribution of social preferences in the Danish population and hence contribute to the discussion on the heterogeneity of these preferences. We document that the empirically most frequent preference type in our sample is (with roughly a third of the population) altruistic. Subjects are classified as altruistic if they are willing to give up own income to increase another person's income both when their income is higher and when it is lower than that of another person. Around a quarter of subjects (23 percent) act in a way that is consistent with inequality aversion – they reveal benevolence when ahead and malevolence when behind; a fifth (20 percent) behaves in a selfish manner; and 14 percent are classified as having maximin preferences – they reveal benevolence when ahead and neutrality when behind. In total, these four types make up 90 percent

[1]See for example Fehr and Schmidt (1999), Bolton and Ockenfels (2000), Fehr and Fischbacher (2002) and Charness and Rabin (2002). We use the terms "distributional" and "social" preferences interchangeably. Distributional preferences explain for instance bargaining behavior (De Bruyn and Bolton, 2008), donations to charities (Derin-Güre and Uler, 2010; Kamas and Preston, 2015), voting decisions (Tyran and Sausgruber, 2006; Höchtl, Sausgruber, and Tyran, 2012; Paetzel, Sausgruber, and Traub, 2014; Fisman, Jakiela, and Kariv, 2017; Kerschbamer and Müller, 2017), as well as competitive behavior (Balafoutas, Kerschbamer, and Sutter, 2012).

of our sample. Thus, while the EET provides a comprehensive framework with nine social preference types, only four of these are empirically relevant in our sample.[2]

All our empirical tests for the explanatory power of distributional preferences follow the same general structure: We first derive individual-level predictions from elicited preferences (and the beliefs about the contributions of others in case of the PGG) using the framework à la Charness and Rabin (2002) and Fehr and Schmidt (1999) to calculate the utility of the decision maker for each option. In addition, we derive necessary conditions for the choices of subjects to depart from the selfish benchmark. In particular, we find that benevolence in the domain of advantageous inequality is a necessary (but not sufficient) condition for a positive contribution in the PGG and the trustworthy choice in the TG. We then show that the three distributional types that are benevolent when ahead (maximin, altruistic and inequality-averse subjects) are indeed significantly more likely to behave in this way than all other types. This result holds even after controlling for a battery of socio-demographics, personality and cognitive characteristics, and individual attitudes.

We make these two advances by using state-of-the-art experimental methodology and high-quality empirical data. Concerning methodology, we use the EET which delivers a parsimonious, nonparametric, comprehensive and mutually exclusive classification of individuals into distributional preference types. Intuitively speaking, the test elicits the slope of an indifference curve when trading off income for oneself versus income for another person. The EET delivers two measures of preference intensity – the x-score and the y-score – which can easily be mapped into the two parameters of a piecewise-linear utility function à la Charness and Rabin (2002) or Fehr and Schmidt (1999). This mapping allows us for the first time to calculate individual-level predictions of behavior across games and hence, unlike existing research, we are able to make precise predictions across games. Moreover, the EET allows us to elicit the benevolence of the decision maker in the domain of advantageous as well as disadvantageous inequality in a straightforward manner in one experimental framework. We consider this property a distinct advantage relative to previous studies as in the EET preferences are unlikely to be contaminated by strategic motives such as reciprocity. Moreover, our empirical implementation of the EET delivers a credible measure of confusion (more than one switching point in the X- or the Y-list) that most existing studies do not deliver. We conduct several robustness checks to ensure that our results are not driven by errors in decision making.[3] In particular, we estimate a finite-mixture model of the four most prevalent types and use posterior probabilities to classify the inconsistent participants into their most likely types. Our conclusions remain unchallenged by this exercise.

Overall, our findings demonstrate that distributional preferences matter for behavior in experimental games and that taking them into account is important to improve the empirical realism of economic models. The results in this paper contrasts with previous experimental evidence that questioned the predictive power of social preference models (Blanco, Engelmann, and Normann, 2011). Our paper also highlights the advantages of using the EET over a standard dictator game (DG), which has frequently been used as a proxy for distributional preferences, in interpreting strategic decision

[2] This finding resonates well with that of Kerschbamer and Müller (2017) who reach similar conclusions in a sample of the German population. However, they find a larger proportion of inequality-averse subjects than in Denmark. This raises intriguing questions about the origins and international differences of social preferences.

[3] Andersson, Holm, Tyran, and Wengström (2016) for example find evidence that errors in decision making can lead to a spurious correlation between cognitive ability and risk preferences.

making. The reason is that behavior in the games studied here does not correlate well with behavior in the DG, see Appendix A.2 for details, but does correlate well with the EET.

The paper is organized as follows. Section 2 discusses related literature. Section 3 provides a short introduction to the EET and explains the experimental design and the iLEE in Denmark in detail. Section 4 discusses the distribution of social preferences in Denmark. Sections 5 and 6 present the evidence for the predictive power of distributional preferences for behavior in the PGG and the TG, respectively. Section 7 concludes. In the appendix, we present additional descriptive statistics, several robustness checks including a finite-mixture model, and a detailed description of the experiment including instructions.

2 Related Literature

Our paper contributes to an ongoing debate on the relevance of social preferences for behavior in experimental games. In general, it is fair to say that the literature has not yet reached a clear verdict on this question.

One of the most prominent contributions is Blanco, Engelmann, and Normann (2011). The authors study behavior in four games – an ultimatum game, a modified dictator game, a sequential prisoner's dilemma game and a public goods game – with the aim of testing the Fehr and Schmidt (1999) model of inequality aversion.[4] They use responder data from the UG to estimate aversion to disadvantageous inequality and the modified dictator game to estimate aversion to advantageous inequality. The resulting measures are used to predict decisions in the other two games. The authors find that the Fehr and Schmidt (1999) model has considerable predictive power at the aggregate level but performs less well at the individual level.

There are other studies that reach similar conclusions to Blanco et al. (2011). Engelmann and Strobel (2010) focus on the predictive power of inequality aversion for behavior in the moonlighting game and do not find any significant correlations. Yamagishi et al. (2012) find that rejection of offers in the UG is not correlated with behavior in other games, including a standard DG. See also Kümmerli et al. (2010) and Burton-Chellew and West (2013) for similar claims.

Several papers find mixed evidence for the predictive power of social preferences for behavior in experimental games. Teyssier (2012) studies the role of inequity aversion and risk preferences for cooperative behavior in two versions of a PGG. She employs the same method to elicit inequity aversion as Blanco et al. (2011) and finds that inequity aversion explains contributions in a sequential PGG, but not in a simultaneous PGG. Dannenberg et al. (2007) classify subjects into Fehr-Schmidt and non-Fehr-Schmidt types based on their choices in a DG and an UG. On the one hand, they find that the composition of groups based on these social preferences significantly influences contribution behavior in a PGG in the sense that inequality averse subjects contribute more. On the other hand, it turns out that information about the players in one's own group is required to raise contributions, such that "fair" groups contribute more to the common good. Harbaugh and Krause (2000) find mixed evidence for the correlation of behavior in a DG and a repeated PGG with children. In

[4] In the literature review, we use the abbreviations DG, PGG, PDG, UG and TG for the Dictator Game, the Public Good Game, the Prisoners Dilemma Game, the Ultimatum Game and the Trust Game, respectively.

particular, there is a correlation between DG behavior and behavior in the first round of the PGG in the expected direction, but no strong correlation to behavior in the last round of the PGG. Finally, Dreber, Fudenberg, and Rand (2014) examine whether giving in a standard DG explains cooperation in a repeated PD. They find evidence for a correlation when no equilibrium involving cooperation exists, but not when cooperation is an equilibrium.

Several studies have found evidence that distributional preferences predict behavior in games. Most closely related to our work is Bruhin, Fehr, and Schunk (2018) who estimate a mixture model of social preference types. They find three preference types in a student sample: strong altruists, moderate altruists and a "behindness averse" type. Their model includes distributional as well as reciprocal concerns. In addition to classifying subjects into types based on the posterior probabilities from the mixture model, they show that the structural parameters from the mixture model predict behavior in a TG and a 'reward and punishment game'. Kamas and Preston (2012) elicit behavior in a DG, an UG and a TG and conclude that their data offers "strong support" for social preferences to matter across games. Yang, Onderstal, and Schram (2016) elicit the two parameters of the Fehr-Schmidt model at the individual-level and find that these parameters matter in explaining choices in a 'production game'. Peysakhovich, Nowak, and Rand (2014) find "strong evidence for a cooperative phenotype", that is, for a correlation of pro-social behavior across five different games, including a DG. Offerman, Sonnemans, and Schram (1996) and Murphy and Ackermann (2017) show that subjects' social value orientation predicts cooperativeness in a PGG, see also Yamagishi et al. (2013). Hernandez-Lagos, Minor, and Sisak (2017) find that social preferences predict effort provision and coordination in a lab experiment. Gächter, Nosenzo, and Sefton (2013) show that Fehr-Schmidt preferences are better able to explain peer effects in a three-person UG than social norms. Holm and Danielson (2005) find that behavior in the DG is significantly related to behavior in the TG in Tanzania and in Sweden.

We shed new light on these mixed findings and make several contributions to the literature. First, we use individual-level measures of distributional preferences to make point predictions of behavior in other games which allows for a sharper test. Second, our results demonstrate that there is a component to distributional preferences that is stable over longer periods of time than the previous literature. This is so because our data come from games that were implemented more than a year apart. Third, we demonstrate the predictive power of social preferences in a representative sample instead of a convenience sample of students. Our findings therefore suggest that the importance of social preferences is more general than previously thought. Forth, we demonstrate that distributional preferences are the most important predictor of behavior relative to a large set of potentially relevant covariates. This demonstration is particular powerful in a heterogeneous sample which exhibits larger variation than convenience samples. Fifth, we find that while the EET predicts well, the standard DG does not. Thus, this finding suggests that the EET is a more appropriate measure of distributional preferences.

3 Experiments in the iLEE and the EET

This section first provides a short introduction to the Equality-Equivalence Test (EET) proposed by Kerschbamer (2015) and then informs about the online experiments conducted in the internet

Laboratory for Experimental Economics (iLEE) that we exploit to gather our data.

3.1 The Equality-Equivalence Test

The EET is a price-list technique that aims at identifying the benevolence, neutrality or malevolence of the decision maker towards an anonymous other subject (the recipient) in two domains of inequality – the domain of advantageous inequality where the decision maker is ahead of the other person, and the domain of disadvantageous inequality where the decision maker is behind. Depending on the revealed benevolence, neutrality or malevolence of the decision maker in the two domains, the decision maker is classified into one of nine social preference types – for instance, as altruistic if the decision maker reveals benevolence towards the recipient in both domains, as inequality averse if the decision maker reveals benevolence in the domain of advantageous and malevolence in the domain of disadvantageous inequality and as selfish if the decision maker reveals neutrality in both domains. See Figure 1 for details.[5]

More specifically, the EET exposes subjects to a number of binary choices between two income distributions (m, o), where m (for "my") stands for the own material payoff of the decision maker while o (for "other") stands for the material payoff of the other person. In each choice problem one of the two alternatives consists of a symmetric reference allocation in which both subjects receive the same material payoffs. In the version of the test we use (this version is displayed in Table 1 and graphically illustrated in Figure 2), the symmetric reference allocation was set to 50 Danish Kroner (Dkr; approximately 7 euros) for each person. The second allocation is always asymmetric. In half of the binary choices (the advantageous inequality block – the *Y-list*) the decision maker gets more than the recipient, in the other half (the disadvantageous inequality block – the *X-list*) the decision maker always gets less. Within each of the two blocks the material payoff of the recipient in the asymmetric allocation is held constant, while the material payoff of the decision maker increases monotonically from one choice to the next.[6] This design feature (together with the fact that the symmetric allocation remains the same in all choices) guarantees that a rational decision maker switches at most once from the symmetric to the asymmetric allocation (and never in the other direction) within each block.[7] As Kerschbamer (2015) shows, the two switching points of a subject can be used to construct a two-dimensional index – the (x, y)-score – representing both archetype and intensity of distributional concerns. A positive (negative) x-score corresponds to benevolence (malevolence) in the domain of disadvantageous inequality, while a positive (negative) y-score corresponds to benevolence (malevolence) in the domain of advantageous inequality. Furthermore, the value of the x-score (y-score, respectively) is an ordinal index of the intensity of distributional preferences in the domain of disadvantageous inequality (advantageous inequality, respectively).

