

Measuring tastes for equity and aggregate wealth behind the veil of ignorance

Jan Heufer^a, Jason Shachat^b, and Yan Xu^c

^aErasmus University Rotterdam and Tinbergen Institute

^bDurham University Business School and Wuhan University

^cUniversity of Vienna

This version: July 2022

Abstract

We propose an instrument to measure individuals' social preferences regarding equity and efficiency behind a veil of ignorance while controlling for idiosyncratic risk preferences. We construct a battery of portfolio and wealth distribution choice problems sharing a common budget set. A given bundle induces the same distribution over an individual's wealth in both problems. The portfolio choice solely reflects an individual's risk attitude, providing a benchmark to evaluate whether their wealth distribution choice exhibits equity or efficiency preferring tastes. Experiments show clusters of social preference types, which are unexpectedly independent of risk preferences.

JEL classification: C14; C91; D11; D12; D63.

Keywords: Inequity aversion; revealed preference; risk preferences; social preferences; veil of ignorance.

^aErasmus School of Economics and Tinbergen Institute, Erasmus University Rotterdam, Burgemeester Oudlaan 50, 3062 PA Rotterdam, Netherlands. Email: heufer@ese.eur.nl. ^b Durham University Business School and Wuhan University, Durham University, Mill Hill Lane, Durham DH1 3LB, United Kingdom. Email: jason.shachat@durham.ac.uk. ^c Corresponding author, Department of Economics, University of Vienna, Oskar-Morgenstern-Platz 1, 1090 Vienna, Austria. Email: yan.xu@univie.ac.at.

1 Introduction

Economic policies often strive to balance maximizing total wealth, what we call efficiency, and delivering sufficiently equal individual shares of that wealth, what we call equity. Failing too miserably in either generating overall prosperity (Acemoglu et al., 2002, Rodrik and Wacziarg, 2005, Joseph, 2008) or delivering sufficiently equitable prosperity (Alesina and Rodrik, 1994, Alesina and Perotti, 1996, Benabou, 2000) can lead to political and social instability. Knowledge of individuals' preferences over efficiency-equity trade-offs is key to performing this balancing act.

In this paper, we present a new instrument to measure these individual efficiency-equity preferences and demonstrate its efficacy. Our instrument measures preferences behind a veil of ignorance, VoI hereafter, and disentangles pure efficiency-equity preferences from those of risk aversion over one's own uncertain wealth. To avoid the ambiguity concerns found in Rawls's (1958) VoI formulation, we adopt a formulation in which each potential placement in society is equally likely (Harsanyi, 1953, 1976). We take a standard economics approach to uncovering individual preferences. If a decision maker, DM hereafter, is a price taker and technology or expenditure constrain her set of feasible choices, then her choice equates her willingness to trade-off one choice element for another with their relative prices. We incorporate this idea by having the DM select her preferred alternative from a linear budget set.

At the core of our instrument is a pair of standard consumer choice problems. Specifically, in each problem the DM chooses a two element commodity bundle, $(x_i, y_i) \in \mathbb{R}_+^2$, where i is the problem type, from a budget set $q x_i + y_i = z$ with relative price $q \geq 1$ and expenditure $z > 0$. In the first type of the problem, VoI, the DM is a member of a two-person economy and chooses a pair of wealth levels, one each for the Poor (x_{VoI}) and Rich (y_{VoI}) individuals. When making this choice, the DM knows there is a fifty percent chance she will be Poor and the other individual Rich and a fifty percent chance it will be vice versa. In this case, the linear budget set is a wealth profile possibilities frontier and the "price" is the constant loss rate in efficiency from redistributing a unit of wealth from the Rich to the Poor position. When this price is strictly greater than one, total wealth is maximized by assigning all wealth to the Rich and none to the Poor position. To maintain the roles of Rich and Poor we further restrict the DM to choosing profiles lying above the forty-five degree line, the locus of equal wealth profiles. Figure 1a depicts the commodity space under this restriction. Figure 1b depicts an example of the DM's choice, $(x_{\text{VoI}}^*, y_{\text{VoI}}^*)$; the tangent indifference curve we draw is not necessary but helps us to set the stage to ask questions regarding her rationality. But for now, using only this choice, what can we conclude about the DM's preferences regarding efficiency and equity?

We posit not much. This choice does not reflect her pure preferences in terms of the trade-off between aggregate wealth levels and its distribution because it is conflated with

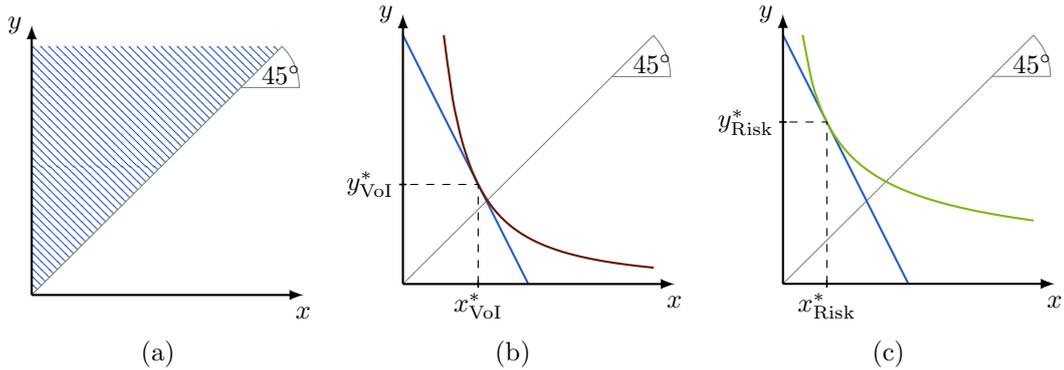


Figure 1: The commodity space.

the DM’s risk preferences over her own terminal wealth. We introduce a second choice problem, Risk, to provide control for such risk preferences. In the Risk problem, the DM selects a state-contingent wealth profile from a set of feasible portfolios. These portfolios generate exactly the same distribution over individual wealth as the one from the VoI problem.

In a Risk problem, the DM chooses a portfolio of contingent claim assets in a world with two equally likely states, Good and Bad.¹ We restrict the DM’s portfolio $(x_{\text{Risk}}, y_{\text{Risk}})$ to satisfy $y_{\text{Risk}} \geq x_{\text{Risk}} \geq 0$, thus preserving the roles of Good and Bad for the two states. The commodity space is the same as in the VoI problem. As before, the DM chooses a portfolio from a linear budget constraint $q x_{\text{Risk}} + y_{\text{Risk}} = z$ where $q \geq 1$; the price of insurance is at best actuarially fair. Figure 1c depicts an example where the consumer makes a choice facing the same price and expenditure as the VoI problem in Figure 1b.

How do these two problems differ? In the VoI problem, the DM receives one of the two possible amounts and somebody else the other, whereas in the Risk problem, the DM receives one of the two possible amounts and there is nobody else. How are these problems similar? For a fixed price and expenditure, a bundle on the VoI budget line generates the same marginal distribution over the DM’s terminal wealth as the same bundle on the Risk budget line. Hence, the choice in the Risk problem provides us a benchmark for the VoI choice problem that controls for the DM’s risk preference.²

¹In our experiments we use the labels High and Low reward for both the VoI and Risk problems. For exposition in this paper, we use the corresponding labels Rich and Poor for VoI problems and Good and Bad for Risk problems. We also use the terms High and Low when we speak generically of a choice problem.

²One could argue that the DM’s VoI and Risk choices differ only due the DM’s custodial duties managing the other person’s risk rather than any equity or efficiency concerns. A number of recent studies suggest that this is not the case. These studies compare the degree of risk aversions exhibited when a DM makes a series of binary lottery choices for themselves and the same series of lottery choices where they and another individual are beneficiaries; we do note the same lottery realization is used to determine the rewards, creating perfect positive rather than negative correlation. When lotteries are solely in the gain domain the majority of studies, such as Andersson et al. (2014), Vieider et al. (2016), Füllbrunn and Luhan (2017, 2015), find no difference in risk aversion. Two notable exceptions are Sujoy et al. (2011), Pahlke et al. (2015).

We define three classifications of social preferences for a DM’s choice by comparing her chosen pair of quantities demanded $x_{\text{VoI}}(q, z)$ and $x_{\text{Risk}}(q, z)$ for a budget with price ratio q and expenditure z . First, when these quantities are the same, then the DM is not bothered by the societal implications of her choice and we call her an *egoist*. Second, if the DM gives more to the Poor in the VoI problem than to the Bad in the Risk problem, $x_{\text{VoI}}(q, z) > x_{\text{Risk}}(q, z)$, we call her *equity preferring*. Third, if the DM gives a smaller amount to the Poor, $x_{\text{VoI}}(q, z) < x_{\text{Risk}}(q, z)$, we call her *efficiency preferring*. In the example we provide in Figure 1 the DM prefers equity for the given budget line. We complete our instrument by compiling forty rounds of such paired problems, where each round differs in the price and the expenditure level.

We implement our instrument by introducing a novel and intuitive graphical interface to our experimental software. Recent studies testing and modeling individual preferences from a revealed preference perspective largely rely upon collecting choice data using an approach introduced by Choi et al. (2007b). In their approach, an individual faces a budget line graph, as commonly presented in microeconomics textbooks, and chooses a bundle by selecting a point on the budget line. In our software interface, by contrast, an individual makes choices by moving a slider to select proportions of an amount of money between High and Low rewards. A matching pie chart depicts the two reward amounts associated with a slider location. As the slider position changes, the pie chart adjusts in both total and reward slice sizes. We feel this presents a more familiar and intuitive approach to participants not having previous experience with the budget line representation.

We evaluate the efficacy of our instrument by conducting financially incentivized experiments with 92 students from Xiamen University. Our instrument necessitates a within-subject experimental design in which each participant makes both VoI and Risk choices. The within-subject design paired with our analytic approaches enable us identify individual heterogeneity in social preference types. This heterogeneity can be misdiagnosed as homogeneous egoist behaviour in population level and between subject experimental studies.

In our first analysis of the data, we evaluate in which of the three social preference categories the average difference between the amounts given to Low rewards in the VoI and Risk problems lies. When examining these differences for all paired choices over all participants we find a very narrow confidence interval around zero. Thus, if we assume homogeneity across individuals then we conclude the population is egoist. However, inspection of each individual’s average difference in the forty paired choices reveals large sized clusters for each social preference class.

An unintended discovery is evidence that social and risk preferences are independent. For each of three clusters we estimate reduced demand functions for both the allocation to Poor, $x_{\text{VoI}}(q, z)$, and Bad, $x_{\text{Risk}}(q, z)$. Surprisingly, their estimated demand curves for x_{Risk}

are statistically insignificant, while their demand curves for x_{VoI} are significantly different from each other. In other words, our classifications capture individuals' distinctive social preference types since all three sub-groupings share a similar risk preference. This result also suggests that there is no correlation between the social preference categorization and the risk preference.

Our second analysis of the data utilizes revealed preference arguments, not requiring the strong statistical assumptions of our first analysis, which allow inferences regarding the deeper structure of preferences and the rationality of individuals. First, we follow the standard practice based on revealed preference axioms and test subjects' choices for consistency and homotheticity with utility maximization. Most of our participants are maximizing a concave (though not necessarily homothetic) utility function. Moreover, our choice data are comparable with other budget experiments, nullifying the ex-ante concerns of our graphic design, price variations, and choice restrictions. Second, we provide a nonparametric classification of participants as efficiency preferring, equity preferring, or egoist based on intra-personal comparisons of each individual's VoI and Risk domain indifference curves. For instance, if the indifference curve recovered from the DM' choices for VoI problems is globally more convex than the one from her choices for Risk problems, her Risk-self will have a relative more concave utility function than her VoI-self, and the DM is more likely to favor equity than efficiency. If otherwise the DM's Risk-self indifference curve is relatively more convex than her VoI-self's, she is more likely to be classified as efficiency preferring.

In this exercise we find that about twenty-eight percent of the subjects' Risk utility functions are globally more concave than their VoI ones, exhibiting a taste for efficiency. Thirty-three percent of the subjects are equity preferring since their VoI utility functions are more globally concave. Seventeen percent of subjects are in the egoist subgroup by having utility functions that are indistinguishable across domains, and finally twenty-two percent of the subjects' utility functions differ in domain but are un-ordable in terms of relative concavity. Using this nonparametric classification results, we re-estimate the reduced demand functions for both the allocation to Poor, $x_{\text{VoI}}(q, z)$, and Bad, $x_{\text{Risk}}(q, z)$. We still find that demand for $x_{\text{Risk}}(q, z)$ is uncorrelated with the assignment to social preference category.

In our third analysis of the data, we provide a structural estimation of a two-parameter subjective expected utility model at the individual level. One parameter reflects an individual's subjective prior for the Low state, which under the appropriate restriction reduces the model to an expected utility model, and the other measures the curvature of the power utility function. We find the majority of participants are optimistic by over-weighting the Good/Rich state. The estimated curvatures are generally higher for VoI than for Risk problems. This is mainly driven by the larger magnitudes of the right tail of VoI-selves' curvature estimates. Though highly noisy in terms of standard errors,

the difference between individuals' estimated curvatures displays significant clusters of the three social preference types. We find seventeen percent of subjects' utility functions are significantly more concave for Risk than for VoI tasks, exhibiting equity preferring while thirteen percent subjects significantly show a taste for equity.

The rest of the paper is structured as follows. In Section 2, we discuss the related literature on measuring individual social preferences with induced budget experiments. In Section 3, we introduce the experimental design and procedures. Section 4 presents the standard statistical data analysis and estimates the reduced demand curves. In Section 5, we present our revealed preference analysis. Section 6 reports our structural estimation results. Section 7 concludes.

2 Literature review

Recent work has recognized the importance of estimating the distributional preferences of those who set policies and of those who follow them (Choudhry et al., 1997, Nelson et al., 2018, Kuhn, 2013). For example, Fisman et al. (2015) and Li et al. (2017) conduct induced budget dictator game experiments respectively with medical and law students at prestigious universities. They presume these students are representative of future policy designers in health and other areas. In the domain of those who are governed by policies, Saez and Stantcheva (2016) and Kuziemko et al. (2015) elicit redistribution preferences from random samples of U.S. taxpayers. Then they explore how these preferences influence optimal taxation policy. Alternatively, Fisman et al. (2017) conduct an induced budget dictator game experiment with a random sample of American voters. They relate these preferences to voting behaviour and political party affiliation.

Most of these studies use a variation of the dictator game to elicit a social preference measure. More broadly, there are three prominent elicitation methods, each imposing an alternative DM perspective. These three perspectives are (1) as a dictator in front of a VoI, (2) as a disinterested social planner behind a VoI and (3) as a society member behind a VoI.³

A standard Dictator game has two players in which one decides how a fixed amount of money will be divided between herself and the other.⁴ Individuals typically do not

³Kariv and Zame (2014) propose an interesting theoretical model in which well behaved preference orderings corresponding to our Risk setting and a Dictator setting over the wealth profiles of our VoI setting are projections from a preference ordering in our VoI setting. Under assumptions of sufficient DM rationality and number of Risk and Dictator choices, one can recover the DM's VoI preferences.