[5] A positively (negatively) sloped indifference curve in a given domain corresponds to malevolence (benevolence) in that domain, while a vertical segment corresponds to neutrality.

[6] We varied the incremental change in m in the asymmetric allocation (the "step size", which is constant in the basic version of the test) so that it is small (2 DKr) close to the reference point but grows larger (up to 10 DKr) when moving away from the reference point. This modification in comparison to the basic version of the test was made to increase the power to discriminate between selfish and non-selfish (that is benevolent or malevolent) behavior without increasing the size or decreasing the discriminatory power at the borders.

[7] The rationality requirements underlying the EET are low. In terms of axioms on preferences the assumptions are ordering (completeness and transitivity) and strict own-money monotonicity – see Kerschbamer (2015) for details.

The EET provides several advantages over alternative approaches to elicit distributional preferences. First, it is derived from a small set of axioms on preferences. Thus, the conditions under which the test holds are well-defined. Second, the same set of assumptions result in a *well-delineated, mutually-exclusive* and *comprehensive* set of distributional types. Thus, the set of distributional types tested for is not ad hoc but rather derived from assumptions about preferences. Third, the test is non-parametric and hence does not rely on any functional form assumption. Fourth, the classification into types is done at the individual-level. It thus is perfectly able to account for individual heterogeneity.

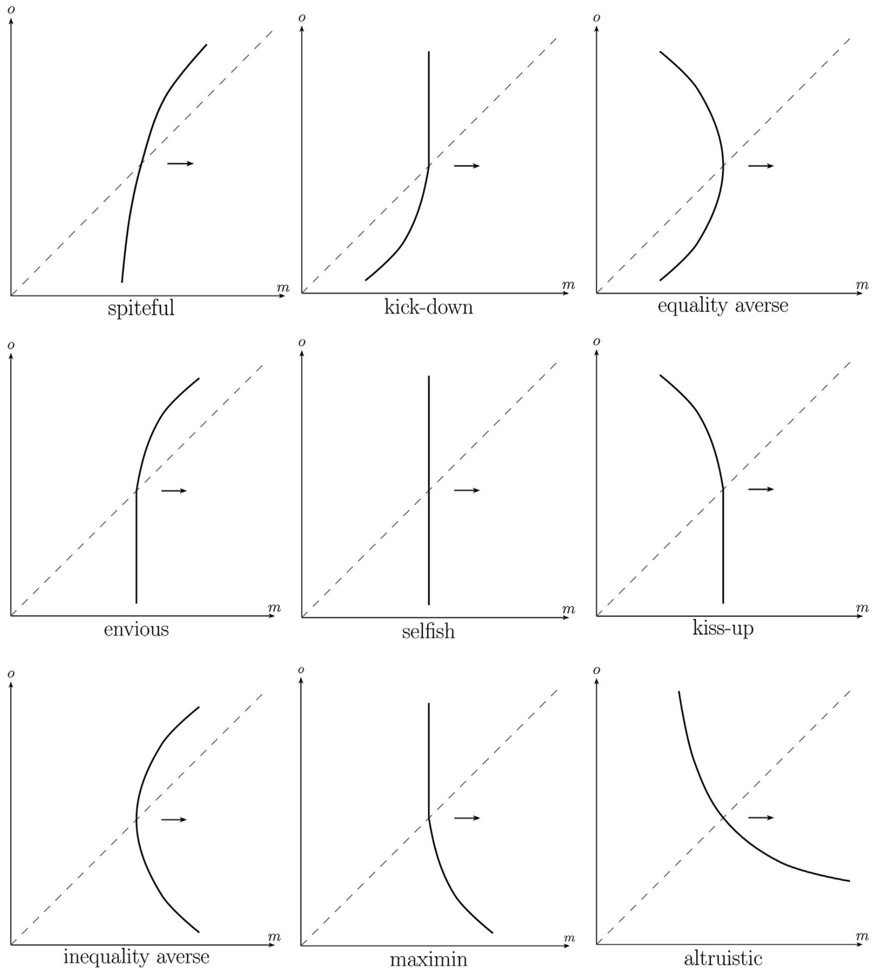

Figure 1: Indifference curves for the nine archetypes of distributional preferences.

The EET was part of wave 3 of the internet Laboratory for Experimental Economics (iLEE). The procedures of the test were as follows. Participants were first explained the rules of the experiment. See Section A.6 in the appendix for experimental instructions and A.7 for screenshots. Choices were made one at the time on separate screens where decision makers choose between Left and Right before moving on to the next choice. Once they have made all 14 choices, subjects saw a confirmation screen. This screen provided an overview of the choices made by the subject in the EET with a horizontal line separating the X- and the Y-list. The chosen distributions were color highlighted and decision makers could go back and change their decisions as many times as they wished. Once they confirmed their decisions, they moved on to the next experiment in the wave.

We employed two conditions that relate to the roles and possible interaction of decision makers and recipients. In the *FixedRoles* condition half of the participants were decision makers, the other half

The X-list				The Y-list			
Left		Right		Left		Right	
m	o	m	o	m	o	m	o
20	75	50	50	42	25	50	50
30	75	50	50	48	25	50	50
42	75	50	50	50	25	50	50
48	75	50	50	52	25	50	50
50	75	50	50	58	25	50	50
52	75	50	50	70	25	50	50
58	75	50	50	80	25	50	50

Table 1: The X- and the Y-list implemented in the iLEE. All numbers in Danish kroner.

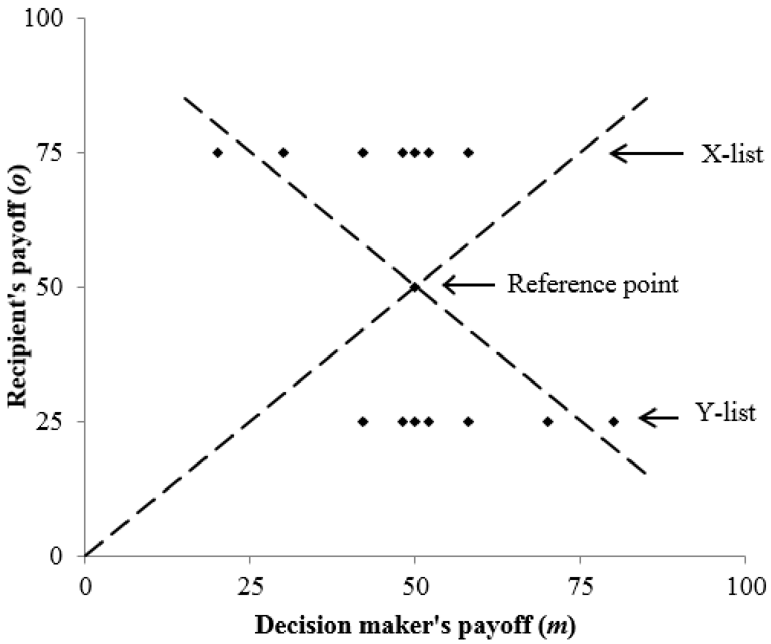

Figure 2: Graphical illustration of the allocations.

were recipients. Roles were randomly assigned and revealed after participants read the instructions but before any decisions were made. Decision makers then made their choices in the EET while recipients made no decisions. At the very end of the experiment, each decision maker was randomly assigned to a recipient and one randomly selected choice was then actually paid out in each pair. In the *RandomRoles* condition all participants made choices as if they were decision makers. A random draw determined ex-post which role each participant was paid. Again, half of the subjects received the decision maker role and half the recipient role, each subject in the decision maker role was randomly assigned to one in the recipient role and one randomly selected choice was then actually paid out. Instructions were kept as similar as possible across conditions and treatment allocation was random with one-third of participants in the *FixedRoles* condition and two-thirds in the *RandomRoles* condition.

3.2 The Internet Laboratory for Experimental Economics

The experiment uses a "virtual lab" approach and is conducted using the platform of the internet Laboratory for Experimental Economics (iLEE) at the University of Copenhagen, Denmark, see Thöni, Tyran, and Wengström (2012). Subjects for the platform are recruited with the assistance of the official statistics agency (Statistics Denmark) who selects a random sample from the general population. The iLEE consists of four different waves, issued between May 2008 and June 2011.[8]

In the linear public good game (PGG) experiment, which was part of the first wave of the iLEE, subjects are endowed with Dkr 50 and decide how much to contribute to a pool of common resources (the public good) and how much to keep for themselves (the private good). Subjects are matched into groups of four. The total amount contributed by the group to the common pool is doubled and shared equally among the group members (the marginal per capita return, MPCR, is hence equal to 0.5). The PGG is played as a one-shot game. While it is socially optimal that all group members contribute the full endowment, individual income is maximized by contributing zero. After the contribution decision, we elicit beliefs about the average contribution of the three other group members incentivized using a quadratic scoring rule.

In the binary trust game (TG), which was part of wave three of iLEE in July 2010, each subject makes two decisions, one in the role of the first mover and one in the role of the second mover. Subjects were informed in the instructions that only one of the two decisions would actually be paid out. For half of the subjects the first-mover decision was selected to be payoff relevant, for the other half the second-mover decision was payoff relevant, and the matching was random and one-to-one. The first mover had to decide between *in* and *out*. *Out* implies payoffs of 50 DKr and 20 Dkr for the first and the second mover. *In* implies that the decision is passed on to the second mover. The second mover then decides between *betrayal* and *honor*, which implements the payoff pair (20,90) or (80,40), respectively. Here, we only consider the decisions of the second mover, as they are clearly distributive in nature. Again, a screenshot can be found in the appendix – see Figure A.7.

Finally, we include three different sets of control variables in our regressions, all taken from the iLEE survey: First, the socio-demographic set consists of age; age squared; a gender dummy; education (coded in four different categories); dummies for employed, retired, student and self-employed status; income (coded in quartiles); and the number of hours working per week. Second, the personality and cognitive controls comprise the IQ score; the score from the cognitive reflection test; and the Big-5 character traits. Third, the set of attitude controls consists of three different variables indicating political preferences and trust. All these variables are explained in more detail in Section A.5.

4 The Distribution of Social Preferences in Denmark

In total, 1067 participants took part in the experiment – with average earnings of 51.8 Dkr. From these 1067 subjects, 885 played the role of a decision maker in the EET, while the rest was only in the role of a recipient (in the *FixedRole* condition) The assumptions of ordering and strict m-monotonicity imply that decision makers switch at most once from Right to Left (and never from Left to Right) in each list.

[8]More detailed information about the iLEE are presented in Section A.5 in the appendix. See also the web page http://www.econ.ku.dk/cee/ilee/description/ for further information.