⁴A variant of the dictator game incorporates a risk component in the dictator or the recipients payoffs (Krawczyk and Le Lec (2010), Brock et al. (2013), Freundt and Lange (2017), Cettolin et al. (2017)). In these experiments, the dictator makes a trade-off between her monetary payoff (certain or risky) and the recipients chance to win a fixed prize. They find that the risk exposure of the recipient affects the dictators' choices, indicating the role of ex-ante fairness concerns in social preference. While these studies are interested in how the presence of risk impacts social preferences, our study treats the inherent risk of choosing from behind the VoI as a nuisance factor.

claim the total amount for themselves and a large literature has identified multiple motivations for the finding. Some studies modify the standard version of the dictator game by changing the set of possible choices to include allocations which vary in both total aggregate wealth and wealth inequality. For instance, Charness and Rabin (2002) and Engelmann and Strobel (2004) limit the dictator's choice sets to a small number of wealth profiles, each corresponding to an alternative theory of social preferences. They establish that individual choices are driven by efficiency concerns as much as equity ones and then propose quasi-maximin preference models, loosely speaking weighted averages of efficiency and the maximin payoff, to explain their aggregate data. Cox (2004) provides a stronger demonstration of efficiency motives with a dictator game in which the two players are endowed with the same level of currency and the dictator can increase the wealth of the other player with a price of one-third; almost two-thirds of the dictators sent positive amounts of money. These studies effectively demonstrate that aggregate choices reflect consideration of efficiency in addition to that of inequity, but they do not provide the opportunity to explore the structure and heterogeneity of individual preferences.

A different strand of the literature examines the rationality and structure of individual dictator preferences through induced budget experiments where participants choose from multiple linear wealth profile possibilities frontiers. Andreoni and Miller (2002) is the first in this strand with a core treatment consisting of eight downward sloping budget sets with prices varying from one-third to three. Over ninety percent of their participants do not violate GARP, as compared to approximately twenty-two percent from a simulation of individuals who simply make choices according to a uniform distribution on a budget set. Moreover, they find three clusters of individual choice patterns: claim all available wealth, choose according to Leontief preferences (which is consistent with inequity aversion) and maximize aggregate wealth. Fisman et al. (2007) introduce a graphical interface allowing for a rapid collection of decisions. In their design, participants complete fifty budget set dictator problems with prices mainly ranging from one-half to two. They find participants' choices are largely consistent with GARP, reflect similar behavioural types as found in Andreoni and Miller, and this is reflected in the various clustering of estimated utility functions in the family of constant elasticity of substitution forms. We should note that both of these two studies differ from ours in that each participant is paid for two randomly selected decisions, once using their dictator choice and once when they are the recipient of a dictator.

A limited number of studies have elicited social preferences using both the perspective of a disinterested social planner behind the VoI and financial consequences.⁵ In treatments relevant to our study, Traub et al. (2005, 2009) present tasks in which participants rank

⁵There are a number of studies using hypothetical scenarios and questionnaires adopting this perspective. For example, see Michelbach et al. (2003), Bernasconi (2002), Amiel and Cowell (2000), Amiel et al. (2009).

twelve alternative income distributions for a five-person economy behind the VoI both as a disinterested social planner and as a society member. Both studies conclude that on average the disinterested social planner perspective yields more equity preferring choices. However at the individual level, there is heterogeneity in this comparative static of moving from the disinterested social planner to a society member; about half the participants become more equity preferring, one-quarter more efficiency preferring and the remaining quarter's preferences do not change.

Hong et al. (2015) reports on the only, as far as we are aware, induced budget experiment for a disinterested social planner behind the VoI. In their experiment participants are randomly matched into trios for each of twenty decisions with linear budget sets for a two-person society. The prices range from one to ten and the largest expenditure level is over four times the lowest level. A single decision is selected, then one member of each trio is selected to be the social planner and the other two members are randomly allocated the rich and poor positions; the social planner is given a fixed payment invariant to the budget set. Hong et al. find the GARP consistency of participants' choices are in line with the previous dictator studies. They then estimate a CES-utility function in which an elasticity of substitution less than -1 is interpreted as efficiency preferring and a value greater than -1 as equity preferring. They report about sixty percent of their participants' estimated elasticities reflect efficiency preferring.

There is a large experimental literature which elicits individual social preferences from the perspective of a society member behind a VoI. We discuss the two most relevant studies which address the conflation of the risk attitude and the social preference behind the VoI. Frignani and Ponti (2012) presents an experiment in which a participant participates in twenty-four decision problems in which they select one two-element wealth profile from four possibilities. Their between-participant design has three treatments: a dictator in front of a VoI, a society member behind a VoI, and a lottery treatment where participants only see their own payoff-relevant information. Their analysis starts from the assumption that individuals maximize a mean-variance expected utility function and the variance term coefficient reflects the strength of inequality aversion in the dictator treatment, a mix of risk and inequality aversion in their behind the VoI treatment and risk aversion in their lottery treatment. They find the distribution of the coefficients for the last two treatments are similar and conclude that choices in the VoI treatment are largely determined by risk attitude. They also provide evidence that the distribution of the estimated variance coefficients has a lower mean than those of the other treatments. These results would appear to stand in contrast to ours. However, the between-subject design does not allow for the intrapersonal comparison of an individual's Risk and VoI choices which is key to identifying their social preference. Indeed we also find there is no treatment effect in the aggregate, but that is because the proportions of equity and efficiency preferring individuals are relatively balanced in our sample. Hence our within-subject design, and our

much weaker assumptions on the form of utility, generates a more informative assessment of individual preferences.

Schildberg-Hörisch (2010) conducts an experiment in which a participant completes a pair of consumer problems like we pose, choosing a contingent claim portfolio and a wealth profile behind a VoI from the same budget set. However, she only considers a single budget set, which has a price of two. A participant receives three payments; the result of the Risk problem, then once as the decision-maker for the VoI and then again as the “other” society member for a different participants’ VoI decision. She observes a difference between the average choices in the Risk and VoI problems, participants allocate more to the poor in the VoI than they do to the Low state in the Risk problem. These results are consistent with our findings, but with the lack of budget set variation there is no ability to evaluate individual-level preferences and the validity of the result is limited to the single price and expenditure level.⁶

Finally, the Risk problem component of our instrument is closely related and contributes to the literature on eliciting risk preferences using linear budget sets experiments. Seminal papers by Choi et al. (2007a,b) introduce an instrument by which individuals select contingent claim portfolios from a random sequence of fifty budget sets. These budgets sets have prices mostly ranging between one-half and two (states did not have the High-Low convention we adopted). Their study consists of treatments in which the states are equally likely and in which one state was twice as likely. While they find individual preferences are strongly consistent with GARP they are heterogeneous. However, many individuals’ choices are not consistent with expected utility; for example not fully insuring when the price is one or selecting portfolios that violate first order stochastic dominance. To accommodate such observations they cleverly use a generalized version of the Gul (1991) disappointment aversion model. Choi et al. (2014) apply this instrument to a representative sample of the Dutch population to demonstrate its external efficacy and find that the level of consistency with GARP is lower than that found with U.S. college students. They also find this consistency is positively correlated with household wealth and negatively correlated with age.

3 Experimental design

We conducted our experiment in the Financial and Experimental Economics Laboratory (FEEL) at Xiamen University. A total of 92 subjects were recruited via ORSEE (Greiner (2004)) and came from a broad range of majors. There were 9 sessions and each one lasted around 100 minutes. The average payment was 55 CNY, not including a 10 CNY show-up

⁶Note this study also has a treatment in which participants make a pair of Dictator and VoI decisions. She finds the average amount given to the poor is less in Dictator problem than in VoI problem. Becker et al. (2013) report similar results based on many more rounds of choices.

fee. At the time of the experiment, the exchange rate was approximately 1 USD = 6.2 CNY.

A session consists of two stages: forty rounds each of a VoI and Risk problem.⁷ Each VoI problem has a matched Risk problem with the same budget set. We construct the forty budget sets by taking the Cartesian product of the set of prices $q \in \{1, 2, 3, 4, 5\}$ and the set of expenditures $z \in \{50, 80, 110, 140, 170, 200, 230, 260\}$. We index all budget sets by (q^j, z^j) with $j = 1, \dots, 40$, and (x_i^j, y_i^j) denotes a chosen allocation from budget set j for choice problem i with $i \in \{\text{Risk, VoI}\}$.

In our computer interface, a participant selects a point using a dynamic pie chart and an accompanying table which she controls with a slider. Figure 2 shows the interface for a typical choice round. The pie indicates both the monetary amounts and shares allocated to High (red) and Low (blue) rewards. The table on the right hand shows the current and nearest divisions of the pie (Low% and High%), the amount of Low and High reward (Low\$ and High\$), and the current size of the pie (Total\$). A participant adjusts allocations by moving the green triangle along the slider, which has a precision of 1%, to the desired position. When she adjusts the slider, the High/Low reward shares and the size of the pie adjust accordingly. Once the participant determines her most favored allocation for the current problem, she can click the “Confirm and Leave” button and start a new round.

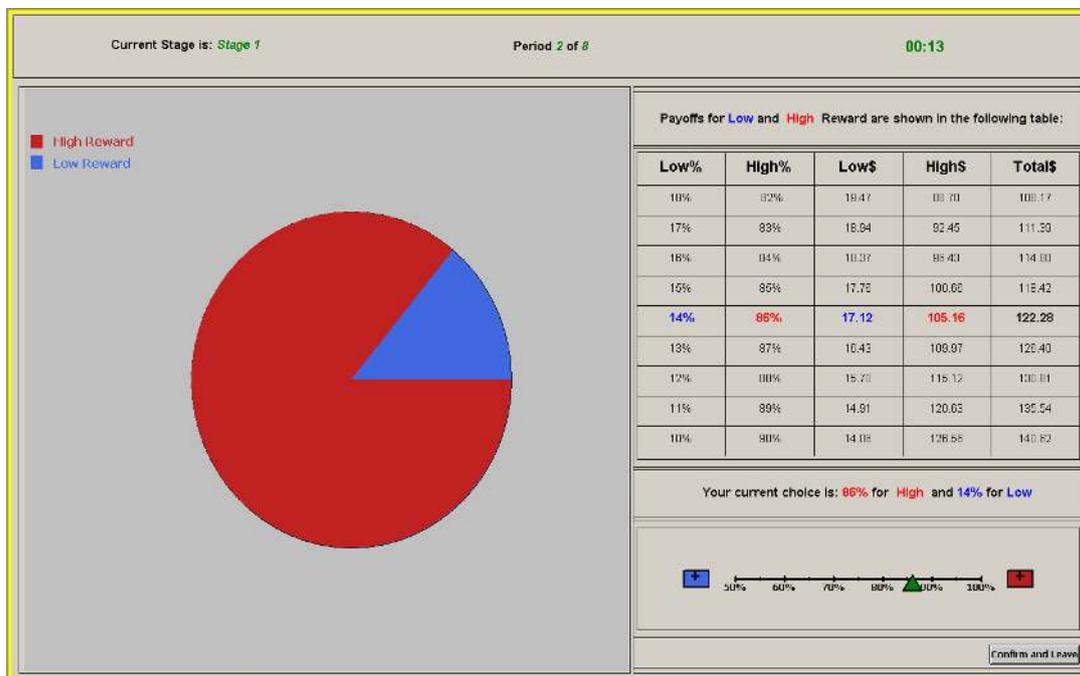


Figure 2: A typical decision screen interface

⁷A subject experiencing two stages is equivalent to a within-subjects experimental design with the VoI and Risk tasks representing two treatments. Within-subject designs sometimes have drawbacks from between-subject designs such as order effects and payoff hedges. However, the crux of this experiment and the variables of interest are differences within an individual, making a within-subjects design a necessity.

To control for potential framing effects we use two frames to describe the tasks. In one frame the slider’s range is 0% to 50%, and in the other frame it is 50% to 100%, indicating proportions allocated to Low or High reward respectively. We set a random initial slider position each time when a participant makes a choice. We also control for order effects by setting half of the sessions as in the sequence VoI-Risk, and the other half as Risk-VoI.

A session with the High reward framing and the Risk-VoI order proceeds through the following steps. First, we collect participants’ informed consents. Second, a monitor reads the stage one instructions in Appendix A.1 aloud while participants silently read along from a hard copy. After reading the instructions, we answer any questions and then participants have to successfully complete a simple quiz to ensure their comprehension. Third, participants complete their stage one decisions and then are given a short break. Fourth, after the break, we distribute the stage two instructions as in Appendix A.2 and repeat the process. Notice that in this VoI stage instructions, we explain to the participant she will be randomly and anonymously matched with another participant in the same session. Before determining payoffs, we ask subjects to complete a short computerized survey. We pay participants for only one of their eighty decisions, and we randomly select this decision round after the completion of the two stages.⁸ If we select a Risk decision round, a participant experiences an individual fair coin toss which determines whether she receives the High or Low reward. If a VoI decision round is chosen, the computer first displays decisions made by both the participant and her counterpart and then randomly selects one of the two decisions to implement. After that, we proceed to determine who of the paired participants receives High/Low reward as if we have selected a Risk decision round.

4 Reduced form analysis

In this section, we evaluate whether participants choose different allocations for Risk and VoI problems. In particular, we assess whether aggregate and individual level data are consistent with one of the three types of social preferences: egoist, equity preferring, or efficiency preferring. Figure 3 provides an array of scatter plots for three subjects’ choices and corresponding budget sets. Each of these subjects corresponds to one of the preference types. Subject 48, whose choices are shown in the top row of the array, is an egoist. She almost allocates the same amount of money in all forty paired problems. Subject 4’s choices, in the middle row of the array, exhibit a taste for equity. In Risk problems, she allocates a vast majority of her expenditure to the Good state but always chooses nearly

⁸To avoid wealth effects in experiments, a common practice is to pay participants based on their decisions in a random round; however, it is also noted that diminishing the probability of any one choice solely determining earnings can undermine financial incentives. Discussions and examination of the two approaches are directly addressed in studies such as Cox et al. (2015), Azrieli et al. (2018) and Charness et al. (2016).

equal wealth distributions in VoI problems. Subject 24's choices, in the bottom row of the array, exhibit a taste for efficiency. In Risk problems, she always insures heavily even when the price is high, but then she selects lopsided wealth profiles in VoI problems. In general, most subjects' choices show a clear inclination towards one of the three categories. However, within the same category decision patterns are highly heterogeneous.⁹

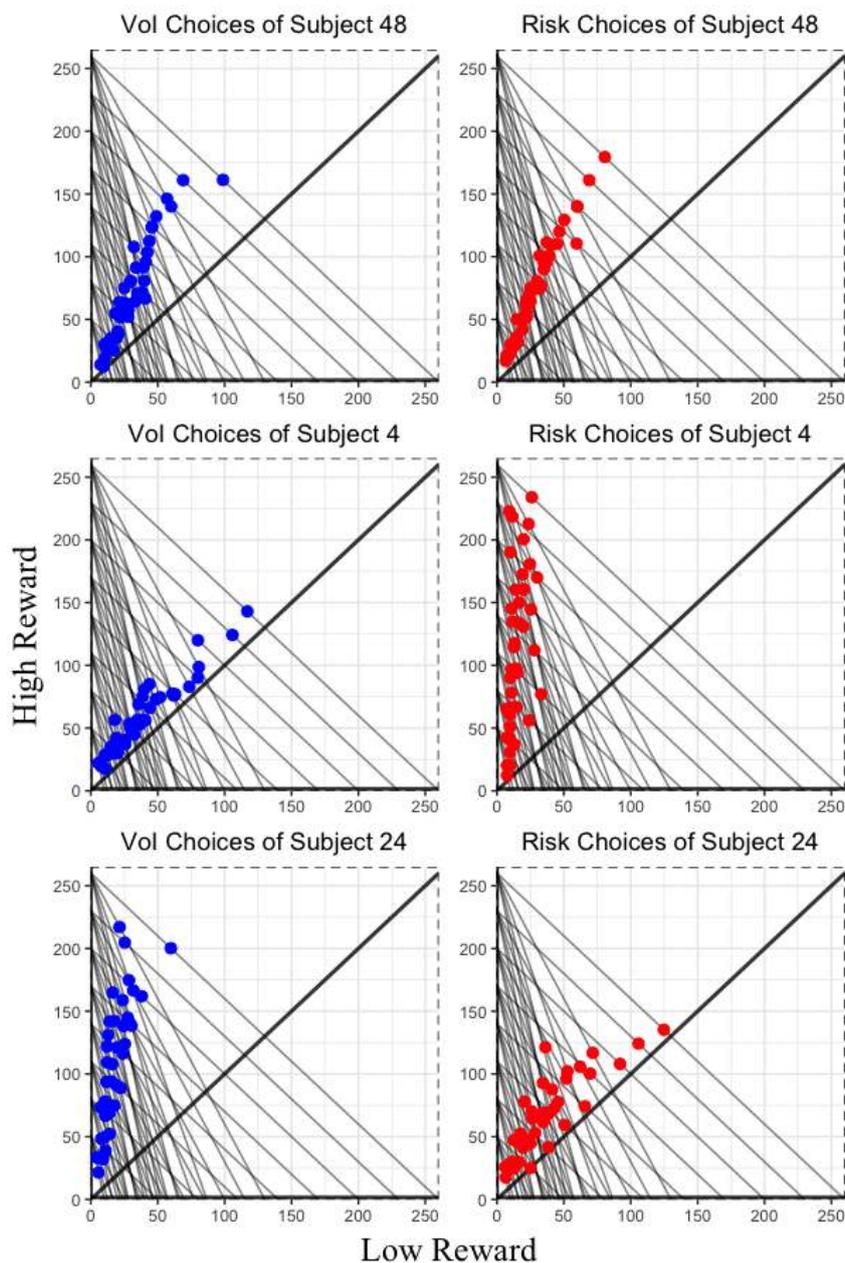


Figure 3: Illustrative examples for three types of choices: egoist (Subject 48), equity preferring (Subject 4) and efficiency preferring (Subject 24)

We proceed by providing statistical evaluations of the treatment effects at multiple levels. In these evaluations, the dependent variable is the paired difference between

⁹The scatter plots for all 92 participants' choices are available from authors on request.

amounts allocated to the Poor individual and to the Bad state, $x_{\text{VoI}} - x_{\text{Risk}}$, in the VoI and Risk problems respectively.

4.1 Aggregate and individual treatment effects

Before aggregating the data, we find neither framing nor order effects in participants' decisions.¹⁰ At the fully aggregated level, we find no significant difference in allocations of Risk and VoI problems; just as Frignani and Ponti (2012). Table 1 reports means and standard errors of monetary and proportional amounts allocated to the Bad and Poor state. Across forty rounds of paired problems, participants on average allocate 30.38 to the Bad state and 30.26 to the Poor member; the difference is statistically insignificant according to a pairwise t -test and a Wilcoxon signed-rank test. Similarly, the average proportion chosen for Low reward account is 27.34% and 27.73% for the Risk and VoI treatments respectively; this difference is statistically significant.¹¹ However, in our judgment this is not economically significant. Since proportional allocations cannot fully capture the influence of varying budget sets, we focus on the monetary amounts going forward.

Variable	Monetary Allocation		Proportional Allocation	
	x_{VoI}	x_{Risk}	$x_{\text{VoI}}\%$	$x_{\text{Risk}}\%$
Mean	30.26	30.38	27.73	27.34
SD	26.09	25.75	15.45	14.69

Table 1: Summary statistics of choices in the VoI and Risk treatments averaged across budgets and individuals

These null results at the aggregate level mask important heterogeneity among participants. Figure 4 reports the mean difference of $x_{\text{VoI}} - x_{\text{Risk}}$ and the corresponding 95% confidence interval for each participant. We plot these means and confidence intervals by stacking from the lowest mean to the highest and then capping the stack with the mean and confidence interval for the aggregate data. This figure allows us to identify three clusters of individuals. First, there are 51 participants who we can not reject as an egoist as their confidence intervals straddle zero. This cluster includes both those who are truly egoist and those for whom this hypothesis test result is a type two error; their unconditional choices have too much variation to precisely measure their non-zero mean. For now, we proceed to label this cluster as egoist. Second, there are 24 participants whose confidence intervals lie above zero. This is evidence for equity-preferring, and we label

¹⁰This result is confirmed by both the t -tests and Wilcoxon rank-sum tests on amounts allocated to the Bad state and the Poor individual. With respect to the framing effects, the p -value is 0.20 (0.22) for the t -test and is 0.09 (0.14) for the Wilcoxon rank-sum test on Risk (VoI) choices. For the order effects, the p -value is 0.17 (0.13) for the t -test and is 0.07 (0.32) for the Wilcoxon rank-sum test on Risk (VoI) choices.

¹¹The p -value is 0.08 for the t -test and 0.02 for the Wilcoxon signed-rank test.

this cluster Equity. Third, there is a cluster of 17 participants whose confidence intervals lie below zero. These participants' choices provide evidence for efficiency-preferring and we label this group Efficiency.¹²

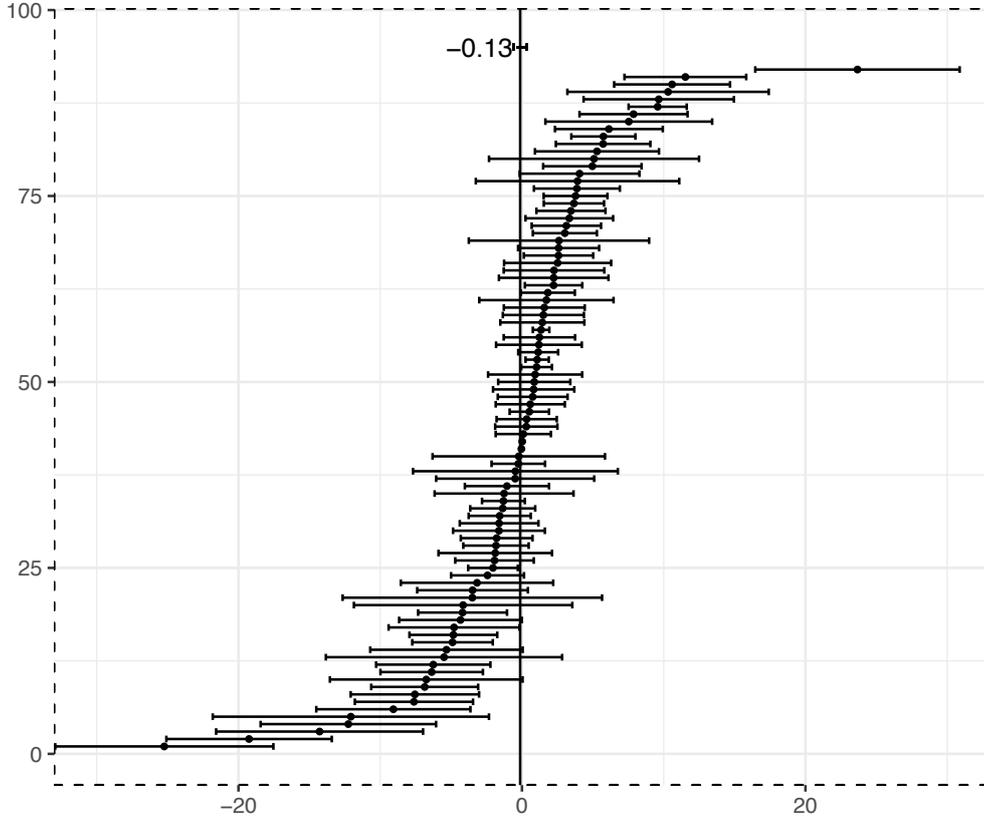


Figure 4: The mean and 95% confidence intervals of the paired difference between Poor and Bad state allocations, $x_{\text{VoI}}(q, z) - x_{\text{Risk}}(q, z)$, for each participant. These confidence intervals are calculated based on paired t -tests for each individual. We stack them from the lowest mean to the highest. The confidence interval for the aggregate data is then stacked at the top.

4.2 Reduced form demand estimation

We estimate reduced form demand functions for the Bad/Poor allocation for each of the three groupings — Equity, Egoist, and Efficiency. Table 2 reports estimated coefficients and robust standard errors clustered at the individual level. We first estimate a simple model where the dependent variable is the demand of the monetary allocation to the Low state, x_i , and the independent variables are the constant, price, and expenditure. We interact each of the terms with a dummy variable for the VoI tasks, D_{VoI} , to create three new factors. The original coefficients are estimates for the Bad state demand function and the interaction terms' coefficients are the estimated deviations for the Poor allocation demand function. We then run a second set of models in which we add a dummy variable

¹²In Appendix B, we provide group classifications based on eight alternative criteria.

$D_{p=1}$ for the price when it equals one. For such choice problems, the price of insurance is actuarial fair in the Risk treatment and there are one-for-one wealth transfers between rich and poor, ergo aggregate wealth is constant, in the VoI treatment. We speculate that behavior takes an abrupt change at this price. This speculation is justified as the intercept terms are greatly reduced, price coefficients more reasonable, and the new coefficients are highly significant for all three groupings. We proceed using the results of this second set of models.

There is evidence that individuals' risk preferences are uncorrelated with their efficiency-equity preferences. Inspection of the estimated coefficients for the Risk treatment shows that they are very close in value for the three sub-groupings of social preferences. We conduct an F -test to evaluate the joint null hypothesis that the three groups' reduced demand functions in the Risk treatment are the same. We fail to reject this null with an F -statistic of 11.79, which is asymptotically χ^2 distributed with eight degrees of freedom and has a p -value of 0.16. This implies that there is no correlation between the social preference categorization and the attitude towards risk.

The coefficients of the Risk demand functions are sensible. Increasing the price of the Bad-state contingent claim by one dollar leads a strong re-balancing of portfolios towards more wealth in the Good-state. Expenditure coefficient estimates also suggest that for every increase of 30, our expenditure step size, only 6 will be used to purchase Bad-state contingent claims and 24 will be added to the Good-state component of the portfolio.

There are notable differences between the estimated VoI demand curves of the Efficiency and Equity groups. First, and almost inevitable given the construction of the sub-samples, the price intercept is greater for the Equity group's VoI demand relative to that of the Efficiency group. Surprisingly when we fix expenditure, the Efficiency group's VoI demand curve is less price sensitive than the Equity group's one. However, when expenditure increases the Equity group's generosity increases at a greater rate than the Efficiency group's. With an expenditure step of 30, the Equity group uses about 7.2 to create a transfer to the Poor, while the Efficiency group uses only 4.2.

The statistical analyses show that there are roughly three types of social preference categories. However, even within the same type, especially the Egoist type, there is a large variation in individual choices. For some subjects, their confidence intervals are quite narrow and we may safely categorize them as egoists. However, many subjects' mean choice differences are consistent with one type but are erroneously classified as egoist due to their wide confidence interval. Some subjects may respond more strongly to expenditure variations or to price variations leading to these wider confidence intervals. To better capture these latent structures, and avoid the myriad of other questionable assumptions underlying these analyses, we utilize nonparametric revealed preference techniques in Section 5 below.

Dependent variable: x_{VoI} and x_{Risk}						
	Equity	Egoist	Efficiency	Equity	Egoist	Efficiency
Constant	28.32** (1.77)	29.79** (1.40)	29.45** (2.33)	11.34** (1.29)	11.90** (1.35)	14.90** (1.20)
Price	-9.57** (0.76)	-10.06** (0.58)	-10.24** (0.98)	-5.32** (0.52)	-5.59** (0.45)	-6.60** (0.59)
Expenditure	0.20** (0.01)	0.20** (0.01)	0.21** (0.01)	0.20** (0.01)	0.20** (0.01)	0.21** (0.01)
Constant* D_{VoI}	5.13** (1.21)	0.08 (0.97)	-9.06** (2.28)	5.13** (1.21)	0.08 (0.97)	-9.06** (2.28)
Price* D_{VoI}	-1.93** (0.42)	0.11 (0.41)	3.71** (0.91)	-1.93** (0.42)	0.11 (0.41)	3.71** (0.91)
Expenditure* D_{VoI}	0.04** (0.01)	-0.004 (0.01)	-0.07** (0.01)	0.04** (0.01)	-0.004 (0.01)	-0.07** (0.01)
$D_{p=1}$				21.22** (2.03)	22.37** (2.08)	18.20** (2.49)
Observations	1,920	4,080	1,360	1,920	4,080	1,360
R ²	0.72	0.62	0.58	0.72	0.62	0.58
Adjusted R ²	0.72	0.62	0.58	0.72	0.62	0.58
Note:	*p<0.05; **p<0.01					

Table 2: Estimated reduced form demand curves for Low/Poor wealth allocations by group classification. OLS regressions with robust standard errors clustered at the individual level. The classification is based on two-sided t -test with 95% confidence level.

5 Nonparametric analysis

We readdress how effectively our instrument elicits efficiency-equity preferences using nonparametric revealed preference techniques. We first test whether participants' choices, made by both Risk- and VoI-selves, are rational in terms of maximizing a monotonic and concave utility function, and if so, whether that function is homothetic. We also check the power of our experimental design and compare it with the power of other induced budget experiments. We then evaluate the relative concavity of each participant's VoI and Risk utility functions and classify participants into different social preference types. We re-estimate the reduced form demand functions according to this classification to further validate our result on the independence of risk and social preference.