Of the $n = 885$ decision makers, 650 fulfill this rationality criterion while 235 (27%) make choices that are not consistent with it.[9] In the main analysis we focus on the consistent decision makers. Later on, in the robustness section, we also estimate mixture models and use posterior probabilities to classify inconsistent participants into one of the four main types. Regarding the two payment protocols, we find very little evidence that these two payment protocols cause differences in behavior. In particular, we do not find evidence that the number of consistent subjects or the frequency of distributional types is affected. Appendix A.3 reports more details. In the following, we therefore merge the data without using dummies for the protocols. The results with the dummies are very similar and available upon request.

Table 2 displays the distribution of social preferences types in Denmark. The first column of the table shows that the empirically most frequent preference type in our sample is (with roughly a third of the population) altruism. Subjects are classified as altruistic if they reveal benevolence both when they are ahead and when they are behind. Around a quarter of subjects (23 percent) act in a way that is consistent with inequality aversion – they reveal benevolence when ahead and malevolence when behind; a fifth (20 percent) behaves is a selfish manner – their behavior seems to be unaffected by the material consequences for others, independently of whether they are ahead or behind; and 14 percent are classified as having maximin preferences – they reveal benevolence when ahead and neutrality when behind. In total, these four most prevalent types make up almost 90 percent of our subject population. Of the remaining, less than six percent act in a way that is consistent with envy and less than three percent are spiteful, while kiss-up, equality averse and kick-down each account for only about one percent of the subjects. Thus, while the EET provides a comprehensive framework which allows for the distinction between nine social preference types, only five of these types attract more than 5% of our subjects. Later on we will run regressions using distributional types as independent variables. In those regressions we do not separately include a dummy for types that hold less than 5% of participants (that is, for spiteful, kick-down, equality-averse and kiss-up). Instead, we merge those types into the category *infrequent*. It is important to note that in total less than 5% of the whole sample falls into this merged category. This is less than the 5.5% of subjects we have in the smallest included non-merged category (envious). Finally, Figure 3 plots the distribution of social preference types in our population in the (x, y) space. In this space, the benevolence of the decision maker in the domain of disadvantageous inequality is measured on the x-axis and the benevolence in the domain of advantageous inequality is measured on the y-axis (in both cases negative values mean malevolence). The figure clearly shows that there are pronounced mass points in the top-left corner (inequality-aversion) and in the center (selfishness), and that there is a densely populated area of somewhat smaller mass points in the positive orthant (with maximin covering the left-hand side of the area and altruism covering the rest).

[9] This share is relatively large – compared to the 5% share reported by Kerschbamer (2015) for a standard lab experiment based on a student subject pool, for instance. A possible reason for the large share of inconsistent subjects is the heterogeneity and representativeness of the sample on which our study is based. Evidence in support of this conjecture comes from an earlier wave of the iLEE: Andersson, Holm, Tyran, and Wengström (2016) report that 35% of the sample had multiple switching points in a variation of the Holt and Laury (2002) risk attitudes elicitation procedure – which is similar to the EET with regards to complexity.

Types	Distribution
Altruist	32.2
Inequality averse	23.2
Selfish	20.0
Maximin	13.7
Envious	5.5
Spiteful	2.6
Kiss-up	1.2
Equality averse	1.1
Kick-down	0.5
N	650

Table 2: Distribution of social preference types in percent.

5 Public Good Game

We now turn to the assessment of the predictive power of distributional preferences across different games. We first investigate how distributional preferences explain behavior in the linear PGG. Section 6 then considers the binary TG.

5.1 Prediction for the PGG

To derive predictions for behavior in the PGG, we use the piecewise-linear social utility function proposed by Charness and Rabin (2002). For the two-agents case, the reciprocity-free version of the Charness-Rabin function reads:

$$U(m,o) = \begin{cases} (1-\sigma)m + \sigma o & \text{if } m \leq o \\ (1-\rho)m + \rho o & \text{if } m > o, \end{cases} \tag{1}$$

where m (for *my*) denotes again the income of the decision maker and o (for *other's*) the income of the second person and where σ and ρ are two parameters that determine the weight the decision maker puts on the income of the other person when she is behind ($m \leq o$) or ahead ($m > o$), respectively. In line with Charness and Rabin (2002), we assume in the main text that $\sigma \leq \rho < 1$. The former inequality means that the decision maker is more benevolent (less malevolent) in the domain of advantageous than in the domain of disadvantageous inequality and it guarantees that indifference curves in the (m, o)-space are convex. The latter inequality makes sure that the preferences of the decision maker are monotone in the own material payoff.

To apply the Charness-Rabin function to the four player game under consideration, we follow the proposal by Fehr and Schmidt (1999) and 'normalize' the parameters σ and ρ by dividing them through the number of other players (in our case three). This normalization guarantees that the relative weight of distributional concerns on the decision maker's total payoff is independent of the number of other players. Furthermore, we assume in line with Fehr and Schmidt (1999) that social preferences are 'self-centered' in the sense that the decision maker compares herself to each other player in her group

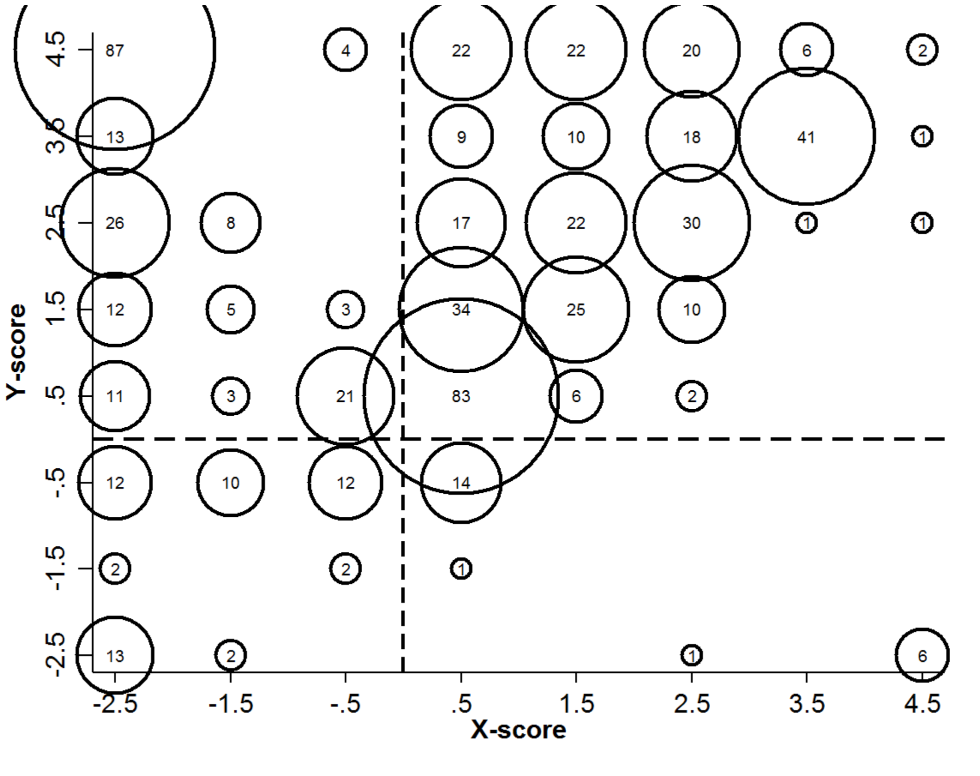

Figure 3: Scatterplot of (x, y) scores.

of four, but does not care about inequalities within the group of others. Finally, we take the elicited belief of the subject about the average contribution of the three other group members as a point belief and we denote this belief by b and the own contribution by c.

Using this notation and taking the budget restriction and the technology of the linear PGG into account, the variables m and o in equation (1) can be written as:

$$m = (E - c) + (c + 3b)\frac{2}{4} = 50 - 0.5c + \frac{3}{2}b \tag{2}$$

$$o = (E - b) + (c + 3b)\frac{2}{4} = 50 + 0.5c + 0.5b. \tag{3}$$

Substituting into the normalized versions of equation (1) and taking into account that $m \leq o \iff c \geq b$ we get utilities of

$$50 - \frac{1}{2}c + \frac{3}{2}b + \sigma(c - b) \tag{4}$$

if $c \geq b$ and

$$50 - \frac{1}{2}c + \frac{3}{2}b + \rho(c - b) \tag{5}$$

if $c < b$.

Given the piecewise linearity of the preferences with a kink at $c = b$ and the linearity of the constraint, each subject has either a unique optimal contribution level at one of the points in $\{0, b, E\}$, or the subject is indifferent among several contribution levels. Specifically, we get the following prediction for the PGG:[10]

[10]The prediction in the main part focuses on the standard case of convex preferences (which translates to the assumption $\rho \geq \sigma$ in the Charness and Rabin framework), as it simplifies the exposition. Subjects with concave distributional

Prediction for the PGG: *Consider the linear PGG with marginal per capita return of one-half. Suppose the decision maker's preferences are of the Charness-Rabin form with parameters* $\rho \geq \sigma$. *Further assume that the decision maker believes that all other group members contribute* b. *Then*

- *if* $\rho \geq \sigma > 0.5$ *then the unique optimal contribution level is at* $c = E$;
- *if* $\rho > \sigma = 0.5$ *then any contribution level in* $[b, E]$ *is optimal;*
- *if* $\rho > 0.5 > \sigma$ *then the unique optimal contribution level is at* $c = b$;
- *if* $\rho = \sigma = 0.5$ *then any contribution level in* $[0, E]$ *is optimal;*
- *if* $\rho = 0.5 > \sigma$ *then any contribution level in* $[0, b]$ *is optimal;*
- *if* $0.5 > \rho \geq \sigma$ *then the unique optimal contribution level is at* $c = 0$.

In our empirical analysis below, we first regress the actual contribution of a subject in the PGG on the predicted value based on the estimated preference parameters. The preference parameters σ and ρ are calculated from the choices in the x- and the y-list for each individual. This is possible because there is a one-to-one relationship between the scores, σ and ρ as well as the willingness-to-pay to increase the other person's income. Tables 9 and 10 in appendix A.1 summarize these relationships. It is important to note that according to the above prediction, a necessary condition for a subject with convex distributional preferences to contribute to the PGG is $\rho \geq 0.5$. A strictly positive ρ means benevolence in the domain of advantage inequality. From the nine distributional types in the EET only altruistic, inequality averse and maximin subjects are benevolent in the domain of advantageous inequality. We will therefore in some regressions work with a dummy indicating that the subject is of one of these three types. It is also important to note that for selfish, envious, spiteful and kick-down subjects $c = 0$ is a dominant strategy. This follows immediately from the fact that these types are non-benevolent in both domains. For equality averse and kiss-up subjects (who have $\rho < \sigma$ by definition) $c = 0$ is a dominant strategy if $\sigma < 0.5 + (\sigma - \rho)b/50$ – which is true for all 15 kiss-up and equality averse subjects in our sample. This is why it also makes sense to classify these two types together with kick-down and spiteful subjects into the *infrequent* category.

5.2 Results: Public Good Game

Figure 4 displays the distribution of actual contributions. There are spikes in contributions at zero, around 50% of the endowment and, most pronounced, at full contribution of 50 kroner. Except for the relatively high level of full contributions observed in this Danish sample, this pattern is quite standard in PGGs.

preferences ($\rho < \sigma$) have either a strict preference for $c = E$ (if $\sigma > 0.5 + (\sigma - \rho)b/50$), or a strict preference for $c = 0$ (if $\sigma < 0.5 + (\sigma - \rho)b/50$), or they are indifferent between the points $c = 0$ and $c = E$ (if the restriction holds as an equality). There are only 47 subjects with strictly concave distributional preferences. For all but one of those subjects the prediction is $c = 0$. For the remaining subject, the prediction is $c = 50$.