5.1 Testing rationality and homotheticity

We test consistency of each participant’s choices with GARP to address the question of rationality.¹³

For price taking DMs and linear budget sets, Afriat (1967) and Varian (1982) show that satisfying GARP is both a necessary and sufficient condition for the existence of a continuous, monotonic, and concave utility function that is maximized by the observed choices. This condition imposes no parametric structure on choice data and is easily testable based on implicit revealed preference relationships. It is a commonly employed test for induced budget experiments.¹⁴

It is common to find GARP violations in choice data. In response, researchers have proposed various ways to measure the extent choices violate GARP. Afriat (1967) propose the Critical Cost Efficiency Index (CCEI), the minimum scalar multiplier one needs to apply to the set of expenditure levels to resolve all GARP violations of a DM. The CCEI is bounded between zero and one. The higher the CCEI of a DM’s choices, the more rational she is.¹⁵ The CCEI can be interpreted as a measure for wasted income: A DM with a CCEI of e could have obtained the same level of utility by spending only a fraction of e of what she actually spent to obtain that level of utility.

We also test choices for consistency with maximizing an increasing homothetic utility function.¹⁶ We do this for two reasons. First, while we believe making allocation choices via a dynamic pie chart is more intuitive than via a graphed budget line, it potentially makes two decision heuristics focal: constant High/Low pie shares and constant budget shares.¹⁷ When a DM uses either heuristic, her choices necessarily satisfy GARP. Second, it is an important validation check when one estimates homothetic utility functions using data from induced budget experiments. For example, Andreoni and Miller (2002) and

¹³In our experiment, when a DM chooses (x^j, y^j) from a budget with price q^j and expenditure z^j she directly reveals that she prefers (x^j, y^j) over all other (x^k, y^k) for which $q^j x^k + y^k \leq z^j$. If the inequality is strict (i.e., $q^j x^k + y^k < z^j$) we say that (x^j, y^j) is strictly directly revealed preferred over (x^k, y^k) . Furthermore, if she chooses (x^k, y^k) from the budget with q^k and z^k , she indirectly reveals that she also prefers (x^j, y^j) to all bundles affordable under q^k and z^k . This is how the indirect revealed preference relation is constructed through the transitive closure of the direct revealed preference relation. GARP states that if a choice (x^j, y^j) is (directly or indirectly) revealed preferred to a choice (x^k, y^k) , then (x^k, y^k) must not be strictly directly revealed preferred to (x^j, y^j) . Varian (1982) shows that GARP is equivalent to the existence of a continuous, monotonic, and concave utility function u which rationalizes the choices, that is, $u(x^j, y^j) \geq u(x, y)$ whenever (x^j, y^j) is revealed preferred over (x, y) .

¹⁴Some notable examples are Cox (1997), Sippel (1997), Harbaugh et al. (2001), Andreoni and Miller (2002). A recent theoretical justification for the validity of using experiments to test utility maximization via GARP was provided by van Bruggen and Heufer (2017).

¹⁵We say that (x^j, y^j) is directly revealed preferred over (x^k, y^k) at efficiency level e if $q^j x^k + y^k \leq e z^j$, where $e \in [0, 1]$. GARP at efficiency level e can be defined accordingly. The CCEI of an individual is the greatest efficiency level e at which that individual’s choices satisfy GARP.

¹⁶A utility function is homothetic if it is a positive monotonic transformation of a utility function that is homogeneous of degree one.

¹⁷Choices made according to a constant High/Low pie share heuristic lie on a single ray through the origin. Choices made using a constant budget shares heuristic lie on a family of rays through the origin, each ray corresponding to a particular price ratio.

Fisman et al. (2007) estimate CES utility functions and Choi et al. (2007a) estimate CRRA utility functions.

Following Heufer (2013), we evaluate participants' choices for consistency with the Pairwise Homothetic Axiom of Revealed Preference (PHARP).¹⁸ Heufer (2013) shows that this condition is equivalent to the constrained maximization of a homothetic utility function in two commodity choice settings. We also calculate Heufer and Hjertstrand's (2017) Homothetic Efficiency Index (HEI) can be interpreted in analogy to the CCEI as a measure of wasted income on resolving violations of PHARP.

While efficiency indices have appealing interpretations, there is no natural scale to measure irrationality. When a DM makes "errors," her efficiency level will depend on the number of budget sets and the variations in both price and expenditure. Bronars (1987) suggests an exercise that simultaneously provides a measure of power for a sequence of linear budget sets and benchmarks for alternative rationality thresholds for an efficiency index. One simulates a large number of DM choice sequences — we will use 10,000 — in which each choice is a random draw from the uniform distribution over the budget line. Then one constructs an empirical density of the resulting efficiency indices. For a given rationality threshold, the proportion of simulated sequences exceeding it is a measure of the corresponding power. Also, the densities of indices from the simulated data serve as benchmarks for those of actual DMs. We report two power calculations: one using uniform random draws from budget lines, and another using uniform draws restricted to the part of the budget above the 45° line. The latter restriction mimics the restriction imposed on the participants in our experiment.

5.2 Rationality results and comparisons with other studies

We now present the results for rationality checks and the power of our experimental design, and compare these with various induced budget experiments.

5.2.1 Rationality and homotheticity of participants' choices

By problem type, we check every possible pair of a participant's choices for GARP violations and count them.¹⁹ The test is binary: a choice pair either satisfies GARP or not. Overall, the number of GARP violations is small, with a median of 3 and 4 respectively for the Risk and VoI problems. Over 30% of participants commit zero or only one violation of GARP. There are 11 participants who fully satisfy GARP in the VoI problems and 12 in the Risk problems; 4 satisfy GARP in both problem types. With respect to the axioms

¹⁸PHARP states that for any distinct pair of choices, indexed by j and k , $(q^j x^k + y^k) (q^k x^j + y^j) \geq z^j z^k$ holds.

¹⁹For each treatment, there are 780 pairs of such choices.

regarding homotheticity, there are 3 participants in both the VoI and the Risk problems who satisfy PHARP; only one of them satisfies HARP in both problem sequences.²⁰

Next, by problem type, we calculate the CCEI and HEI for each participant and report summary statistics for the resulting distribution of values in the first two rows of Table 3. The mean CCEIs are 0.96 and 0.95 for the Risk and VoI problems respectively, and the median CCEIs are both 0.98. Also, 73%, 87%, and 99% of the participants' Risks (VoI) CCEIs exceed the respective thresholds of 0.95, 0.90, and 0.80.²¹ While the majority of participants commit at least one GARP violation, there is compelling evidence almost all are highly rational. Most participants' violations are not severe because small adjustments in expenditures brings full compliance with GARP. In contrast, the HEI results suggest that many participants do not make choices consistent with homothetic preferences. The average HEI is 0.85 in both problems, and only about one-third of the HEIs, 32% and 36% for Risk and VoI problem respectively, exceed the threshold of 0.90.

5.2.2 The power of the collection of linear budget sets

We report on the distributions of the two efficiency indices generated by Bronars's suggested simulation — adjusted for the budget sets of our instrument. These distributions provide more context to our claims of high rationality and an evaluation of the power of our linear budget sets. Figure 5 compares the distribution of CCEI for our 92 participants and for 10,000 hypothetical participants who choose randomly from the part of the budget above the 45° line. The consistency levels of the participants are overwhelmingly higher than those of random choice. Figure 6 similarly compares the distribution of HEI between real and simulated participants. While the HEI values of the participants are considerably lower than their CCEI values, the figure shows that the participants still have much higher homothetic efficiency than the random choices.

The third row of Table 3 reports summary statistics for the distributions of CCEIs and HEIs generated by simulated choices. These values are strikingly lower than those for the Risk and VoI problems. The simulated HEI rarely exceed any of our benchmark thresholds. In the fourth row we report the same summary statistics for simulations of random draws from a uniform distribution over the portion of the budget line above the 45° line. There is little evidence of this censoring leading to critically low test power, except perhaps for the CCEI at the 80% benchmark. We confidently rule out the possibility of random behaviour generating the observed high consistency levels.

²⁰We do not report violation counts for PHARP as it only considers pairwise violations; a count of violations is therefore not comparable to a count of GARP violations.

²¹There is no consensus in the literature regarding the thresholds for efficiency indices. For example, Andreoni and Miller (2002) uses a CCEI threshold of 0.95 for their dictator game with eight budgets, while Fisman et al. (2007) omit participants with a CCEI below 0.8 from further analysis in their dictator game with fifty budgets.

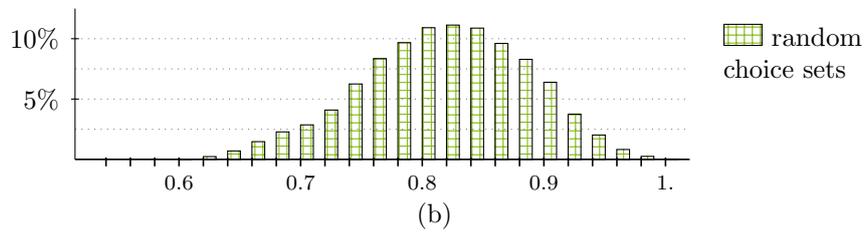
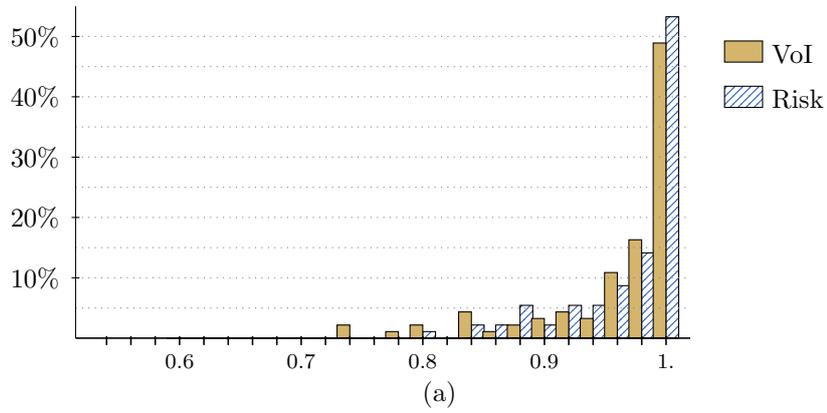


Figure 5: Histogram of the CCEI (a) for 92 subjects and (b) for 10,000 random choices from uniform distributions over the portion of the budget set above the 45° line.

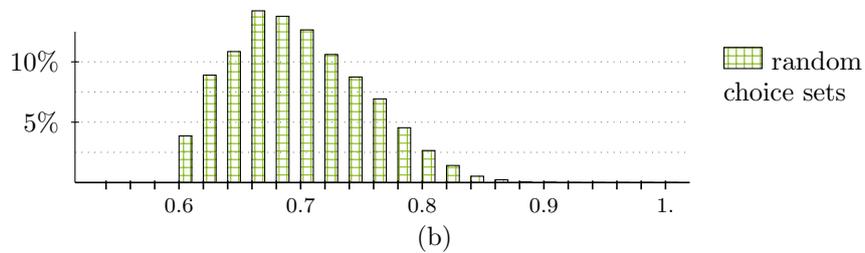
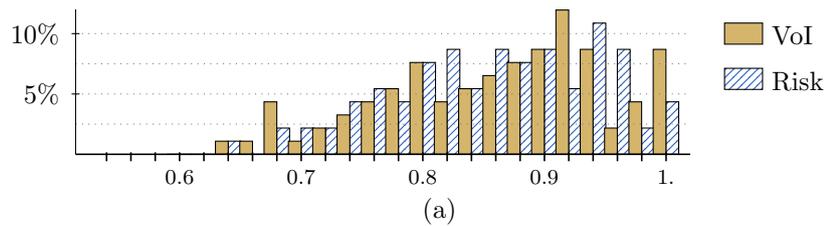


Figure 6: Histogram of the HEI (a) for 92 subjects and (b) for 10,000 random choices from uniform distributions over the portion of the budget set above the 45° line.

Study	Experiment	Decision maker	CCEI					HEI				
			Mean	Median	≥ 0.95	≥ 0.90	≥ 0.80	Mean	Median	≥ 0.95	≥ 0.90	≥ 0.80
Current	Risk VoI	Subjects	0.96	0.98	73%	87%	99%	0.85	0.86	11%	32%	72%
		Subjects	0.95	0.98	73%	84%	95%	0.85	0.87	13%	36%	70%
	Random	0.75	0.75	0%	3%	30%	0.60	0.59	0%	0%	0%	
	Random above 45°	0.80	0.81	1%	7%	53%	0.68	0.68	0%	0%	2%	
Choice under risk	CFGK	Subjects	0.94	0.97	68%	81%	89%	0.83	0.86	19%	38%	68%
		Subjects censored at 45°	0.96	0.98	79%	87%	94%	0.85	0.87	19%	40%	72%
	Random	0.65	0.66	0%	0%	6%	0.52	0.52	0%	0%	0%	
	Random above 45°	0.83	0.84	1%	16%	71%	0.70	0.71	0%	0%	9%	
Dictator giving	FKM	Subjects	0.88	0.93	45%	58%	76%	0.78	0.80	22%	33%	50%
		Subjects censored at 45°	0.95	0.98	67%	81%	96%	0.86	0.87	28%	43%	70%
	Random	0.72	0.74	1%	6%	30%	0.58	0.58	0%	0%	1%	
	Random above 45°	0.89	0.90	19%	48%	88%	0.75	0.76	0%	2%	30%	
Dictator giving	FKM	Subjects	0.96	0.99	71%	87%	97%	0.88	0.92	41%	55%	82%
		Random	0.60	0.61	0%	0%	4%	0.47	0.47	0%	0%	0%
	Subjects	0.99	1.00	98%	99%	99%	0.98	1.00	83%	94%	99%	
	Random	0.87	0.90	37%	50%	74%	0.76	0.77	4%	12%	41%	

Table 3: Comparison of efficiency indices and test power results for various comparable experiments: FKM (Fisman et al., 2007), AM (Andreoni and Miller, 2002), CFGK (Choi et al., 2007a) and CKMS (Choi et al., 2014).

5.2.3 Test power and efficiency levels compared to other experiments

We compare our efficiency indices and corresponding power assessments with those of other similar budget experiments, namely the decision under risk studies of Choi et al. (2007a; CFGK) and Choi et al. (2014; CKMS) as well as the dictator studies of Andreoni and Miller (2002; AM) and Fisman et al. (2007; FKM). These experiments, including ours, differ in the subject pool, the number of budgets, and the price and expenditure variations; all of which may affect economic indices and test powers. In particular, our experiment restricts participants' choices to be above the 45° line, which may deprive chances of violating GARP and PHARP and thus result in higher indices. To better evaluate this concern, we also additionally impose similar restrictions to the data of CFGK and CKMS. Note that CFGK considers two different priors on the Good state of the world; probabilities $\frac{2}{3}$ and $\frac{1}{2}$. We only include data on the latter in our analysis. Further note that in CFGK and CKMS either good could be the cheaper one, while in our framework, good x is never cheaper than y . Accordingly we first transform CFGK's and CKMS's data by labeling the cheaper good as y . We then censor choices below the 45° line by replacing them with a choice on the 45° line, thereby "correcting" choices that violate first order stochastic dominance (FOSD).²²

Overall our experiment's power and participant's rationality are in line with these other experiments. With respect to the choice under risk studies, inspection of the middle rows of Table 3 reveals our participants' CCEIs are slightly higher for the mean, median and all threshold levels when compared to the original CFGK and CKMS data, and slightly lower when compared to the censored data. As we ex-ante believed our design would nudge participants to be more homothetic, is it somewhat surprising that our participants do not universally exhibit higher homothetic efficiency; the differences are relatively minor. The power of all of these designs is generally very high.

The results for the two dictator games are quite different due to the large difference in the number of budgets; FKM asked participants to make 50 choices, while AM's participants only made 8 choices. Accordingly, for the CCEI, the results for FKM are similar to our results for the VoI part. With respect to the HEI, our participants do considerably worse. Part of the reason is that many participants in dictator games keep all or almost all of the money to themselves which reduces the chance of violating PHARP. Our participants have lower efficiency compared with AM, which can be explained by the fact that the power of our experiment is substantially greater.

²²For dictator games, this restriction does not make sense, as ungenerous dictators have no reason to allocate higher amounts to the other recipient even if it is relatively cheap to do so. Therefore, we do not impose this restriction on FKM and AM. Also when we report power calculations for linear budget dictator game studies we only report on random selection with the full budget line as the support.

5.3 The relative convexity of Risk-self versus VoI-self preferences

We return to the idea of classifying individual participants' social preferences into the categories of equity preferring, efficiency preferring, and egoist. We refine these classifications based on a nonparametric approach proposed by Heufer (2014). This approach allows the evaluation of the relative global convexity of two families of indifference curves when each is used to generate optimal choices against the same budget sets. We use this approach to evaluate the relative convexity of the family of indifference curves — i.e. concavity of the utility function — for each participants' Risk- and VoI-selves. Note that the more concave a DM's Risk-self utility function is relative to that of her VoI-self, the smaller the difference between her allocation to the Poor individual in a VoI problem and her allocation to the Bad state in a Risk problem when facing the same budget set.

Application of this approach requires an additional rationality condition on a set of linear budget set choices: a DM's choices must be consistent with *second order stochastic dominance* (SOSD). A lottery (wealth profile) A has SOSD over a lottery (wealth profile) B if and only if every expected (average) utility maximizer prefers A over B (Hadar and Russell, 1969). A DM's choices are consistent with SOSD if B is never revealed preferred over A whenever A has second order stochastic dominance over B .²³

We check each of our participant's choices, by problem type, for consistency with SOSD-GARP and compute the SOSD-CCEI as defined by Heufer (2014), which are the SOSD analogies to GARP and the CCEI. We find that 11 (9) subjects in the Lottery (VoI) treatment satisfy SOSD-GARP compared to 12 (11) subjects who satisfy GARP. The SOSD-CCEI values are exactly the same as the CCEI values for all but three subjects who satisfy GARP but not SOSD in one of problem types. This is not surprising as subjects were required to make choices above the forty-five degree line, making it difficult to violate stochastic dominance without also violating GARP. As in Heufer (2014) we use efficiency-adjusted revealed preference relations in order to meet the condition.

The key concept for our approach is a revealed binary relation we call Partially Revealed More Convex (PRMC), where convex refers to the shape of the indifference curves.²⁴ We say a DM is Risk-PRMC at an observed Risk choice, denoted A , if there exists at least one allocation, denoted B , from the DM's set of VoI choices such that

- (i) the DM's expected wealth level from the wealth profile B is higher than her expected wealth of A ;
- (ii) A is revealed preferred to B by the Risk-self;
- (iii) B is revealed preferred to A by the VoI-self; and,

²³Note that consistency with SOSD does not require consistency with expected utility.

²⁴The original approach in Heufer (2014) was developed to compare the risk aversion of two individuals. PRMC corresponds to Heufer's partially revealed more risk averse (PMRA) relation.

(iv) at least one of those two revealed preference relations is strict.

VoI-PRMC is analogously defined.

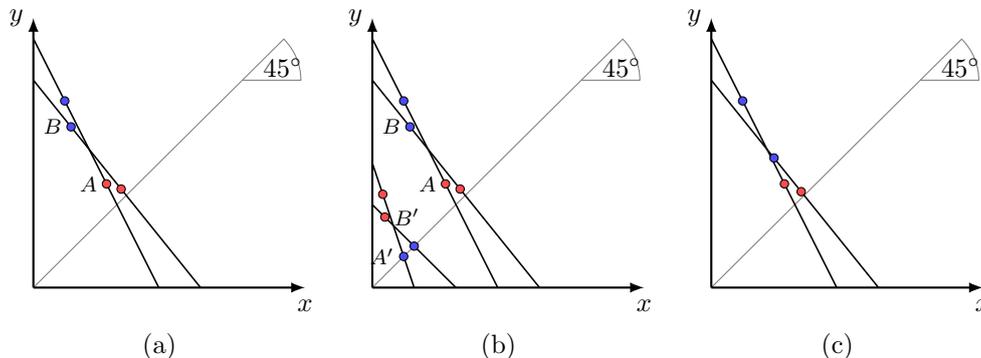


Figure 7: Examples of PRMC relationships. Choices made in Risk and VoI problems are in red and blue, respectively. In (a), the DM is Risk-PRMC at the observed Risk choice bundle A . In (b), the DM is Risk-PRMC at Risk choice bundle A and VoI-PRMC at the VoI choice bundle A' . In (c), the DM's choices show neither Risk-PRMC nor VoI-PRMC.

Figure 7a illustrates a Risk-PRMC relationship through a pair of Risk and VoI choices. The DM's Risk-self prefers A over B even though B has a higher expected value but greater variance. In other words, the DM's VoI-self is “more risk-tolerant” as she is willing to accept a “greater variance in individual wealth” when compensated by sufficiently greater aggregate wealth. This extra “risk tolerance” by her VoI-self reflects a taste for efficiency. Therefore, passing the Risk-PRMC test at A establishes that the indifference curve of DM's Risk-self is more convex than that of her VoI-self at A . The test for VoI-PRMC is analogously defined. If a DM passes the VoI-PRMC test at an observed choice from a VoI problem, her VoI-self is willing to take a more equal but less efficient allocation than her Risk-self, and therefore the indifference curve for VoI-self is more convex at this particular choice.

For each participant we test eighty PRMC relations, forty each for Risk and VoI choices. Each test result is an observation-specific binary indicator, which jointly provides us a global picture of the relative convexity of Risk and VoI-self's preferences. We impose the most strict requirement on these 80 test results and refine our social preference classifications into four categories:

- (i) unorderable: both Risk-PRMC and VoI-PRMC hold at least once;
- (ii) efficiency-preferring: Risk-PRMC holds at least once while VoI-PRMC never holds;
- (iii) equity-preferring: VoI-PRMC holds at least once while Risk-PRMC never holds;
- (iv) egoist: both Risk-PRMC and VoI-PRMC never hold.

In three out of the four cases, we reach a definitive conclusion about a DM's global social preference implied by a straightforward reinterpretation of Theorem 3 in Heufer

(2014). Case (i) is the inconclusive one in which we can conclude a DM’s Risk-self and VoI-self utility functions are distinct but there is no global relative rank of their concavities. Figure 7b provides an example of such un-orderable Risk- and VoI-self preferences. The DM is Risk-PRMC at the Risk choice A , revealing a taste for efficiency (locally); while she is VoI-PRMC at the VoI choice bundle A' , revealing a taste for equity (locally). The un-ordable preferences imply that either (1) people have different notions of the trade-offs in Risk and VoI problems, or (2) people’s preferences are not stable over the entire expenditure range. In case (ii) the DM’s Risk-self choices can be represented by a more concave utility function than the VoI-self choices but not vice versa. Case (iii) is the reverse. Case (iv) implies that the same utility function can represent both selves. For instance, in Figure 7c, we cannot find a PRMC relationship based on the DM’s choices. When all her choices follow such a pattern, the underlying indifference curves for her Risk- and VoI-self are globally similar.

Table 4’s first row summarizes the classification results. Out of our 92 participants, 76 participants (83%) demonstrate a non-trivial social preference and the other 16 participants (17%) are classified as egoist. Among participants exhibiting social preferences, 26 clearly exhibit a taste for efficiency; 30 are equity-preferring; the other 20 participants have distinct Risk- and VoI-self utility functions but we can’t establish a global concavity ranking. Compared to our earlier type clusters based on t -tests, participants in the group of efficiency-preferring and equity-preferring largely overlap. However, the revealed preference approach allowed us to refine many of the those who we labeled as egoist into a pair of clusters: those who are egoist and those having a social preference which is not globally consistent with respect to trade-offs between efficiency and equity. Thus, we provide some resolution to the issue of identifying those individuals for whom the difference between their Bad and Poor quantities demanded is zero from those individuals for whom the difference is not zero but the sign differs by budget sets — and is a possible source of the instances of wide confidence intervals around zero in Figure 4.

Tolerance level	Different			Similar
	Equity preferring	Efficiency preferring	Un-orderable	Egoist
Zero tolerance	30 (32.6%)	26 (28.3%)	20 (21.7%)	16 (17.4%)
.0001	30 (32.6%)	26 (28.3%)	20 (21.7%)	16 (17.4%)
.025	16 (17.4%)	22 (23.9%)	7 (7.6%)	47 (51.1%)
.05	13 (14.1%)	13 (14.1%)	1 (1.1%)	65 (70.7%)

Table 4: Nonparametric categorization of social preference types

Notice our nonparametric classifications are based on participants’ pre-adjusted revealed preference relations. We now evaluate how these classifications change when we allow for varying levels of irrational choice. We allow for more irrationality by slackening

each participant’s budget by SOSD-CCEI by different tolerance levels. For each tolerance level we re-tabulate the preference classifications. For a tolerance limited to a 0.0001 expenditure adjustment, the classification results are reported in the second row of Table 4. As the expenditure slack grows more tolerant, a stronger difference between Risk- and VoI-self preferences are required to pass the PRMC tests. If a DM is persistently classified in one of the different preference groups, we are more certain about her preference type in terms of economic significance. The third and fourth row of Table 4 show that when the tolerance level increases, more participants are classified as egoists. With a tolerance level of 0.05, we only identify distinct Risk and VoI preferences for approximately 30% of the participants.²⁵

5.4 Social preference classification stability

Our nonparametric social preference classifications overcome some of the issues associated with our previous statistical approach. In particular, it allows refinement of the egoist classification into those whose preferences are globally similar and those whose preferences are not comparable. Further, it relaxes potentially invalid assumptions underlying the statistical tests used to generate classifications. To evaluate the impact of these factors, we show how individual participants’ classifications differ under the two approaches. Then we examine how the nonparametric approach impacts estimates of type specific reduced form demand functions.

We first examine how participants’ classifications vary under the two approaches. Table 5 lists the number and fraction of subjects who are classified into different social preference types under both the statistical and nonparametric approach. The rows are the classification results under the statistical t -test, and the columns are those under the nonparametric test. We first note participants with preferences classified as un-orderable originally come from all three preference types. These are pairwise not significantly different from each other under the proportion test. This counters our expectation that those classified as un-orderable would have previously been classified as egoist. We secondly note there is little switching of participants between Equity Preferring and Efficiency under the two classification approaches. However, subjects who are categorized as Egoist under t -test are now almost equally distributed among three social preference clusters. This result confirms that (1) the three social preference types elicited by our experimental

²⁵In unreported analyses, we adjust expenditure levels using an alternative efficiency index, the Varian Efficiency Vector (VEV) (Varian, 1993, Heufer and Hjertstrand, 2017). The VEV corrects the revealed preference relation by relaxing the expenditure level only when necessary, thus retaining more information implied by people’s choices. We follow the same procedure and re-classify our subjects into four preference types. Compared to the classification results in Table 4, more subjects (93.4%) demonstrate distinctive preferences for Risk and VoI problems. All three non-trivial social preference types (equity-preferring, efficiency-preferring, and un-ordable) also show greater tolerance for the expenditure slack. With a tolerance level of 0.05, 20 (21.7%) participants are equity-preferring; 15 (16.3%) are efficiency-preferring; and 16 (17.4%) participants’ Risk- and VoI-self preferences are un-ordable.

instrument are robust; (2) our nonparametric method provides a more refined classification result.

	Nonparametric categorization				Sum
	Equity preferring	Efficiency preferring	Egoist	Un-orderable	
Equity preferring	15 (88.2%)	0 (0.0%)	2 (11.8%)	7	24
Efficiency preferring	1 (7.7%)	11 (84.6%)	1 (7.7%)	4	17
Egoist	14 (33.3%)	15 (35.7%)	13 (31.0%)	9	51
Sum	30	26	16	20	92

Table 5: The cross table for the categorization under the statistical and nonparametric approach. The rows are the classification results based on two-sided t -test, with 95% confidence interval. The fractions in the parentheses are calculated by excluding subjects who are classified as un-orderable.

We conclude this section by reporting in Table 6 the re-estimation of reduced form demand functions for each social preference type under the nonparametric classification. Note that the interaction terms with the treatment dummy are less significant compared to the estimation based on t -test. This is due to that two-thirds of the subjects who were previously egoist are now assigned to the equity and efficiency group, reducing the precision of our treatment effect estimates. However, a joint F -test for the reduced demand functions for the Risk treatment across three clusters (excluding the un-orderable subjects) shows that they are not significantly different. The F -statistic is 13.08, and the p value is 0.11 with a degree of freedom of 8. At the same time, a joint F -test for the reduced demand functions for VoI treatment across three clusters (excluding the un-orderable subjects) shows that they are significantly different. The F -statistic is 40.71, and the p value is 0.00 with a degree of freedom of 6. Thus, our initial evidence that social preference types form distinct clusters but risk preferences are independent of these types persists in this alternative nonparametric approach.

6 Structural equation estimation and analysis

Estimates of structural utility models are useful in counter-factual policy analysis when the assumed utility model accurately approximates DMs' preferences. In this section, we measure the efficiency-equity trade-offs through a parametric approach. We estimate a two-parameter subjective expected utility model for each participant. One parameter reflects the relative subjective prior of the Bad/Poor state. The other reflects the curvature of the utility function. We find the estimates of subjective priors are consistent with overconfidence in the Rich/Good state of the world. We also show that individual differences between estimated Risk and VoI utility curvatures form clusters of social preferences similar to those found in our reduced form and nonparametric analyses.