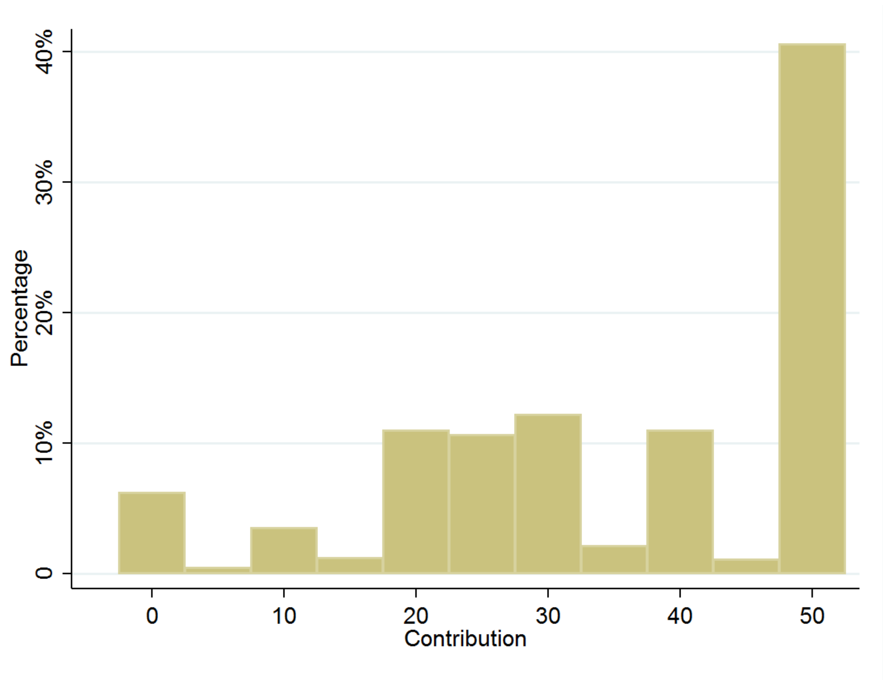

Figure 4: PGG contributions.

Our main results are presented in Table 3 in which we regress the actual individual contributions in the PGG on the predicted contributions, with and without controls.[11] We divide our controls into three distinct sets as: Socio-demographic, personality, cognitive and attitude set of controls.[12] We find that the calculated predictor is highly significant in those regressions, independently of whether we include control variables or not.[13] The size of the coefficient is between 0.15 and 0.2 and it can be interpreted as a partial correlation.

Furthermore, Table 4 displays in columns (1) - (5) results from regressions that include a dummy that indicates inequality averse, maximin as well as altruistic subjects. All other types consequently constitute the reference group. This dummy is significant in the expected direction in all cases. The estimates indicate that subjects who are benevolent in the domain of advantageous inequality contribute, on average, around 4 to 5 tokens more to the public good than other participants. Columns (6) - (10) show that altruists, and to a lesser extent maximin and inequality averse, contribute more

[11]There are however 49 subjects (7.5%) who we predict to be indifferent over all possible contribution levels. Another 14.8% are predicted to be indifferent either between all contributions in $[0, b]$ or between all contributions in $[b, 50]$. Hence, in total, 22.3% of our participants are affected by a theoretical indifference based purely on distributional preferences. We treat this issue in the following way. In the main text, we assign to each indifferent subject as the prediction the highest contribution level that is optimal for this subject. The appendix presents results where we assign the lowest optimal value. Table 17 in the appendix shows that our conclusions remain unaffected.

[12]Specifically, the socio-demographic set includes the variables: age, age squared, gender, education, employed-dummy, retired-dummy, student-dummy, self-employed-dummy, income quartiles and hours worked. The personality and cognitive controls include IQ score, score in the cognitive reflection test and the Big-5 (one variable for each of the five traits). The attitude controls are political left-right assessment, responsibility of the individual versus the government, attitudes toward competition (all three variables coded between one and ten) and the generalized trust question (a binary indicator). Moreover, we always control for the role treatment in the EET and also, whenever possible, for the framing of the PGG.

[13]We show OLS regressions throughout the paper. Two-limit Tobit models (in case of the PGG) and Logit models (in case of the TG) deliver very similar results. Results are available upon request.

than all other types to the public good. In all these regressions, selfish types serve as the reference category. All this holds after controlling for a large battery of covariates. All things considered, these results support the conclusion that distributional preferences matter for behavior in the PGG.[14]

Table 5 shows results of a dominance analysis (Azen and Budescu, 2003), which allows to assess the relative importance of predictors in a multiple regression framework. The table shows the standardized dominance statistic which compares the relative contribution of each predictor to the overall predictive power of the model.[15] We find that the variable *prediction* predicts between almost half – in column (1) – and two-thirds – in column (3) – of the total explained variation across the three subsets of covariates. Finally, the last column of the same table shows that *prediction* still explains almost one-third (27%) of the total explained variation in behavior when pitched against *all three* subsets of covariates together. It is also noteworthy that *prediction* is never dominated by any other of the 21 control variables as predictor of actual behavior in the PGG and always has the highest dominance statistic, i.e., *prediction* predicts better than any other variable across all subsets of models.

Result PGG: *Distributional preferences do affect cooperation in the PGG. Altruistic, inequality averse or maximin subjects contribute more than selfish ones.*

Contribution	(1)	(2)	(3)	(4)	(5)
Prediction	0.208***	0.165***	0.186***	0.176***	0.150***
	(0.03)	(0.04)	(0.03)	(0.03)	(0.04)
Socio-demographics	No	Yes	No	No	Yes
Cognition & Personality	No	No	Yes	No	Yes
Attitudes	No	No	No	Yes	Yes
Observations	650	443	650	603	412
R^2	0.064	0.091	0.105	0.074	0.136

Table 3: Dependent variable is the individual contribution in the PGG. *Prediction* is the predicted contribution of the piecewise linear model. OLS, robust standard errors in brackets. *, ** and *** indicate significance at the 10%, 5% and 1% level, respectively. A constant is included in all cases but not displayed here.

[14]Moreover, we also note that non-parametric Fisher tests – on the correlation between a dummy that indicates a positive (a dummy that indicates full contributions respectively) and the IA-Alt-Maximin dummy – also deliver results significant at the 5%-level. The nonparametric results are not reported here but are available from the authors upon request.

[15]In the case of p regressors, this measure represents the average difference in fit between all $p!$ subsets of models that include a regressor x_i and those that do not (Azen and Budescu, 2003).

Contribution	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
IA-Alt-Maximin	5.077***	4.736***	4.245***	3.932***	4.343**					
	(1.41)	(1.61)	(1.44)	(1.51)	(1.76)					
Altruist						6.841***	5.152**	6.202***	6.191***	5.280**
						(1.88)	(2.09)	(1.87)	(1.98)	(2.26)
Inequality averse						3.652*	2.397	2.685	1.978	1.437
						(1.96)	(2.16)	(1.94)	(2.05)	(2.23)
Maximin						3.946*	3.792	3.621*	3.743*	4.042
						(2.20)	(2.76)	(2.17)	(2.24)	(2.71)
Envious						1.287	-1.839	2.591	2.716	-1.426
						(3.19)	(3.52)	(3.27)	(3.39)	(3.94)
Infrequent						-0.644	-2.755	-1.268	-0.629	-2.746
						(3.21)	(4.11)	(3.26)	(3.45)	(4.42)
Socio-demographics	No	Yes	No	No	Yes	No	Yes	No	No	Yes
Cognition & Personality	No	No	Yes	No	Yes	No	No	Yes	No	Yes
Attitudes	No	No	No	Yes	Yes	No	No	No	Yes	Yes
Observations	650	443	650	603	412	650	443	650	603	412
R^2	0.023	0.070	0.072	0.043	0.119	0.030	0.076	0.081	0.054	0.128

Table 4: Dependent variable is the individual contribution in the PGG. *IA-Alt-Maximin* is a dummy indicating that the subject is either inequality averse, maximin or altruistic. *Infrequent* is a dummy indicating that the subject is either spiteful, kick-down, equality-averse or kiss-up. OLS, robust standard errors in brackets. *, ** and *** indicate significance at the 10%, 5% and 1% level, respectively. A constant is included in all cases but not displayed here.

	Socio-demographics	Cognition & Personality	Attitudes	All Controls
Prediction	0.47	0.53	0.66	0.27
Socio-demographics	0.53	-	-	0.35
Cognition & Personality	-	0.47	-	0.35
Attitudes	-	-	0.34	0.04

Table 5: Table displays standardized dominance statistics (in %). Dependent variable is the individual contribution in the PGG in an OLS regression model. The R^2 serves as the fit-statistic.

6 Trust Game

We now analyze the predictive power of distributional preferences for second-mover behavior in the binary trust game contained in wave 3 of the iLEE. A screenshot of the TG is shown in Figure 6 in the appendix.

6.1 Prediction for the TG

In order to make an individual-level prediction for our binary TG, we again use the Charness and Rabin utility function given in (1). A second mover in the TG faces the decision between *betrayal* and *honor*, implying the allocations $(20, 90)$ and $(80, 40)$, respectively.[16] Inserting these payoffs into (1), we see that the second mover prefers $(80, 40)$ over $(20, 90)$ iff

$$(1 - \sigma)40 + \sigma 80 \geq (1 - \rho)90 + \rho 20, \tag{6}$$

which yields the prediction:

Prediction for the TG: *Consider a binary TG, in which the second mover has the choice between the payoff allocations* $(20, 90)$ *and* $(80, 40)$*. Suppose the second mover's preferences are of the Charness-Rabin form with parameters* ρ *and* σ*. Then*

- *if* $4\sigma + 7\rho > 5$ *the second mover's uniquely optimal move is to pick honor;*
- *if* $4\sigma + 7\rho = 5$ *the second mover is indifferent between betrayal and honor; and*
- *if* $4\sigma + 7\rho < 5$ *the second mover's uniquely optimal move is to pick betrayal.*

In our main analysis we test the above prediction in two ways: First, we use a dummy that indicates a higher utility from the *honor* allocation (*Prediction-honor* dummy) and second we use the actual utility difference between *honor* and *betrayal* (Δ-*honor*) as predictor. Since $\sigma \leq 1$, a necessary, but not sufficient, condition for a decision maker to pick *honor* is $\rho > 0$. Thus, to pick *honor* the decision maker must be benevolent when ahead and consequently be either inequality averse, maximin or altruistic. Based on this observation, we regress in a complementary analysis a dummy indicating

[16] In those vectors, the first (second) entry is the first- (second-) mover payoff in the respective allocation.

whether the subject picked *honor* on the set of types (and covariates), to assess whether subjects who are benevolent in the domain of advantageous inequality are indeed more likely to pick *honor*.

6.2 Results: Trust Game

Table 6 shows that both the *Prediction-honor* dummy, which is equal to one if the individual-level Charness-Rabin utility of *honor* is larger than that of *betrayal*, as well as Δ-*honor*, the actual utility difference, are robust predictors in the expected direction of actual choices of second movers.[17] In addition, Table 7 shows in columns (1) to (5) that the dummy *IA-Alt-Maximin* – again indicating that the subject is either altruistic, inequality averse or maximin – is also significant, confirming our earlier conclusions. Looking at the individual types – again merging spiteful, kiss-up, equality averse and kick-down into an *infrequent* category – we observe in columns (6) to (10) that types that are benevolent when ahead are indeed all more likely to pick *honor*. This holds true whether we include our standard set of controls or not.[18]

Table 8 shows the results from the dominance analysis for the TG. It displays the standardized dominance statistics for the *prediction* variable in each model. In particular, *prediction* contributes between 55% and 73% to the overall model fit compared to the three sets of covariates, see columns (1) - (3). When pitched against all 21 covariates, this variable remains important and contributes with around 36% to a large degree to the total model fit. Similar to the PGG, it is also never dominated by any other of the 21 control variables as predictor of behavior. This exercise corroborates the finding that distributional preferences are the single most important predictor of behavior across games within an extensive set of covariates. In fact, *prediction* is more important than any of the three *sets of covariates.*

Result TG: *Distributional preferences are a significant determinant of second mover behavior in a TG. Altruistic, maximin and inequality averse are more likely to pick honor than all other subjects.*

[17]All results reported in this section again use OLS regressions and heteroscedasticity robust standard errors and also use the same set of controls as in the previous section. The choice of the empirical model is again inconsequential for the conclusions.

[18]Moreover, we also again note that non-parametric Fisher exact tests, on the correlation between actual choice and i) the IA-Alt-Maximin dummy and ii) the predicted choice based on piecewise-linear preferences, are significant at the 1% level and thus confirm our conclusions (results not reported here).

Subject picked *honor*	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Prediction-honor	0.233***	0.198***	0.221***	0.222***	0.195**					
	(0.06)	(0.08)	(0.06)	(0.06)	(0.08)					
Δ-honor						0.359***	0.331***	0.354***	0.334***	0.315***
						(0.08)	(0.10)	(0.08)	(0.08)	(0.10)
Socio-demographics	No	Yes	No	No	Yes	No	Yes	No	No	Yes
Cognition & Personality	No	No	Yes	No	Yes	No	No	Yes	No	Yes
Attitudes	No	No	No	Yes	Yes	No	No	No	Yes	Yes
Observations	650	443	650	603	412	650	443	650	603	412
R^2	0.027	0.043	0.042	0.037	0.066	0.036	0.052	0.051	0.042	0.071

Table 6: Dependent variable is a dummy that indicates whether subject picked *honor* in trust game. *Prediction-honor* is a dummy that is equal to one for subjects who are predicted to pick honor on the basis of the elicited Charness and Rabin parameters. Δ*-honor* is the actual utility difference between the two allocations. OLS regression, robust standard errors in brackets. *, ** and *** indicate significance at the 10%, 5% and 1% level, respectively. A constant is included in all cases but not displayed here.

Subject picked *honor*	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
IA-Alt-Maximin	0.139***	0.124***	0.129***	0.126***	0.103**					
	(0.03)	(0.04)	(0.03)	(0.04)	(0.05)					
Altruist						0.182***	0.170***	0.178***	0.167***	0.139**
						(0.05)	(0.06)	(0.05)	(0.05)	(0.06)
Inequality averse						0.146***	0.149**	0.125**	0.130**	0.128*
						(0.05)	(0.06)	(0.05)	(0.05)	(0.07)
Maximin						0.106*	0.031	0.111*	0.119**	0.044
						(0.06)	(0.07)	(0.06)	(0.06)	(0.08)
Envious						-0.041	-0.038	-0.029	-0.010	-0.008
						(0.06)	(0.08)	(0.06)	(0.07)	(0.09)
Infrequent						0.132	0.125	0.130	0.111	0.100
						(0.08)	(0.10)	(0.08)	(0.08)	(0.10)
Socio-demographics	No	Yes	No	No	Yes	No	Yes	No	No	Yes
Cognition & Personality	No	No	Yes	No	Yes	No	No	Yes	No	Yes
Attitudes	No	No	No	Yes	Yes	No	No	No	Yes	Yes
Observations	650	443	650	603	412	650	443	650	603	412
R^2	0.022	0.041	0.037	0.029	0.058	0.029	0.054	0.044	0.033	0.065

Table 7: Dependent variable is a dummy that indicates whether subject picked *honor* in trust game. *IA-Alt-Maximin* is a dummy indicating that the subject is either inequality averse, maximin or altruistic. *Infrequent* is a dummy indicating that the subject is either spiteful, kick-down, equality-averse or kiss-up. OLS, robust standard errors in brackets. *, ** and *** indicate significance at the 10%, 5% and 1% level, respectively. A constant is included in all cases but not displayed here.

	Socio-demographics	Personality	Attitudes	All Controls
Prediction	0.55	0.67	0.73	0.36
Socio-demographics	0.45	-	-	0.27
Cognition & Personality	-	0.33	-	0.15
Attitudes	-	-	0.27	0.22

Table 8: Table displays standardized dominance statistics (in %). Dependent variable is a dummy indicating whether subject picked *honor* in the TG in an OLS regression model. The R^2 serves as the fit-statistic. The last column considers for computational reasons all variables in the set *socio-demographics, personality* and *attitudes* respectively as single independent variable.

7 Concluding Remarks

Evidence for the predictive power of social preferences for behavior in strategic decisions is surprisingly sparse and the available evidence is inconclusive. The present paper contributes to this literature by showing that social preferences elicited at one point in time significantly predict behavior in two experimental games played more than a year apart. We infer from this predictive success that social preferences exhibit a stable component. This finding is noteworthy on two accounts. First, the predictive success is remarkably strong because it is greater than the joint success of alternative predictive measures like socio-demographics, measures of cognitive ability, personality, and attitudes. Second, it sheds new light on the debate about whether social preferences are context dependent (Levitt and List, 2007). There is indeed evidence that behavior in the dictator game is motivated by a desire to signal that one is not entirely selfish or by a desire to follow a social norm that is choice-set dependent, see List (2007) and Bardsley (2008). This finding has been replicated by Cappelen et al. (2013) using the same subject pool and the same "virtual lab" approach as in the current study. Consistent with these results, we find that behavior in the standard dictator game has no predictive power for the two experimental games under consideration here. However, we do not use the standard dictator game but Kerschbamer's (2015) Equality-Equivalence Test to elicit distributional preferences in a systematic and comprehensive way. Hence, our findings caution against the use of the standard dictator game to elicit social preferences.

Another interesting finding arising from this study, is the characterization of the distribution of social preferences in a large and heterogeneous sample. We find that almost 90% of subjects are classified into one of just four preference types: altruism, inequality aversion, maximin and selfishness. This finding is in line with Bruhin, Fehr, and Schunk (2018) as well as Kerschbamer and Müller (2017). Both studies present evidence indicating that four preference types are sufficient to classify the vast majority of people. However, there are also nuances in the findings. In particular, we find that altruistic concerns are a more important driver of behavior than inequality aversion which contrasts with the results in Kerschbamer and Müller (2017) who use representative German data.

In all, our findings suggest a reconsideration of the relevance of distributional preferences for behavior in strategic interactions and highlight the importance of using a theory-driven approach to measure distributional preferences.

References

Andersson, O., H. J. Holm, J.-R. Tyran, and E. Wengström (2016): "Risk aversion relates to cognitive ability: Preferences or Noise?," *Journal of the European Economic Association*, 14(5), 1129–1154.

Azen, R., and D. V. Budescu (2003): "The dominance analysis approach for comparing predictors in multiple regression," *Psychological Methods*, 8(2), 129–148.

Balafoutas, L., R. Kerschbamer, and M. Sutter (2012): "Distributional preferences and competitive behavior," *Journal of Economic Behavior & Organization*, 83(1), 125–135.

Bardsley, N. (2008): "Dictator game giving: altruism or artefact?," *Experimental Economics*, 11(2), 122–133.

Blanco, M., D. Engelmann, and H. T. Normann (2011): "A within-subject analysis of other-regarding preferences," *Games and Economic Behavior*, 72(2), 321–338.

Bolton, G. E., and A. Ockenfels (2000): "ERC: A theory of equity, reciprocity, and competition," *American Economic Review*, 90, 166–193.

Bruhin, A., E. Fehr, and D. Schunk (2018): "The many faces of human sociality: Uncovering the distribution and stability of social preferences," *Journal of the European Economic Association*, pp. 1–55.

Burton-Chellew, M. N., and S. A. West (2013): "Prosocial preferences do not explain human cooperation in public-goods games," *Proceedings of the National Academy of Sciences*, 110(1), 216–221.

Cappelen, A. W., U. H. Nielsen, E. Ø. Sørensen, B. Tungodden, and J.-R. Tyran (2013): "Give and take in dictator games," *Economics Letters*, 118(2), 280–283.

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

Dannenberg, A., T. Riechmann, B. Sturm, and C. Vogt (2007): "Inequity aversion and individual behavior in public good games: An experimental investigation," *mimeo.*

De Bruyn, A., and G. E. Bolton (2008): "Estimating the influence of fairness on bargaining behavior," *Management Science*, 54(10), 1774–1791.

Derin-Güre, P., and N. Uler (2010): "Charitable giving under inequality aversion," *Economics Letters*, 107(2), 208–210.

Dreber, A., D. Fudenberg, and D. G. Rand (2014): "Who cooperates in repeated games: The role of altruism, inequity aversion, and demographics," *Journal of Economic Behavior & Organization*, 98, 41–55.

Engelmann, D., and M. Strobel (2010): "Inequality aversion and reciprocity in moonlighting games," *Games*, 1(4), 459–477.

Fehr, E., and U. Fischbacher (2002): "Why social preferences matter – The impact of non-selfish motives on competition, cooperation and incentives," *Economic Journal*, 112(478), C1–C33.

Fehr, E., and K. M. Schmidt (1999): "A theory of fairness, competition, and cooperation," *The Quarterly Journal of Economics*, 114(3), 817–868.

Fisman, R., P. Jakiela, and S. Kariv (2017): "Distributional preferences and political behavior," *Journal of Public Economics*, 155, 1–10.

Frederick, S. (2005): "Cognitive reflection and decision making," *Journal of Economic Perspectives*, 19(4), 25–42.

Gächter, S., D. Nosenzo, and M. Sefton (2013): "Peer effects in pro-social behavior: Social norms or social preferences?," *Journal of the European Economic Association*, 11(3), 548–573.

Harbaugh, W. T., and K. Krause (2000): "Children's altruism in public good and dictator experiments," *Economic Inquiry*, 38(1), 95–109.

Hernandez-Lagos, P., D. Minor, and D. Sisak (2017): "Do people who care about others cooperate more? Experimental evidence from relative incentive pay," *Experimental Economics*, 20(4), 809–835.

Höchtl, W., R. Sausgruber, and J.-R. Tyran (2012): "Inequality aversion and voting on redistribution," *European Economic Review*, 56(7), 1406–1421.

Holm, H. J., and A. Danielson (2005): "Tropic trust versus Nordic trust: experimental evidence from Tanzania and Sweden," *Economic Journal*, 115(503), 505–532.

Holt, C. A., and S. K. Laury (2002): "Risk aversion and incentive effects," *American Economic Review*, 92(5), 1644–1655.

Kamas, L., and A. Preston (2012): "Distributive and reciprocal fairness: What can we learn from the heterogeneity of social preferences?," *Journal of Economic Psychology*, 33(3), 538–553.

——— (2015): "Can social preferences explain gender differences in economic behavior?," *Journal of Economic Behavior & Organization*, 116, 525–539.

Kerschbamer, R. (2015): "The geometry of distributional preferences and a non-parametric identification approach: The Equality Equivalence Test," *European Economic Review*, 76, 85–103.

Kerschbamer, R., and D. Müller (2017): "Social preferences and political attitudes: An online experiment on a large heterogeneous sample," *Innsbruck Working Papers in Economics and Statistics*.

Kümmerli, R., M. N. Burton-Chellew, A. Ross-Gillespie, and S. A. West (2010): "Resistance to extreme strategies, rather than prosocial preferences, can explain human cooperation in public goods games," *Proceedings of the National Academy of Sciences*, 107(22), 10125–10130.