	Dependent variable: x_{VoI} and x_{Risk}			
	Equity	Egoist	Efficiency	Un-orderable
Constant	11.51** (1.00)	14.93** (2.02)	11.99** (2.08)	11.80** (1.89)
Price	-5.24** (0.46)	-7.51** (0.57)	-5.34** (0.70)	-5.43** (0.60)
Expenditure	0.19** (0.01)	0.23** (0.01)	0.20** (0.01)	0.18** (0.01)
Constant* D_{VoI}	2.44* (0.99)	-2.41 (1.90)	-3.49 (1.97)	1.44 (2.18)
Price* D_{VoI}	-0.76 (0.47)	1.36* (0.64)	1.15 (0.77)	-0.34 (0.91)
Expenditure* D_{VoI}	0.02* (0.01)	-0.02 (0.01)	-0.03* (0.01)	0.002 (0.01)
$D_{p=1}$	18.84** (2.05)	24.61** (3.12)	20.21** (2.61)	23.75** (3.38)
Observations	2,400	1,280	2,080	1,600
R ²	0.65	0.68	0.60	0.59
Adjusted R ²	0.65	0.68	0.60	0.59
Note:	*p<0.05; **p<0.01			

Table 6: Estimated reduced form demand curves for Low/Poor wealth allocations by group classification. The classification is based on the non-parametric approach. OLS regressions with robust standard errors clustered at the individual level.

Finally, we find a similar range of curvature estimates, for both Risk and VoI tasks, as Choi et al. (2014) but greater estimates than Choi et al. (2007b).

6.1 Subjective expected utility specification and estimation approach

We assume a DM's choices for a decision round with relative price q and expenditure z are governed by the solution of the constrained maximization problem

$$\begin{aligned} \arg \max_{(x,y)} \quad & \alpha u(x) + u(y) \\ \text{subject to} \quad & qx + y = z \\ & y \geq x \geq 0. \end{aligned} \tag{1}$$

The parameter α is the ratio of her subjective priors of the Low and High state of the world, $\text{Pr}(\text{Low})/\text{Pr}(\text{High})$.²⁶ When $\alpha = 1$, the DM has an objective prior in which two states are equally likely, and consequently, is an expected utility (EU) maximizer. When $\alpha \neq 1$, the DM's prior differs from the objective one and she is a subjective expected utility (SEU) maximizer. For $\alpha > 1$ she believes the low state is more likely and vice versa when $\alpha < 1$. Note, a DM's subjective prior of the Bad/Poor state is $\alpha/(1 + \alpha)$.²⁷

We assume the utility function for the same DM is the commonly employed power utility function,

$$u(x) = \frac{x^{1-\rho_i}}{1-\rho_i},$$

where $i \in \{\text{Risk}, \text{VoI}\}$. The parameter ρ_i defines the concavity of the DM's problem specific utility function. For the Risk choice problem ρ_{Risk} is a constant coefficient of relative risk aversion. In terms of empirical execution, the power function is not well-defined for boundary allocations $(x_i, y_i) = (0, z)$. However, these choices might be realized when there is an unobserved noise component in the DM's choices. When evaluating these corner choices we replace 0 with a small number ωz , where $\omega = 0.001$.²⁸

²⁶Notice for the VoI task, we are following the arguments of Harsanyi (1953, 1976) that a social welfare function closely approximates to an average expected utility function.

²⁷An alternative interpretation for α is a relative weight the DM attaches to a state with a less favorable outcome while attaining the objective probability over the two states. When $\alpha > 1$, she over-weights the Low state due to disappointment aversion (or loss aversion if assuming equal allocations as reference points) as described in Gul (1991), and vice versa when $\alpha < 1$.

²⁸All following estimation results are robust to reasonable alternative choices of ω .

Solving the maximization problem (1) yields the following equation characterizing the optimal commodity bundle choice,

$$\ln(x_i^*/y_i^*) = f(q, \omega; \alpha, \rho_i) = \begin{cases} \ln(\omega) & \text{if } \ln(\alpha) - \rho_i \ln(\omega) \leq \ln(q), \\ -\frac{1}{\rho_i} [\ln(q) - \ln(\alpha)] & \text{if } \ln(\alpha) < \ln(q) < \ln(\alpha) - \rho_i \ln(\omega), \\ 0 & \text{if } \ln(q) \leq \ln(\alpha). \end{cases}$$

The first case is the corner solution $(x_i^*, y_i^*) = (\omega z, z)$; the second case is an interior solution and the last case is for a corner solution in which $x_i^* = y_i^*$. We use a dummy variable formulation to indicate whether a decision round j is either a Risk or VoI problem. With respect to the utility curvature parameter ρ we define,

$$\rho_{\text{VoI}} = \rho_{\text{Risk}} + D_{\{j=\text{VoI}\}} \cdot \rho_{\text{diff}}.$$

The parameter ρ_{diff} is the difference of the curvature between two utility functions: one recovered from the DM's VoI-self and the other from her Risk-self. With positive ρ_{diff} , the DM is globally more equity preferring and a negative ρ_{diff} implies efficiency-preferring.

We obtain estimates of the DM's utility parameters by minimizing the square sum of the distances between $\ln(x/y)$ and $\ln(x^*/y^*)$, where the former vector is composed of her choices in eighty decision rounds and the latter from the corresponding optimal choices. Specifically, we solve the problem

$$\arg \min_{\alpha, \rho_{\text{Risk}}, \rho_{\text{diff}}} \sum_{i \in \{\text{Risk}, \text{VoI}\}} \sum_{j=1}^{40} \left[\ln \left(\frac{x_i^j}{y_i^j} \right) - \ln \left(\frac{x_i^{j*}}{y_i^{j*}} \right) \right]^2 = \sum_{i \in \{\text{Risk}, \text{VoI}\}} \sum_{j=1}^{40} \left[\ln \left(\frac{x_i^j}{y_i^j} \right) - f(q^j, \omega; \alpha, \rho_{\text{Risk}}, \rho_{\text{diff}}) \right]^2.$$

We estimate belief and curvature parameters by the nonlinear least square (NLLS) method using the Levenberg-Marquardt (Moré, 1978) algorithm.²⁹

6.2 Individual level parametric estimation

We estimate structural models for 64 (70%) participants and exclude 28 participants. First, we omit 6 participants whose choices are not sufficiently consistent with utility maximization as indicated by CCEI values of less than 0.8 for either Risk or VoI treatment.

²⁹As with many optimization algorithms, this algorithm only guarantees the convergence upon local optima. Accordingly, we first run the OLS on the interior optimal condition to get reasonable starting parameters for the NLLS estimation. For subjects who choose few boundary allocations, the initial OLS estimation will be close to the region where the global optimum lies.

Second, we exclude subjects who fail to make choices varying enough to ensure convergence. For instance, 6 participants chose at least 15 boundary allocations in which $x = 0$ in both VoI and Risk problems; 2 participants chose too many boundary allocations in which $x = y$; and 14 participants either frequently distributed fixed proportions or amounts to one asset/wealth component, regardless of the budget set.

We report a summary of estimation results in Table 7.³⁰ By subject, we estimate both an expected utility, by restricting $\alpha = 1$, and a subjective expected utility specification. For the expected utility estimations, we report summary statistics of the individual estimates of ρ_{Risk} and ρ_{diff} , and the calculated values of ρ_{VoI} . For the subjective utility estimations we additionally report results on the individual estimates of α .^{31,32}

We note several observations from Table 7. First, for both EU and SEU, the mean of estimated utility curvatures is greater for Risk than VoI. The mean estimates of ρ_{Risk} are 1.16 and 2.77 for EU and SEU respectively, and the corresponding mean estimates of ρ_{VoI} are 1.61 and 3.81. Second, the variance of estimated curvatures is higher for the VoI than Risk choice problems - for both EU and SEU specifications. This is driven by the right tail of the distribution of ρ_{VoI} estimates. The median, the twenty-fifth and the fifth quantiles are virtually the same for ρ_{Risk} and ρ_{VoI} and for higher quantiles we find ρ_{VoI} exceeds ρ_{Risk} . Third, for the SEU specification, 87.5% (56 out of 64 participants) of the estimated α 's are smaller than one, reflecting optimism for the higher payoff. Since the EU specification is nested within SEU, we use a log-likelihood ratio test to evaluate the hypothesis a participant is an EU-maximizer against the alternative of a SEU-maximizer. For 42 out of 64 subjects, we reject EU model at the 5% level of significance.

For both utility specifications, we evaluate whether a participant has the same preferences in the Risk and VoI choice problems by testing whether their estimated ρ_{diff} is statistically different from zero. We present a visualization of the hypothesis tests results in the pair of stacked confidence interval plots presented in Figure 8. Despite the lack of precision in many estimates, as demonstrated by the numerous wide 95% confidence intervals, we find more evidence supporting the same heterogeneity of social preferences found in earlier reduced form and non-parametric analyses. For the EU specification, 26 out of 64 subjects' ρ_{diff} are significantly different from zero: sixteen larger and ten smaller

³⁰The formulation of the dummy variable for a Risk/VoI problem allows us to jointly estimate the DM's utility parameters for Risk- and VoI-selves. The joint estimation procedure implicitly assumes two error terms associated with Risk and VoI problems respectively follow the same distribution, which is reasonable under our experimental design of matched pairs of forty decision rounds. In unreported results, We have estimated utility parameters for Risk and VoI problems separately, and the resulting percentiles are identical to those reported in Table 7.

³¹In unreported results, we also estimated another version of the subjective expected utility model in which a DM's α can vary between the Risk and VoI problems. For 62 out of 64 participants, we cannot reject the null hypothesis of the same prior parameter at a 5% significance level.

³²We attempted to conduct the same procedure for the constant absolute risk aversion (CARA) utility function. We do not report these results as the numerical optimization algorithm did not find parameter estimates for a large majority of subjects. The estimation results based on a sub-sample of convergence subjects are consistent with those with CARA utility functions and comparable with other experiments.

	Expected Utility ($\alpha = 1$)			Subjective Expected Utility ($\alpha > 0$)			
	ρ_{Risk}	ρ_{VoI}	ρ_{diff}	α	ρ_{Risk}	ρ_{VoI}	ρ_{diff}
Mean	1.16	1.61	0.45	0.53	2.77	3.81	1.04
Variance	0.34	4.46	3.55	0.25	8.69	40.25	27.42
p5	0.44	0.44	-0.51	0.01	0.34	0.47	-0.93
p25	0.73	0.72	-0.10	0.09	1.09	1.15	-0.21
p50	1.04	1.16	0.07	0.46	1.80	1.82	0.11
p75	1.49	1.67	0.46	0.76	3.17	4.20	0.69
p95	2.16	2.92	1.48	1.51	7.12	12.12	3.11

Table 7: Summary of NLLS estimation of individual level parameter estimates of ρ_i and α . The quantile a is denoted as pa .

than zero. For the SEU specification, 20 out of 64 subjects' ρ_{diff} are significantly different from zero: eleven larger and nine smaller than zero.

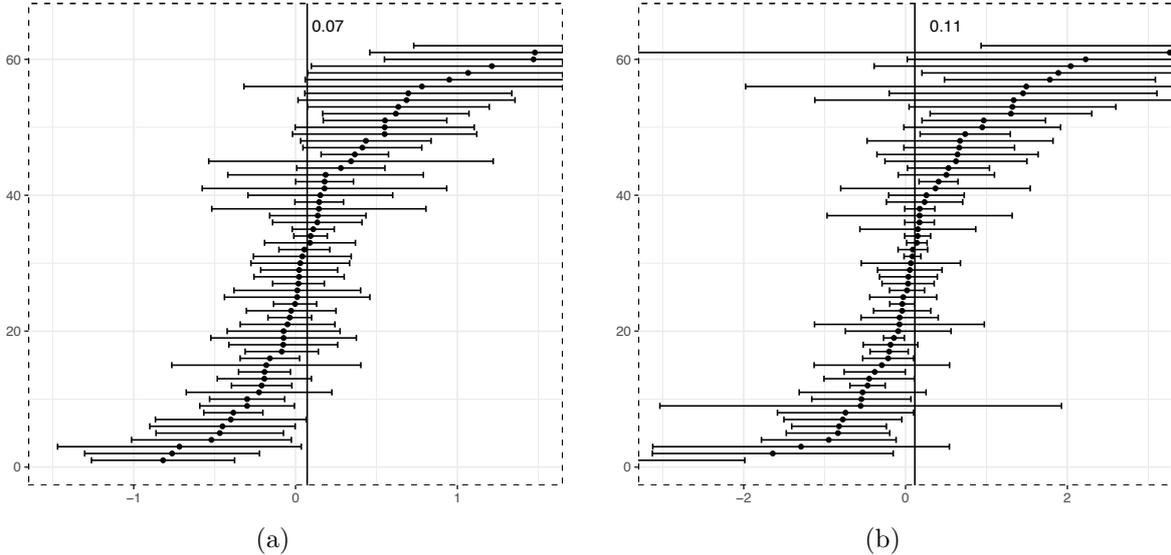


Figure 8: The estimated curvature differences ρ_{diff} based on (a) expected utility and (b) subjective expected utility model for each participant. We stack them from the lowest to the highest. Also notice that estimates from two subjects from each model are excluded because their values and confidence intervals are out of figures' ranges. The median ρ_{diff} for all 64 subjects is 0.07 in EU and 0.11 in SEU model, and both are stacked at the top.

6.3 Comparison with other experiments

Next we report and compare structural estimates from the data of two previously considered induced budget set experiments preferences: CFGK and CKMS.³³ This meta-analysis is designed to identify how structural estimates can vary across subject pools and choice

³³We do not conduct structural estimates for the formally considered linear budget dictator experiments, AM and FKM, because participants knew the realized states and therefore the (subjective) expected utility framework is not applicable.

interfaces. To do this we censor or filter data in different ways to make the variety of choice sets comparable across studies.

We start with “Full” data sets which are not censored, but for CFGK and CKMS we relabel the state-contingent assets so the cheaper one is the x good. Next, for CFGK and CKMS, we replace choices that violate first order stochastic dominance (FOSD) with the fully insured portfolio on the budget line. We also consider “Price” adjusted data sets, including ours, by selecting decision rounds for which relative prices range between one and three. Then we consider simultaneously FOSD censored and price filtered “FOSD & Price” data. We estimate the EU and SEU specifications for each individual and report the quartiles of estimate in Table 8.³⁴

First, we find the curvature estimates from our experiment are much larger than those from CFGK but more comparable with those from CKMS. A closer examination of the estimate of the subjective prior parameter reveals that our participants are the most optimistic, CKMS’s follow, and those in CFGK approximately assign equal probabilities for both states. As CFGK and CKMS share the same experimental protocols, the distinctions of both curvature and weight estimates may largely be attributed to the difference in subject pool: the former was conducted among US college student and the latter among a representative sample of the Dutch population.

How sensitive are the estimations to choice censoring, smaller price variation, and both? The corrections for first order stochastic dominance result in minimal changes in the curvature estimates of CFGK and CKMS but do increase the estimates of the subjective prior ratio of CKMS. Filtering for smaller relative prices has a differential impact across studies. In our study, estimates of ρ increase while those of α increase, both slightly. In CFGK ρ estimates increase but are almost unaffected in CKMS. In sum, both types of manipulation on choices affect the parametric estimation results to some degree, but they exhibit no unambiguous direction across three studies and furthermore, they do not abruptly alter the baseline magnitude of the estimates of each study.