Levitt, S. D., and J. A. List (2007): "What do laboratory experiments measuring social preferences reveal about the real world?," *Journal of Economic Perspectives*, 21(2), 153–174.

List, J. A. (2007): "On the interpretation of giving in dictator games," *Journal of Political Economy*, 115(3), 482–493.

McCrae, R. R., and P. T. Costa Jr (2004): "A contemplated revision of the NEO five-factor inventory," *Personality and Individual Differences*, 36(3), 587–596.

Moffatt, P. G. (2015): *Experimetrics: Econometrics for experimental economics*. Palgrave Macmillan.

Murphy, R. O., and K. A. Ackermann (2017): "Explaining behavior in public goods games," *Academy of Management Proceedings*, 2015.

Offerman, T., J. Sonnemans, and A. Schram (1996): "Value orientations, expectations and voluntary contributions in public goods," *Economic Journal*, 106, 817–845.

Paetzel, F., R. Sausgruber, and S. Traub (2014): "Social preferences and voting on reform: An experimental study," *European Economic Review*, 70, 36–55.

Peysakhovich, A., M. A. Nowak, and D. G. Rand (2014): "Humans display a 'cooperative phenotype' that is domain general and temporally stable," *Nature Communications*, 5, 4939.

Teyssier, S. (2012): "Inequity and risk aversion in sequential public good games," *Public Choice*, 151(1-2), 91–119.

Thöni, C., J.-R. Tyran, and E. Wengström (2012): "Microfoundations of social capital," *Journal of Public Economics*, 96(7-8), 635–643.

Tyran, J.-R., and R. Sausgruber (2006): "A little fairness may induce a lot of redistribution in democracy," *European Economic Review*, 50(2), 469–485.

Yamagishi, T., Y. Horita, N. Mifune, H. Hashimoto, Y. Li, M. Shinada, A. Miura, K. Inukai, H. Takagishi, and D. Simunovic (2012): "Rejection of unfair offers in the ultimatum game is no evidence of strong reciprocity," *Proceedings of the National Academy of Sciences*, 109(50), 20364–20368.

Yamagishi, T., N. Mifune, Y. Li, M. Shinada, H. Hashimoto, Y. Horita, A. Miura, K. Inukai, S. Tanida, and T. Kiyonari (2013): "Is behavioral pro-sociality game-specific? Pro-social preference and expectations of pro-sociality," *Organizational Behavior and Human Decision Processes*, 120(2), 260–271.

Yang, Y., S. Onderstal, and A. Schram (2016): "Inequity aversion revisited," *Journal of Economic Psychology*, 54, 1–16.

A Appendix

A.1 The Relation between Scores, WTP and Charness-Rabin Parameters

X-list: subject chooses LEFT for the first time in row	X-score	Parameter range of σ	WTP_d
1	+ 4.5	$0.545 \leq \sigma$	$1.2 \leq WTP_d$
2	+ 3.5	$0.444 \leq \sigma < 0.545$	$0.8 \leq WTP_d < 1.2$
3	+ 2.5	$0.242 \leq \sigma < 0.444$	$0.32 \leq WTP_d < 0.8$
4	+ 1.5	$0.074 \leq \sigma < 0.242$	$0.08 \leq WTP_d < 0.32$
5	+ 0.5	$0 \leq \sigma < 0.074$	$0 \leq WTP_d < 0.08$
6	- 0.5	$-0.087 \leq \sigma < 0$	$-0.08 \leq WTP_d < 0$
7	- 1.5	$-0.471 \leq \sigma < -0.087$	$-0.32 \leq WTP_d < -0.08$
Never	- 2.5	$\sigma < -0.471$	$WTP_d < -0.32$

Table 9: WTP_d: amount of own material payoff the decision maker is willing to pay in the domain of **disadvantageous** inequality in order to increase the other's material payoff by one unit. The parameter σ is the weight on the other's income in the piecewise linear model. Note that the parameter α in the Fehr-Schmidt model corresponds to $-\sigma$ here.

Y-list: subject chooses LEFT for the first time in row	Y-score	Parameter range of ρ	WTP_a
1	- 2.5	$\rho \leq -0.471$	$WTP_a \leq -0.32$
2	- 1.5	$-0.471 < \rho \leq -0.087$	$-0.32 < WTP_a \leq -0.08$
3	- 0.5	$-0.087 < \rho \leq 0$	$-0.08 < WTP_a \leq 0$
4	+ 0.5	$0 < \rho \leq 0.074$	$0 < WTP_a \leq 0.08$
5	+ 1.5	$0.074 < \rho \leq 0.242$	$0.08 < WTP_a \leq 0.32$
6	+ 2.5	$0.242 < \rho \leq 0.444$	$0.32 < WTP_a \leq 0.8$
7	+ 3.5	$0.444 < \rho \leq 0.545$	$0.8 < WTP_a \leq 1.2$
Never	+4.5	$0.545 < \rho$	$1.2 < WTP_a$

Table 10: WTP_a: amount of own material payoff the decision maker is willing to pay in the domain of **advantageous** inequality in order to increase the other's material payoff by one unit. The parameter ρ is the weight on the other's income in the piecewise linear model. Note that the parameter β in the Fehr-Schmidt model corresponds to ρ here.

A.2 Dictator Game

In this section, we present evidence for the lack of correlation of behavior in the standard dictator game (DG) with behavior in other games. The DG was part of wave 2 in the iLEE. In this game, the dictator is endowed with 150 Dkr and decides on passing any amount of money from her endowment to the recipient. We use the amount kept by the dictator as dependent variable in the regressions below. In particular, in Table 11 we regress the y-score in columns (1) and (2), and the x-score in columns (3) and (4) on the amount kept by the dictator. In Table 12 we use the contribution in the

PGG as dependent variable, while in Table 13 a dummy, equal to one if the participant picked the right allocation as the second mover in the TG, serves as the dependent variable. In all cases, we present results with and without the usual set of controls.

As it turns out, in no case is this measure a significant predictor of behavior across games. In fact, it is not even related to behavior in the EET, which is a modified dictator game. This finding adds to the growing evidence suggesting that behavior in the standard dictator game is not reliable, see for example List (2007) and Bardsley (2008).

Equality-	Y-score		X-score	
Equivalence Test	(1)	(2)	(3)	(4)
Amount kept by Dictator	-0.004	-0.006	0.002	-0.002
	(0.00)	(0.00)	(0.00)	(0.01)
Controls	No	Yes	No	Yes
Constant	2.541***	8.000***	-0.101	3.794
	(0.42)	(3.04)	(0.45)	(3.55)
Observations	207	146	208	146
R^2	0.004	0.231	0.001	0.237

Table 11: Dependent variable is the y-score in columns (1) an (2) and the x-score in columns (3) and (4). OLS, robust standard errors in brackets. *, ** and *** indicate significance at the 10%, 5% and 1% level, respectively. A constant is included in all cases but not displayed here.

Contribution	(1)	(2)
Amount kept by Dictator	-0.047	-0.048
	(0.03)	(0.03)
Controls	No	Yes
Constant	40.646***	2.285
	(2.92)	(21.07)
Observations	314	213
R^2	0.009	0.167

Table 12: Dependent variable is the individual contribution in the PGG. OLS, robust standard errors in brackets. *, ** and *** indicate significance at the 10%, 5% and 1% level, respectively. A constant is included in all cases but not displayed here.

Subject picked *honor*	(1)	(2)
Amount kept by Dictator	-0.000	0.000
	(0.00)	(0.00)
Controls	No	Yes
Constant	0.268***	0.034
	(0.08)	(0.49)
Observations	314	213
R^2	0.000	0.104

Table 13: Dependent variable is a dummy that indicates whether subject picked *honor* in trust game. OLS, robust standard errors in brackets. *, ** and *** indicate significance at the 10%, 5% and 1% level, respectively. A constant is included in all cases but not displayed here.

A.3 Payments in the EET

The EET was carried out using two different payment protocols that vary whether there is uncertainty about the final role (decision maker or recipient) a subjects takes on in the EET. In one condition, roles are determined ex-ante and participants chosen to be decision makers know that their choices will affect the own material payoff and the payoff of a recipient for sure – while recipients make no choices and cannot affect outcomes. In the other condition, all participants take decisions as if they are decision makers and actual roles are randomly determined ex-post. We find that the two treatments do not affect the distribution of social preference types. We find however subtle evidence suggesting that the degree of benevolence in one domain might be affected.

Using Fisher exact tests, we do not find any evidence that the two payment protocols FixedRoles and RandomRoles influence the number of inconsistent decision makers ($p = 0.64$). Next, we test whether the payment protocol influences the distribution of types. Here, too, we are unable to find any evidence that supports the hypothesis that the payment protocol influences the decisions of subjects in the EET. The corresponding p-value of the likelihood ratio test is 0.49.[19] Looking at the intensity of social preferences – by considering the two scores, the x-score representing the benevolence of the decision maker in the domain of disadvantageous inequality and the y-score measuring the benevolence in the domain of advantageous inequality – we find some evidence suggesting that people exhibit a higher y-score (but not x-score) in the RandomRoles than in the FixedRoles treatment. A Fisher exact test yields a p-value of 0.02 (0.96) for the y-score (the x-score, respectively). In particular, the average y-score (x-score) is 1.88 (0.24) in the FixedRole treatment and 2.28 (0.19) in the RandomRole treatment, indicating that benevolence in the advantageous domain might be somewhat higher in the RandomRoles protocol. Nevertheless, all of this is of course inconsequential for the main results of the paper.

[19]The Fisher exact test delivers basically same result ($p = 0.55$). The finding that the payment protocol does not influence the distribution of social preference types is also backed up by nine different Fisher exact tests with the null hypothesis that a specific type is as frequent in the FixedRoles as in the RandomRoles treatment.

A.4 Robustness Section

We follow several different approaches to evaluate the robustness of our findings. First, we estimate a finite-mixture model of the four most prevalent distributional types – altruists, inequality averse, selfish and maximin – and use the posterior probability to classify inconsistent people into one of those four types. This robustness check confirms that the predictive power of distributional preferences does not depend on inconsistent subjects. Second, Section A.4.2 presents results from lowest prediction for those who are indifferent over some interval.

A.4.1 Mixture Model

In this section, we present robustness checks of our results using a finite-mixture of types model. Finite-mixture models have recently become increasingly popular in experimental economics (less so in the literature on social preferences, though), as they allow for several data-generating processes at the same time and are consequently a way to account for individual heterogeneity (Moffatt, 2015). We use this model to classify inconsistent subjects into distributional types based on posterior probabilities. We then include these subjects in our previous analysis. It is however important to note one main methodological difference to most other studies: The EET allows us to *perfectly* account for individual heterogeneity. Hence, we use the mixture model purely as a robustness check. Most other papers, like e.g. Bruhin, Fehr, and Schunk (2018), need to rely on these models for the main part of their analysis because their design does not allow to classify subjects into types at the individual-level.

In particular, we estimate a mixture of the four most prevalent distributional types – altruists, inequality averse, selfish and maximin subjects – in a random utility framework. Together these four types describe the behavior of almost 90% of the (consistent) subjects. Random utility models are based on the assumption that the utilities of all options are perturbed by a random error term. The decision maker then picks the option in which this perturbed utility is highest. That is, the decision maker has the highest *probability* of picking the option with the highest utility. The parametric structure that we impose on utility is again that of the piecewise linear utility functions of Charness and Rabin (2002) introduced earlier in equation (1).

The Fechner error version of the random utility model then assumes that the decision maker picks allocation (m^A, o^A) over (m^B, o^B) iff $U(m^A, o^A) + \varepsilon^A > U(m^B, o^B) + \varepsilon^B$. Denote the utility difference between allocation A and B by $\Delta = U(m^A, o^A) - U(m^B, o^B)$. Given the normally distributed Fechner error ε, the index Δ is transformed into a cumulative probability via the normal linking function $\Phi(\Delta)$. The log-likelihood for any given values of ρ and σ is then given by

$$lnL\left(\rho, \sigma | d\right) = \sum_{i=1}^{N} \left[d_i ln\left(\Phi(\Delta)\right) + (1 - d_i) ln\left(\Phi(-\Delta)\right)\right] \tag{7}$$

where d_i is a dummy that is equal to 1 if the DM picked allocation A.

The four main distributional types in our sample emerge from the Charness-Rabin via the following restrictions: $\rho > 0$ and $\sigma > 0$ (altruist), $\rho > 0$ and $\sigma < 0$ (inequality averse), $\rho = 0$ and $\sigma = 0$ (selfish) and finally $\rho > 0$ and $\sigma = 0$ (maximin). Let l_i^a, l_i^{ia}, l_i^s and l_i^m denote the individual likelihood contribution of observation i for the altruistic, inequality averse, selfish and maximin model, respectively. Then the grand log-likelihood of the mixture model is given by:

$$lnL\left(\sigma^a, \sigma^{ia}, \rho^a, \rho^{ia}, \rho^m, p^a, p^{ia}, p^s, \lambda|d\right) = \sum_{i=1}^{N} ln\left[p^a l_i^a + p^{ia} l_i^{ia} + p^s l_i^s + (1 - p^a - p^{ia} - p^s) l_i^m\right] \quad (8)$$

where p^t denotes the mixing proportion (that is, the relative frequency in the sample) of type $t \in \{a, ia, s, m\}$. The mixing proportion of maximin types is given by $p^m = 1 - p^a - p^{ia} - p^s$, without loss of generality. Finally, $\lambda > 0$ denotes the variance parameter. We restrict the variance to be equal across types, that is, we assume a homoscedastic error which eases the computational burden considerably relative to the heteroscedastic case.

Given parameter estimates and d, the posterior probability of subject j being of type t can be calculated using Bayes rule:

$$post_j^t = \frac{p^t L_j^t}{p^a L_j^a + p^{ia} L_j^{ia} + p^s L_j^s + (1 - p^a - p^{ia} - p^s) L_j^m} \quad (9)$$

where $L_j^t = \prod_{i=1}^{14} l_{j,i}^t$ and $t \in \{a, ia, s, m\}$. A subject that is inconsistent in the EET, is classified into one of the four types according to the highest posterior probability of her choices.

	Coefficient	Standard Error	Z-statistic	p-value
σ^a	0.413	.0155	26.57	0.000
ρ^a	0.125	.0153	8.18	0.000
σ^{ia}	-0.999	0.0001	-9219	0.000
ρ^{ia}	0.125	.017	7.49	0.000
ρ^m	0.554	.011	49.03	0.000
p^{ia}	0.205	.0193	11.02	0.000
p^s	0.112	.038	2.92	0.004
p^a	0.210	.023	9.12	0.000
λ	0.879	.004	198.21	0.000

Table 14: Parameters of the 4-type mixture model. Standard errors clustered at the individual-level are calculated using the delta method. $N = 12,390$.

Table 14 reports the parameter estimates. Tables 15 and 16 replicate the previous analysis.[20] As these tables show, this robustness check strongly confirms our previous conclusions.

[20] The other results where the individual-level prediction serves as a regressor can of course not be replicated because the mixture model only allows for the classification into types.

Contribution	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
IA-Alt-Maximin	3.712***	3.662***	3.132***	3.096***	3.532***					
	(1.04)	(1.24)	(1.06)	(1.08)	(1.31)					
Altruist						5.571***	3.870**	5.046***	5.276***	4.095**
						(1.57)	(1.82)	(1.54)	(1.64)	(1.93)
Inequality averse						2.587	1.397	1.706	1.816	0.525
						(1.62)	(1.86)	(1.60)	(1.69)	(1.92)
Maximin						3.991**	3.162	3.588**	4.058**	3.131
						(1.64)	(2.02)	(1.61)	(1.68)	(2.02)
Envious						0.906	-2.052	2.119	2.497	-1.075
						(3.05)	(3.39)	(3.13)	(3.21)	(3.79)
Infrequent						-1.147	-2.954	-1.736	-0.789	-2.790
						(3.06)	(3.95)	(3.12)	(3.27)	(4.32)
Socio-demographics	No	Yes	No	No	Yes	No	Yes	No	No	Yes
Cognition & Personality	No	No	Yes	No	Yes	No	No	Yes	No	Yes
Attitudes	No	No	No	Yes	Yes	No	No	No	Yes	Yes
Observations	885	588	885	820	549	885	588	885	820	549
R^2	0.015	0.052	0.059	0.030	0.098	0.021	0.053	0.066	0.038	0.101

Table 15: Robustness Check **Mixture Model**. Dependent variable is the individual contribution in the **PGG**. *IA-Alt-Maximin* is a dummy indicating that the subject is either inequality averse, maximin or altruistic. *Infrequent* is a dummy indicating that the subject is either spiteful, kick-down, equality-averse or kiss-up. OLS, robust standard errors in brackets. *, ** and *** indicate significance at the 10%, 5% and 1% level, respectively. A constant is included in all cases but not displayed here.

Subject picked **honor**	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
IA-Alt-Maximin	0.065**	0.084**	0.071**	0.057*	0.076**					
	(0.03)	(0.04)	(0.03)	(0.03)	(0.04)					
Altruist						0.151***	0.127**	0.158***	0.128***	0.107**
						(0.04)	(0.05)	(0.04)	(0.04)	(0.05)
Inequality averse						0.135***	0.150***	0.123***	0.123***	0.134**
						(0.04)	(0.05)	(0.04)	(0.05)	(0.06)
Maximin						0.096**	0.039	0.094**	0.094**	0.041
						(0.04)	(0.05)	(0.04)	(0.05)	(0.05)
Envious						-0.072	-0.055	-0.076	-0.044	-0.042
						(0.06)	(0.08)	(0.06)	(0.07)	(0.09)
Infrequent						0.103	0.111	0.098	0.086	0.087
						(0.08)	(0.10)	(0.08)	(0.08)	(0.10)
Socio-demographics	No	Yes	No	No	Yes	No	Yes	No	No	Yes
Cognition & Personality	No	No	Yes	No	Yes	No	No	Yes	No	Yes
Attitudes	No	No	No	Yes	Yes	No	No	No	Yes	Yes
Observations	885	588	885	820	549	885	588	885	820	549
R^2	0.007	0.029	0.016	0.013	0.045	0.023	0.042	0.031	0.024	0.053

Table 16: Robustness Check **Mixture Model**. Dependent variable is a dummy that indicates whether subject picked *honor* in **trust game**. *IA-Alt-Maximin* is a dummy indicating that the subject is either inequality averse, maximin or altruistic. *Infrequent* is a dummy indicating that the subject is either spiteful, kick-down, equality-averse or kiss-up. OLS, robust standard errors in brackets. *, ** and *** indicate significance at the 10%, 5% and 1% level, respectively. A constant is included in all cases but not displayed here.

A.4.2 Robustness: Prediction in PGG

Contribution	(1)	(2)	(3)	(4)	(5)
Prediction-low	0.224***	0.202***	0.205***	0.196***	0.199***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Socio-demographics	No	Yes	No	No	Yes
Cognition & Personality	No	No	Yes	No	Yes
Attitudes	No	No	No	Yes	Yes
Observations	650	443	650	603	412
R^2	0.049	0.088	0.096	0.067	0.140

Table 17: Dependent variable is the individual contribution in the PGG. OLS, robust standard errors in brackets. *Prediction-low* is the predicted contribution of the piecewise linear model using the **lowest value** for indifferent subjects. *, ** and *** indicate significance at the 10%, 5% and 1% level, respectively. A constant is included in all cases but not displayed here.

A.5 General iLEE Procedures

The experiment is conducted using the platform of the internet laboratory for experimental economics (iLEE) at the University of Copenhagen, Denmark. Subjects for the platform are recruited with the assistance of the official statistics agency (Statistics Denmark) who select a random sample from the general population. Statistics Denmark sends the selected individuals physical letters, inviting them to participate in an online scientific experiment that is jointly organized by the University and Statistics Denmark. Participants log in to the experiment using a personal identification code provided by Statistics Denmark. Payments are executed by electronic bank transfer and participants remain anonymous to the researchers at the University throughout the experiment. The EET is part of the third wave of experiments conducted on the iLEE platform. All three waves are run using the same set of participants, thus creating a panel data set useful for cross game analysis. For the third wave, we invite the 2291 people who completed the first wave. In total, 1067 participants complete the third wave between July and September, 2010. Participants can log on at any point during this period and are free to log out and continue later at their convenience. The third wave consists of a total of six different parts. The first part of the third wave consisted of a trust game, followed by four other, smaller parts: a real effort task, a voting game, measures of risk and loss aversion and our application of the EET. The order of these four parts is random. The final part is a questionnaire which includes questions on age, gender and education. In total, the median person spends 63 minutes completing the entire wave and earns 279 DKr (37 euros). Cooperation with Statistics Denmark is necessary to obtain the names and addresses of participants needed to send out invitations but our cooperation also yields additional advantages. First, it allows us to target a representative sample of the population. Combined with the high penetration of internet access in Denmark, this means that we have participants from all walks of life, which enables us to investigate how experimental behavior is correlated with self-reported socio-economic variables such as age, education and employment. Second, our procedures entail double blindness in the sense that participants are anonymous not only to other participants but also to us, the experimenters. Anonymity is important to minimize potential experimenter-demand effects. Levitt and List (2007) survey evidence that shows how the lack of anonymity between experimenters and participants increases the level of pro-social behavior when measuring distributional preferences. Double-blindness should decrease such effects. Questionnaire Participants answer questions regarding their basic socio-economic background, including their age, gender and level of education. In the analysis below, we group education in four categories: primary (no more than 10 years, 6 percent), secondary (vocational and high school, 22 percent), short tertiary (50 percent) and long tertiary (22 percent). In addition, we ask participants to answer five attitude questions from the World Values Survey. Participants had the option of not answering the questions. About 8 percent choose to not answer at least one of the following five questions:

LeftRight: "In political matters, people talk of 'the left' and 'the right'. How would you place your views on this scale if 1 means the left and 10 means the right?" Possible answers are integers ranging from 1: "left" to 10: "right".

Responsibility: "We would like your opinion on important political issues. How would you place your views on a scale from 1 to 10?" Possible answers are integers ranging from 1: "People should take more responsibility to provide for themselves" to 10: "The government should take more responsibility to ensure that everyone is provided for".

Competition: How would you place your views on a scale from 1 to 10?" Possible answers are integers ranging from 1: "Competition is good. It stimulates people to work hard and develop new ideas" to 10: "Competition is harmful. It brings out the worst in people".

Trust: "Generally speaking, would you say that most people can be trusted or that you cannot be too careful in dealing with people?" Possible answers are 0: "Cannot be too careful" and 1: "Most people can be trusted".

Fairness: "Do you think most people would try to take advantage of you if they got a chance, or would they try to be fair?" Possible answers are integers ranging from 1: "Would take advantage of you" to 10: "Would try to be fair".