Overall we find larger estimates for the curvature of power utility functions and smaller estimates of the subjective prior ratio between Low/High states than those from CFGK, but these estimates are more comparable with CKMS. These distinctions are more likely to be driven by different subjects pools rather than different experiment protocols.

³⁴Note the constrained utility maximization problems for FOSD censored data sets are the same as in problem 1. For the full and the price only adjusted data sets, we cannot impose the second constraint as in problem 1 that the money allocated to y is higher than to x account. Hence the optimal conditions for demands of the two commodities are slightly different. Details can be found in Appendix C.

Study	Experiment	Sample	Expected Utility ($\alpha = 1$)			Subjective Expected Utility ($\alpha > 0$)					
			ρ			α			ρ		
			p25	p50	p75	p25	p50	p75	p25	p50	p75
Current	Risk	Full	0.73	1.04	1.49	0.09	0.46	0.76	1.09	1.80	3.17
	VoI		0.72	1.16	1.67				1.15	1.82	4.20
	Risk	Price	0.71	0.99	1.24	0.23	0.53	0.71	1.10	1.75	2.76
	VoI		0.68	1.00	1.44				1.07	1.80	2.88
Choice under risk	CFGK	Full	0.31	0.58	0.79	0.77	1.01	1.24	0.25	0.48	0.93
		FOSD	0.31	0.59	0.84	0.91	1.06	1.17	0.30	0.52	0.96
		Price	0.36	0.66	0.94	0.80	0.95	1.05	0.36	0.72	1.18
		FOSD & Price	0.36	0.70	0.99	0.91	1.01	1.11	0.36	0.75	1.01
	CKMS	Full	0.65	1.03	1.51	0.45	0.78	1.01	0.78	1.27	2.17
		FOSD	0.69	1.13	1.72	0.73	1.00	1.22	0.75	1.12	1.86
		Price	0.54	0.91	1.56	0.50	0.82	1.00	0.69	1.28	2.64
		FOSD & Price	0.62	1.11	2.05	0.85	1.02	1.20	0.57	1.06	2.08

Table 8: Meta-comparison of structural estimates of EU and SEU specifications: Risk, VoI, CFGK (Choi et al., 2007a) and CKMS (Choi et al., 2014).

7 Conclusion

In this paper, we construct an instrument to identify and decompose social preferences behind a VoI. In particular, we examine pairs of contingent claim portfolio and VoI income distribution problems with a common budget constraint. A fixed commodity bundle induces the same marginal distribution of a DM’s wealth in both choice problems. This allows us to use a DM’s chosen contingent claim portfolio as a benchmark to establish whether her wealth distribution choice reflects equity preferring, efficiency preferring, or egoist social preferences. Our instrument consists of a set of forty linear budget sets for paired choice problems.

We applied this instrument in an experimental lab setting. Aggregate analysis suggests that there is no difference between behavior in the Risk and the VoI choice problems. However, our within-subject design allows us to demonstrate this is a fallacy. The analysis of intrapersonal differences points to clusters of social preference types. We establish the robustness of this conclusion using three empirical methodologies: reduced form statistical analysis, revealed preference analysis, and structural estimations of individual subjective utility functions. We found consistent classifications of individuals into groups exhibiting equity preferring, efficiency preferring, and egoist preferences.

Our results are informative to the heterogeneity of social preferences, with respect equity and efficiency, to the population of our sub-sample. However, we caution extrapolating these conclusions to other populations. We believe the intuitive interface and the structure of the data provided by our instrument lends itself well to more applied contexts. We intend to use it to measure the social and risk preferences of policymakers and beneficiaries

in the developing world where the impacts of efficiency-equity trade-offs are acute. We also intend to use the instrument to measure risk and social preferences in organizational contexts where such information is key to develop effective contract designs. Finally, we are developing extensions to the instrument incorporating income mobility.

References

- Acemoglu, D., Johnson, S., and Robinson, J. A. (2002), “Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution,” *Quarterly Journal of Economics*, 117, 1231–1294.
- Afriat, S. N. (1967), “The Construction of Utility Functions from Expenditure Data,” *International Economic Review*, 8, 67–77.
- Alesina, A. and Perotti, R. (1996), “Income Distribution, Political Instability, and Investment,” *European Economic Review*, 40, 1203–1228.
- Alesina, A. and Rodrik, D. (1994), “Distributive Politics and Economic Growth,” *Quarterly Journal of Economics*, 109, 465–490.
- Amiel, Y. and Cowell, F. (2000), “Attitudes to Risk and Inequality: A New Twist on the Transfer Principle,” Tech. rep., Distributional Analysis Discussion Paper.
- Amiel, Y., Cowell, F. A., and Gaertner, W. (2009), “To Be or not To Be Involved: A Questionnaire-Experimental View on Harsanyi’s Utilitarian Ethics,” *Social Choice and Welfare*, 32, 299–316.
- Andersson, O., Holm, H. J., Tyran, J.-R., and Wengström, E. (2014), “Deciding for Others Reduces Loss Aversion,” *Management Science*, 62, 29–36.
- Andreoni, J. and Miller, J. (2002), “Giving According to GARP: An Experimental Test of the Consistency of Preferences for Altruism,” *Econometrica*, 70, 737–753.
- Azrieli, Y., Chambers, C. P., and Healy, P. J. (2018), “Incentives in Experiments: A Theoretical Analysis,” *Journal of Political Economy*, 126, 1472–1503.
- Becker, N., Häger, K., and Heufer, J. (2013), “Revealed Notions of Distributive Justice II: Experimental Analysis,” *Ruhr Economic Papers*, #444, TU Dortmund University, Discussion Paper.
- Benabou, R. (2000), “Unequal Societies: Income Distribution and the Social Contract,” *American Economic Review*, 90, 96–129.
- Bernasconi, M. (2002), “How Should Income Be Divided? Questionnaire Evidence from the Theory of “Impartial Preferences”,” *Journal of Economics*, 9, 163–195.
- Brock, J. M., Lange, A., and Ozbay, E. Y. (2013), “Dictating the Risk: Experimental Evidence on Giving in Risky Environments,” *American Economic Review*, 103, 415–37.
- Bronars, S. G. (1987), “The Power of Nonparametric Tests of Preference Maximization,” *Econometrica*, 55, 693–698.

- van Bruggen, P. and Heufer, J. (2017), “Afriat in the Lab,” *Journal of Economic Theory*, 169, 546–550.
- Cettolin, E., Riedl, A., and Tran, G. (2017), “Giving in the Face of Risk,” *Journal of Risk and Uncertainty*, 55, 95–118.
- Charness, G., Gneezy, U., and Halladay, B. (2016), “Experimental Methods: Pay One or Pay All,” *Journal of Economic Behavior & Organization*, 131, 141–150.
- Charness, G. and Rabin, M. (2002), “Understanding Social Preferences With Simple Tests,” *Quarterly Journal of Economics*, 117, 817–869.
- Choi, S., Fisman, R., Gale, D., and Kariv, S. (2007a), “Consistency and Heterogeneity of Individual Behavior under Uncertainty,” *American Economic Review*, 97, 1921–1938.
- (2007b), “Revealing Preferences Graphically: An Old Method Gets a New Tool Kit,” *American Economic Review*, 97, 153–158.
- Choi, S., Kariv, S., Müller, W., and Silverman, D. (2014), “Who Is (More) Rational?” *American Economic Review*, 104, 1518–1550.
- Choudhry, N., Slaughter, P., Sykora, K., and Naylor, C. D. (1997), “Distributional Dilemmas in Health Policy: Large Benefits for a Few or Smaller Benefits for Many?” *Journal of Health Services Research & Policy*, 2, 212–216.
- Cox, J., Sadiraj, V., and Schmidt, U. (2015), “Paradoxes and Mechanism for Choice under Risk,” *Experimental Economics*, 18, 215–250.
- Cox, J. C. (1997), “On Testing the Utility Hypothesis,” *The Economic Journal*, 107, 1054–1078.
- (2004), “How to Identify Trust and Reciprocity,” *Games and Economic Behavior*, 46, 260–281.
- Engelmann, D. and Strobel, M. (2004), “Inequality Aversion, Efficiency, and Maximin Preferences in Simple Distribution Experiments,” *American Economic Review*, 94, 857–869.
- Fisman, R., Jakiela, P., and Kariv, S. (2017), “Distributional Preferences and Political Behavior,” *Journal of Public Economics*, 155, 1–10.
- Fisman, R., Jakiela, P., Kariv, S., and Markovits, D. (2015), “The Distributional Preferences of An Elite,” *Science*, 349.
- Fisman, R., Kariv, S., and Markovits, D. (2007), “Individual Preferences for Giving,” *American Economic Review*, 97, 1858–1876.
- Freundt, J. and Lange, A. (2017), “On the Determinants of Giving under Risk,” *Journal of Economic Behavior & Organization*, 142, 24–31.
- Frignani, N. and Ponti, G. (2012), “Risk versus Social Preferences under the Veil of Ignorance,” *Economics Letters*, 116, 143–146.

- Füllbrunn, S. and Luhan, W. (2015), “Am I my Peer’s Keeper? Social Responsibility in Financial Decision Making,” *Ruhr Economic Papers* #551.
- Füllbrunn, S. C. and Luhan, W. J. (2017), “Decision Making for Others: The case of Loss Aversion,” *Economics Letters*, 161, 154–156.
- Greiner, B. (2004), “The Online Recruitment System ORSEE 2.0 – A Guide for the Organization of Experiments in Economics,” Tech. rep., University of Cologne, technical Report 10, Working Paper Series in Economics.
- Gul, F. (1991), “A Theory of Disappointment Aversion,” *Econometrica*, 59, 667–686.
- Hadar, J. and Russell, W. R. (1969), “Rules for Ordering Uncertain Prospects,” *American Economic Review*, 59, 25–33.
- Harbaugh, W. T., Krause, K., and Berry, T. R. (2001), “GARP for Kids: On the Development of Rational Choice Behavior,” *American Economic Review*, 91, 1539–1545.
- Harsanyi, J. C. (1953), “Cardinal Utility in Welfare Economics and in the Theory of Risk-Taking,” *Journal of Political Economy*, 61, 434–435.
- (1976), “Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility,” in *Essays on Ethics, Social Behavior, and Scientific Explanation*, Springer, pp. 6–23.
- Heufer, J. (2013), “Testing Revealed Preferences for Homotheticity with Two-Good Experiments,” *Experimental Economics*, 16, 114–124.
- (2014), “Nonparametric Comparative Revealed Risk Aversion,” *Journal of Economic Theory*, 153, 569–616.
- Heufer, J. and Hjertstrand, P. (2017), “Homothetic Efficiency: Theory and Applications,” *Journal of Business & Economic Statistics*.
- Hong, H., Ding, J., and Yao, Y. (2015), “Individual Social Welfare Preferences: An Experimental Study,” *Journal of Behavioral and Experimental Economics*, 57, 89–97.
- Joseph, W. (2008), “Do Authoritarian Institutions Constrain? How Legislatures Affect Economic Growth and Investment,” *American Journal of Political Science*, 52, 322–343.
- Kariv, S. and Zame, W. (2014), “Character and Candidates: A View from Decision Theory,” Working Paper.
- Krawczyk, M. and Le Lec, F. (2010), “‘Give Me a Chance!’ An experiment in Social Decision under Risk,” *Experimental Economics*, 13, 500–511.
- Kuhn, A. (2013), “Inequality Perceptions, Distributional Norms, and Redistributive Preferences in East and West Germany,” *German Economic Review*, 14, 483–499.
- Kuziemko, I., Norton, M. I., Saez, E., and Stantcheva, S. (2015), “How Elastic Are Preferences for Redistribution? Evidence from Randomized Survey Experiments,” *American Economic Review*, 105, 1478–1508.

- Li, J., Dow, W. H., and Kariv, S. (2017), “Social Preferences of Future Physicians,” *Proceedings of the National Academy of Sciences*, 114, E10291–E10300.
- Michelbach, P. A., Scott, J. T., Matland, R. E., and Bornstein, B. H. (2003), “Doing Rawls Justice: An Experimental Study of Income Distribution Norms,” *American Journal of Political Science*, 47, 523–539.
- Moré, J. J. (1978), “The Levenberg-Marquardt Algorithm: Implementation and Theory,” in *Numerical Analysis*, Springer, pp. 105–116.
- Nelson, K. M., Schlüter, A., and Vance, C. (2018), “Distributional Preferences and Donation Behavior Among Marine Resource Users in Wakatobi, Indonesia,” *Ocean & Coastal Management*, 162, 34–45.
- Pahlke, J., Strasser, S., and Vieider, F. M. (2015), “Responsibility Effects in Decision Making under Risk,” *Journal of Risk and Uncertainty*, 51, 125–146.
- Rawls, J. (1958), “Justice as Fairness,” *The Philosophical Review*, 67, 164–194.
- Rodrik, D. and Wacziarg, R. (2005), “Do Democratic Transitions Produce Bad Economic Outcomes?” *American Economic Review*, 95, 50–55.
- Saez, E. and Stantcheva, S. (2016), “Generalized Social Marginal Welfare Weights for Optimal Tax Theory,” *American Economic Review*, 106, 24–45.
- Schildberg-Hörisch, H. (2010), “Is the Veil of Ignorance only a Concept about Risk? An Experiment,” *Journal of Public Economics*, 94, 1062–1066.
- Sippel, R. (1997), “An Experiment on the Pure Theory of Consumer’s Behaviour,” *The Economic Journal*, 107, 1431–1444.
- Sujoy, C., W., H. G., E., H. E., and Elisabet, R. E. (2011), “Are You Risk Averse over Other People’s Money?” *Southern Economic Journal*, 77, 901–913.
- Traub, S., Seidl, C., and Schmidt, U. (2009), “An Experimental Study on Individual Choice, Social Welfare, and Social Preferences,” *European Economic Review*, 53, 385–400.
- Traub, S., Seidl, C., Schmidt, U., and Levati, M. V. (2005), “Friedman, Harsanyi, Rawls, Boulding—or Somebody Else? An Experimental Investigation of Distributive Justice,” *Social Choice and Welfare*, 24, 283–309.
- Varian, H. R. (1982), “The Nonparametric Approach to Demand Analysis,” *Econometrica*, 50, 945–972.
- (1993), “Goodness-of-Fit for Revealed Preference Tests,” Working Paper.
- Vieider, F. M., Villegas-Palacio, C., Martinsson, P., and Mejía, M. (2016), “Risk Taking for Oneself and Others: A Structural Model Approach,” *Economic Inquiry*, 54, 879–894.

A Experimental Instructions

Preliminary Remarks

You are participating in an experiment investigating individual decision-making. Contingent on your decisions in this experiment, you can earn money in excess of your show-up fee of 10 RMB. Please pay careful attention to the instruction as a considerable amount of money is at stake.