Psychological measures:

We are also able to include psychological measures in the survey that participants answer. These measures consist of a cognitive reflection test (CRT), an IQ test and a personality test. The CRT is due to Frederick (2005) and consists of three short questions that all have incorrect but "intuitive" answers. Hence, the CRT is aimed at capturing participants ability to reflect upon a question and resist the temptation of giving the first (wrong) answer that comes to mind. Frederick finds that the CRT is predictive of behavior in a number of decision making environments. The three questions are: 1: "A bat and a ball cost 110 Dkr in total. The bat costs 100 Dkr more than the ball. How much does the ball cost?" Answer is given in Dkr 2: "If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets?" Answer is given in number of minutes 3: "In a lake, there is a patch of lily pads. Every day, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover half of the lake?" Answer is given in number of days.

The intuitive answers to the questions are (10, 100 and 24) while the correct answers are (5, 5 and 47). The variable CRT score is calculated as the number of correct answers, i.e. 0, 1, 2 or 3. Our measure of IQ is based on the I-S-T 2000R intelligence structure test (which we use by permission of Dansk Psykologisk Forlag who administers it in Denmark). The test is based on Raven's Progressive Matrices and participants have 10 minutes to solve 20 puzzles. As our IQ score variable, we use the number of correct answers, from 0 to 20.

Our measure of personality is based on the Five Factor Model (McCrae and Costa Jr, 2004) which describes human personality according to the "Big Five" dimensions or traits: Openness (to experience), Conscientiousness, Extraversion, Agreeableness and Neuroticism. Openness is related to creativity, to being curios and original and to the person's ability to contemplate new ideas. Conscientiousness is related to having a will to achieve, to being conscientious, hard-working and well-organized and to being ambitious. Extraversion is related to being social, passionate, talkative and dominating in groups. Agreeableness is related to kindness and altruism and to being good-natured and trusting. Neuroticism is related to being emotional, worried, self-conscious and temperamental. We use the short version of the NEO PI-R test with 60 questions. The test yields scores for each of the five dimensions on a scale from 1 to 48. A higher score means that a personality is correlated with a higher degree of the particular trait. For example, a person who scores 40 on Neuroticism is likely to be more emotional than a person who scores 5.

A.6 Instructions

The first part of the instructions are identical for both the FixedRoles and RandomRoles treatments.

In this part of the experiment, there are two roles: decision makers and recipients. A decision maker makes 14 choices on behalf of the person him-/herself and a randomly selected second participant (the recipient). Every choice is between two alternatives: LEFT and RIGHT. The alternative chosen by the decision maker will determine the payment for both the decision maker and the recipient.

Here is an example:

If the decision maker chooses LEFT, he/she gets 70 kr. and the recipient gets 25 kr. If the decision maker chooses RIGHT, he/she gets 50 kr. and the recipient gets 50 kr.

The continued instructions differ depending on the treatment:

Fixed role treatments:

Only decision makers are asked to make the 14 choices. Recipients make no decisions. Half the participants will be decision makers and the other half will be recipients. What role you get is determined randomly before the decisions are made. It is as likely that you will be decision maker as it is that you will be recipient. Once the roles are determined, each decision maker is randomly matched with a recipient. Only one of the decision maker's 14 choices will be selected for payment. All choices have the same probability of being selected.

On the next screen you will be told whether you have been chosen to be a decision maker or a recipient. Remember that if you are selected to be a decision maker, your choices will determine both your own and a recipient's earnings from this part of the experiment. The recipient will only get a payment from your decisions and no further payment. If you are selected to be a recipient, your earnings will be solely determined by another participant's choices. In this case, you will not yourself make any choices in this part of the experiment.

Random roles treatment:

All participants are asked to make the 14 choices as if they are decision makers. Half the participants will actually be decision makers whose choices will count whereas the other half will be recipients whose choices will not count. What role you get is determined randomly after the experiment has ended. It is as likely that you will be decision maker as it is that you will be recipient. Once the roles are determined, each decision maker is randomly matched with a recipient. Only one of the decision maker's 14 choices will be selected for payment. All choices have the same probability of being selected. On the next screens you will make the 14 choices between LEFT and RIGHT. Remember that if you are selected to be a decision maker, your choices will determine both your own and a recipient's earnings from this part of the experiment. The recipient will only get a payment from your decisions and no further payment. If you are selected to be a recipient, your earnings will be solely determined by another participant's choices. In this case, your choices in this part of the experiment will have no effect on anyone's payment (neither on your own payment nor on anybody else's payment).

Subjects in the FixedRoles treatment see an additional screen informing them of the outcome of the random draw that determines their role:

Fixed Roles – Subjects chosen to be decision makers:

You have been randomly selected to be a decision maker. On the next screens you will make the 14 choices between LEFT and RIGHT. Remember that your choices will determine both your own and a recipient's earnings from this part of the experiment. The recipient will only get a payment from your decisions and no further payment.

Fixed Roles – Subjects chosen to be recipients:

You have been randomly selected to be a recipient. Your earnings will be solely determined by another participant's choices. You will not make choices in this part of the experiment yourself.

A.7 Screenshots Experiment

Translation Figure 5:

Confirm your choices. You now have the option to examine your choices and possibly to revise them. Your selections are pointed out by colors in the table below. If you wish to revise a decision, click Revise (Revider). You will then again see the decision screen for this decision. Afterwards, you will return here and your revised choice will be apparent below.

VENSTRE = LEFT, HØJRE = RIGHT

Du får = you get, modtageren får = the recipient gets.

Du valgte = You chose.

Revider dette valg? = Revise this decision.

Bekræft valg = Confirm decisions

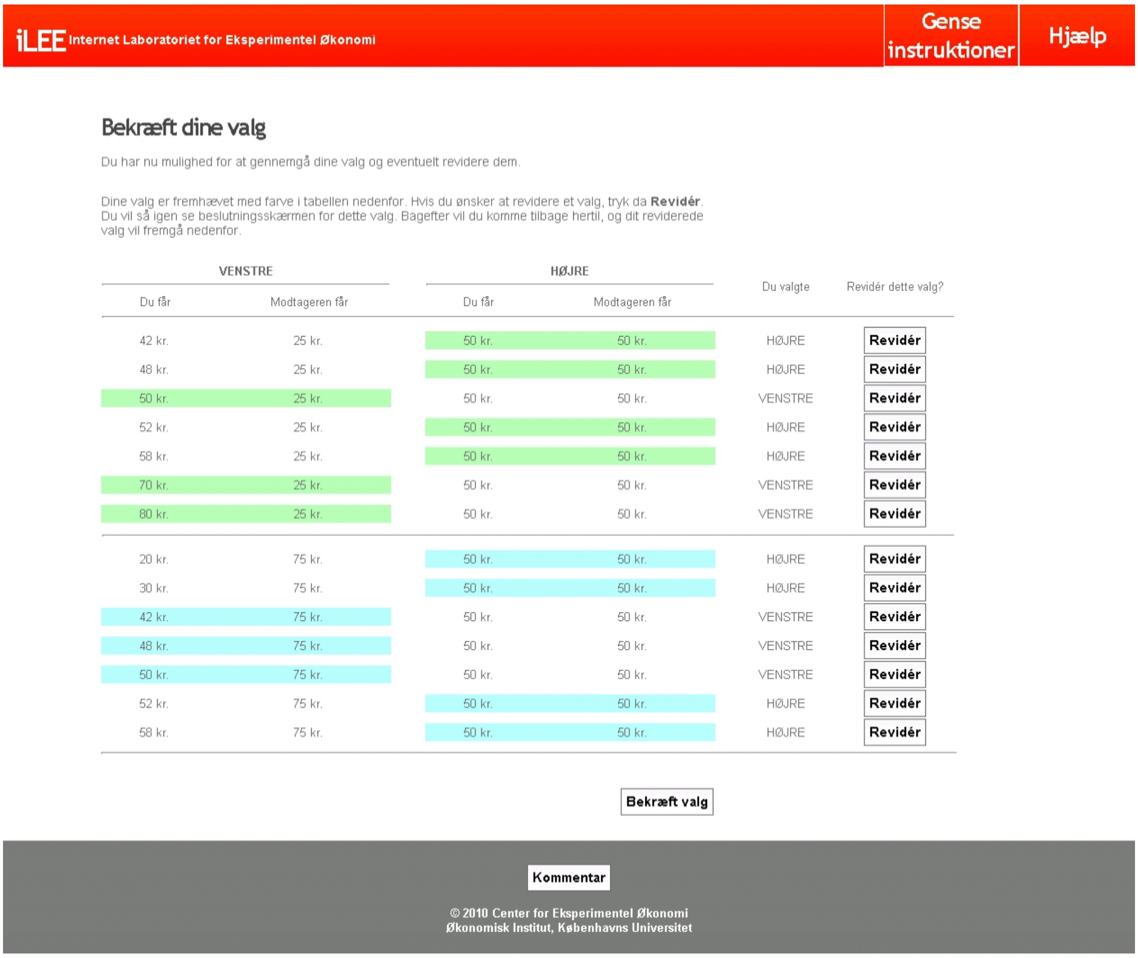

Figure 5: Screenshot EET.

Top bar: Gense instruktioner = Repeat instructions.

Hjælp = Help.

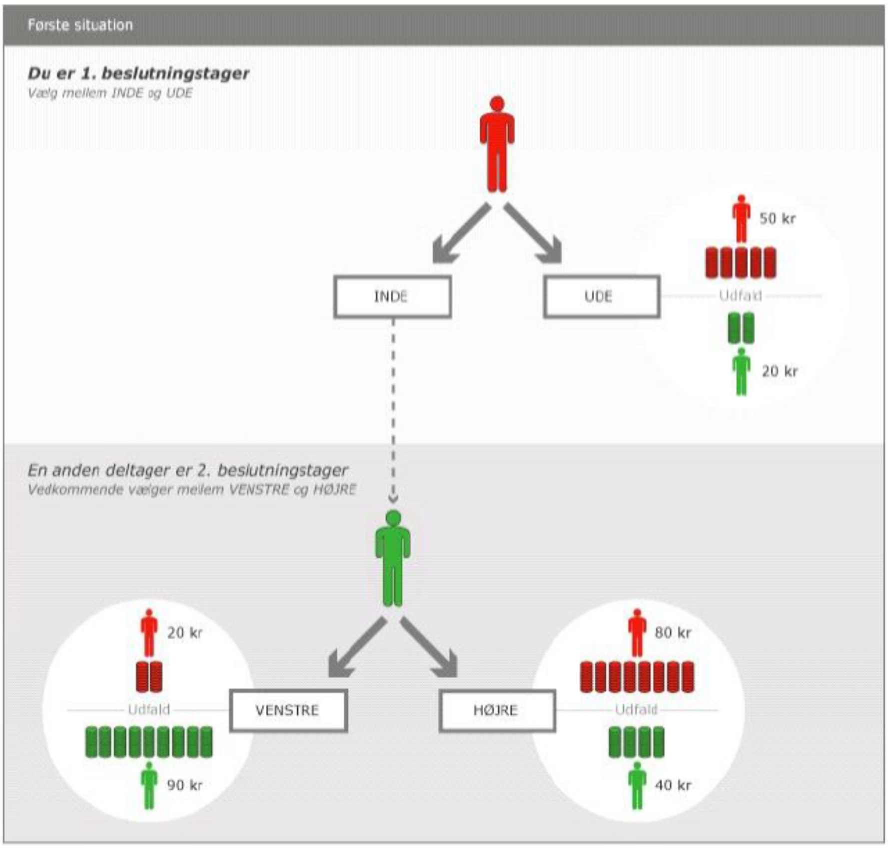

Figure 6: Screenshot Trust Game.