During the experiment, please turn off your cell phones, laptop computers and other communication equipment. Do not turn on any software on the desktop other than the experimental application. Please do not communicate with any of the other participants or look at their computer monitors. If you have questions anytime, please raise your hand and we will address it as soon as possible. If you do not obey these rules, we will ask you to leave without any payment.

Overview

In today's experiment, you are going to complete 80 decision periods. The 80 periods are broken into two stages of 40 periods each Stage 1 and Stage 2. There is a break of 10 minutes between two stages. At the end of the 80th period, only one decision period will be randomly selected as payment period. Therefore, your decision in each period might determine your final payoff. In today's experiment, there is a 50% chance you will receive a High Reward and a 50% chance you will receive a Low Reward. Whether you receive the High or Low reward will be determined only at the end of the experiment. Your decisions can't influence these probabilities; however, you can influence the amount of money associated with these rewards.

In a Stage 1 period, you will simply be asked to divide a pie of money between the High and Low Rewards. This division only affects the potential amount of money you will receive. Every other participant will make their own respective divisions, and whether you receive a High or Low reward has no bearing on what reward they receive. At the end of experiment, if the randomly selected period is in Stage 1, each subject will toss their own individual coin to determine whether they receive the High or Low Reward. Note that this determination will only be made if a Stage 1 period is selected.

All monetary amounts in the experiment are Yuan. At the end of the experiment, you will be paid privately. When we call your ID number, please come forward to the sign-in counter and receive your earnings in the experiment plus your show-up fee. We ensure you that your participation in the experiment, your name, ID number and any information about your payoffs are confidential.

How to make decisions?

In each period in Stage 1, you will be asked to choose your most preferred division of a pie of money between the High (indicated by red) and Low (indicated by blue) Rewards by moving a slider. You may receive the money from either High Reward or Low Reward as your final payment.

Your division is restricted in that High Reward must be assigned at least 50% of the pie and Low Reward no more than 50%. It is important to note that, the pie size (total reward) is not fixed and grows as you allocate a higher percentage of the pie to High Reward. In different periods, the pie size and growth rate vary. You should also note that

even in the same period, the rate that pie size grows is not the same for different divisions. For example, the following hypothetical table shows how pie size change differently for the same 5% increase of share for two different divisions. For a (44%, 56%) division, a 5% increase of the High Reward share increases the pie size by \$8.97, and for the (21%, 79%) division the same 5% increase of the High Reward share increase the total pie size by \$16. To understand the pie growth rate you need to check the pie size at various divisions each period.

Low%	High%	Low\$	High\$	Total\$
44%	56%	70.21	89.36	159.57
39%	61%	65.73	102.81	168.54
21%	79%	44.37	166.90	211.27
16%	84%	36.36	190.91	227.27

Table 9: Example of Varying Growth Rate for Same Proportional Increase

Figure 9 gives an example of what your computer screen may look like in one period. The top of the screen is a status bar showing current stage, current period and time elapsed in this screen. The left part is a pie indicating shares allocated to High (red) and Low (blue) Rewards. When you make your division choice, shares for High and Low in the pie will change accordingly, as will the pie size. The top of the right part is a 10 row table shows the current and the nearest divisions of the pie (Low% and High%), the amount of Low Reward and High Reward (Low\$ and High\$) and current pie size (Total\$). The right bottom is the slider where you make decisions. The calibration of slider in this screen is 50% – 100%, which means points on the slider are proportions for High Reward. You can make your decision by either dragging the green triangle along the slider to your most desired position or clicking ”+” button at the two ends of the slider. Once you are satisfied with your division, click on the ”Confirm and Leave” button. After that, you will leave the current period and won’t be able to change your decision in this period. So please confirm your decision before clicking it.

How are earnings determined?

At the end of the experiment, the computer will randomly select one out of 80 periods for all subjects as Payment Period. Every period is equally likely to be chosen and the period selected is same for all subjects.

If the Payment Period belongs to Stage 1, the computer will randomly assign you either the High or Low Reward. Then your individual earnings will be determined according to the division you chose in the Payment Period.

Figure 10 shows the screen you will see if the Payment Period belongs to Stage 1. The top of the screen shows the Payment Period and which stage it belongs to. After you click ”Random Reward” button, the computer select the Reward Amount, displaying your Reward level and payoff. After that, a ”Next” button will appear and you can click it to leave this screen. In the example shown in Figure 10, the Payment Period is Period 1 and the reward is High, according to his division in Period 1 (18%, 82%), his payoff is \$45.22.

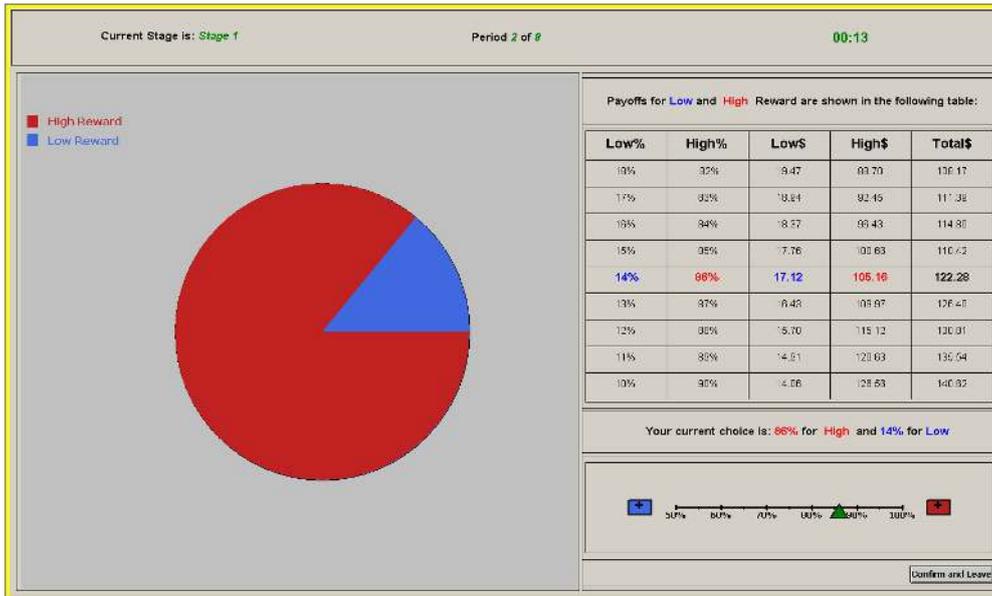


Figure 9: Decision Screen

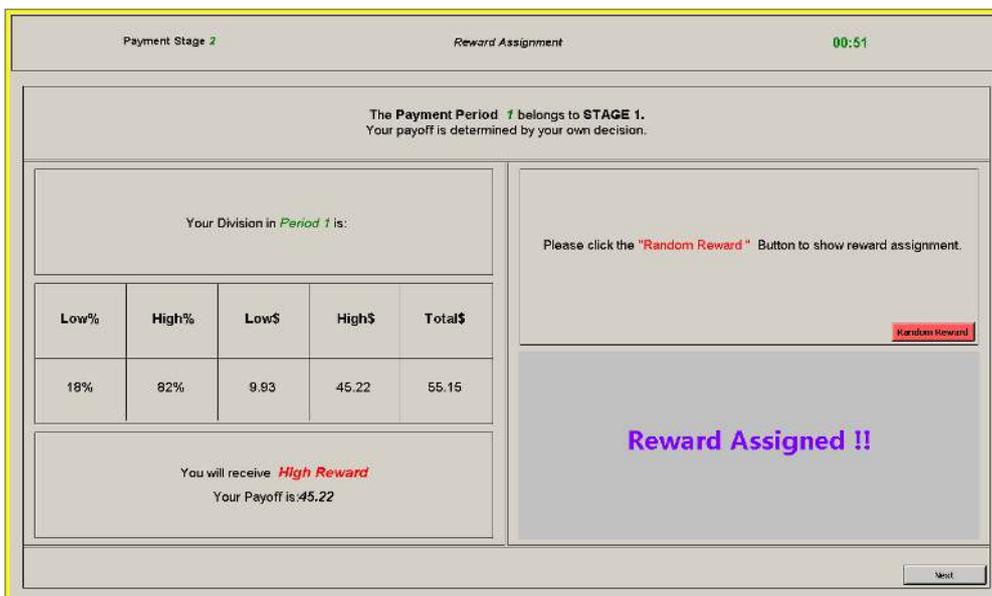


Figure 10: Payment Screen in Stage 1

Appendix A.2: Instructions (Stage 2)

Review

In today's experiment, you are going to complete 80 decision periods, which are broken into two stages Stage 1 and Stage 2. At the end of the 80th period, only one decision period will be randomly selected as payment period. Your payment is a reward and there is a 50% chance you will receive a High Reward and a 50% chance you will receive a Low Reward. Whether you receive the High or Low reward will be determined only at the end of the experiment. Your decisions can't influence these probabilities; however, you can influence the amount of money associated with these rewards.

Overview

In a Stage 2 period, your task is still to divide a pie of money between High and Low Rewards. This stage is different because you will be matched with another participant in the experiment (called your counterpart), but we will not reveal who. Since you and your counterpart both make divisions of High and Low Reward, we only select one of two divisions to determine payoffs for the pair. There is a 50% chance your division will be carried out (and your counterparts' not) and 50% of chance your counterpart's division will be carried out (and yours not). This stage also differs in that one of you will receive the High Reward and the other the Low Reward. So there are two possibilities, each equally likely: (1) you receive the High Reward and you counterpart the Low Reward, and (2) you receive the Low Reward and your counterpart the High Reward. Again, note that the division selection and reward determination will only be made at the end of the experiment if a Stage 2 period is selected.

How to make decisions?

In each period in Stage 2, you will be asked to choose your most preferred division of a pie of money between the High (indicated by red) and Low (indicated by blue) Rewards by moving a slider. The allocation you choose may determine the payoff of yourself and your counterpart. If one of the pair receives High Reward, the other one will get Low Reward surely.

The interface you face is similar to that of Stage 1. The pie size (total reward) is not fixed and grows as you allocate a higher percentage of the pie to High Reward. In different periods, the pie size and growth rate also vary. Even in the same period, the rate that pie size grows is not the same for different divisions. You need to check the pie size at various divisions to understand the pie growth rate you each period.

How are earnings determined?

At the end of the experiment, the computer will randomly select one out of 80 periods for all subjects as Payment Period. Every period is equally likely to be chosen and the period selected is same for all subjects.

If the Payment Period belongs to Stage 2, the payment process will contain three parts. You will firstly be randomly paired with another participant. Then the computer will randomly select one of your two division choices to use to determine the money associated with the High and Low Reward amounts. Next, the computer will randomly assign the High and the Low Reward amounts to you and your counterpart.

Figure 11 shows the screen you will see if the Payment Period belongs to Stage 2. The top of the screen shows the Payment Period and which stage it belongs to. After you click "Random Match" button, the computer will end the random match process and show both of your chosen divisions. After you click the "Division Selection" button, the computer will end the random process to select one of the divisions, and a red check will appear under the selected table. In the example shown in Figure 3, your division will be carried out. It means one of you will receive \$28.57 and the other one \$42.86. After you click "Next" button on the right bottom of the screen, you will see a similar screen for random reward. When you click "Random Reward" button, you can see who was assigned the High Reward and who the Low Reward, as well as the corresponding payoffs.

Payment Stage 2 Match and Execution 00:32

The Payment Period 3 belongs to **STAGE 2**.
Your payoff is determined by either your decision or your partner's decision.

Your Division in <i>Period 3</i> is:					Please click the "Random Match" Button to show your and your partner's decisions.					Your Partner's Division in <i>Period 3</i> is:				
Low%	High%	Low\$	High\$	Total\$	<p style="color: purple; font-weight: bold; font-size: 1.2em;">Match Finished !!</p> <p style="color: purple; font-weight: bold; font-size: 1.2em;">Division Selected !!</p>					Low%	High%	Low\$	High\$	Total\$
40%	60%	28.57	42.86	71.43						14%	86%	19.81	121.70	141.51
Your Division is executed.					Please click the "Division Selection" Button to show who's decision will be executed.									

Figure 11: Payment Screen in Stage 2

B Type classifications based on alternative criteria

Table 10 lists type-group classifications for eight criteria differing in test types (paired t -test or paired Wilcoxon signed-rank test), confidence levels (95% or 90%) and hypothesis symmetries (two tail or one tail). With only 40 observations, one may have concerns that the empirical distributions of the reports t -tests have yet to approach their asymptotic distributions. Rather than testing for normality of the distribution of paired differences, which would establish the validity of the t -tests, we report classifications based upon the nonparametric Wilcoxon signed-rank test. We find this generates qualitatively similar clusters. But we note that the Wilcoxon signed-rank tests offers lower power and results in slightly larger egoist clusters.

Criteria	t -test				W -test			
	$CL = 95\%$		$CL = 90\%$		$CL = 95\%$		$CL = 90\%$	
	2 Tail	1 Tail						
Equity	24	28	28	31	26	31	31	33
Egoist	51	43	43	35	48	36	36	32
Efficiency	17	21	21	26	18	25	25	27

Note: This table provides the number of participants in each type based on eight classification criteria.

Table 10: Type classifications by alternative criteria

Notice that the results we reported in the Section 4 are robust to alternative type classification criteria. In unreported results, we ran the reduced form analyses for each of the three groups and for each of these type classification results from the Table 10. we find the estimated demand functions are similar to those reported in Table 2. Furthermore, we cannot reject the null hypothesis that the demand functions for Risk problems are the same across these three groups.

C Demand functions for uncensored CFGK and CKMS datasets

Notice the DM in CFGK and CKMS can choose allocations on the whole budget line, therefore, the optimization conditions are slightly different from those listed in subsection 6.2. With CRRA utility, we first similarly replace all boundary allocations of 0 with a small number such that the ratio between x and y is either ω or $1/\omega$. However, when there is no censoring constraint, the DM can choose boundary allocation $(x, y) = (0, z)$ or $(x, y) = (z/q, 0)$. For the first case, we replace 0 with ωz and for the second case, we replace 0 with $\omega z/q$. Since ω is much smaller than q , this procedure will not affect the estimation. Solving the utility maximization problem yields the following optimization conditions:

$$\ln(x^*/y^*) = f[\ln(q), \omega; \alpha, \rho]$$

$$= \begin{cases} \ln(\omega) & \text{if } \ln(\alpha) - \rho \ln(\omega) \leq \ln(q), \\ -\frac{1}{\rho} [\ln(q) - \ln(\alpha)] & \text{if } \ln(\alpha) + \rho \ln(\omega) < \ln(q) < \ln(\alpha) - \rho \ln(\omega), \\ -\ln(\omega) & \text{if } \ln(q) \leq \ln(\alpha) + \rho \ln(\omega), \end{cases}$$

where the first and third condition are derived from two corner solutions, one on the y -axis and the other on the x -axis; the second condition is similarly for the interior solution.