Eliciting Preferences for Risk and Altruism: Experimental Evidence*

Romain Gauriot
Department of Economics, Deakin University

Stephanie A. Heger
Department of Economics, University of Bologna

Robert Slonim
School of Economics, University Technology of Sydney

September 29, 2022

ABSTRACT: We apply the basic lessons and insights learned in the elicitation and estimation of risk and time preferences literature to the literature on social preferences. Following Andersen et al. (2008), we design a laboratory experiment to jointly elicit risk preferences and preferences for altruism. Consistent with theory, we find that the standard simplifying assumptions about risk preferences lead to significantly biased estimates of altruism. This is particularly problematic when comparing altruism across relevant sub-groups, such as gender and wealth, leading to possibly erroneous conclusions about which is the more generous sex and the self-regarding rich.

KEYWORDS: Altruism, risk aversion, experiment.

*Our experimental design benefited from insightful comments from seminar participants at the University of Bologna, as well as Lisa Spantig, Marie Claire Villeval, and Nate Wilcox. Alex Berger and Patrick Hendy provided excellent research assistance. We are grateful to Glenn W. Harrison for providing his code and data.

1 Introduction

The standard model of other-regarding preferences allows a decision-maker to receive utility from his own payoff and utility from giving to a recipient (e.g., another individual or a charitable organization), where a preference parameter governs the relative intensity between these two utility components. Previous literature has taken an interest in two aspects of these types of models: (1) the magnitude of the social preference parameter and (2) potential differences in the shape of utility functions over payoffs to self versus payoffs to recipients that may result in different responses to incentives for self versus the recipient. And while the importance of joint estimation of preference parameters has been shown to matter in a wide range of applications (Harrison, 2018) and most notably to time preferences (Andersen et al., 2008),[1] this insight and methodology have not been widely applied to the social preferences literature.

Following the joint estimation approach of Andersen et al. (2008), we estimate three functional forms of utility and show that the estimate of a social preference parameter is significantly biased by standard assumptions on the curvature of the decision-maker's utility function. Further, across all three specifications, our data reject the assumption that the curvature over utility for self and the curvature over utility of the recipient's payoff (e.g., charity) are equal and that incorrectly assuming equality leads to significant underestimation of the intensity of social preferences (Gauriot, Heger, and Slonim, 2020). This marks a departure from the bulk of the literature measuring altruism which tends to assume a single curvature parameter.[2]

To the best of our knowledge, there are two exceptions in the social preferences literature—DellaVigna, List, and Malmendier (2012) and Exley (2016)—to which our paper is closely related. DellaVigna, List, and Malmendier (2012) estimate a structural model of giving and social pressure in which they assume a quasilinear utility function where the utility over self payoffs is linear and then estimate a separate concave curvature over the payoffs that go to the charity (i.e., donations).[3] By doing so, they recognize that the curvatures over self and charity payoffs may represent differ-

[1]Other examples include, the estimation of subjective probabilities (Andersen et al., 2014a), correlation aversion (Andersen et al., 2018), and bid functions in first price sealed bid auctions (Harrison and Rutström, 2008a).

[2]See Tables A1 and A2, respectively, in Gauriot, Heger, and Slonim (2020).

[3]This quasilinear framework is also used in Null (2011).

ent aspects of utility. For example, the concavity over the donation may represent the idea that individuals receive a boost in utility from the first dollar donated but that utility increases at a diminishing rate due to warm glow motives.[4] While they allow for curvature over self and curvature over charity payoffs to differ, they *assume* linear utility rather than fully estimating the model. We contribute beyond DellaVigna, List, and Malmendier (2012) by jointly estimating both curvature terms.

Similarly, Exley (2016) reports results from an experiment in which she asks whether self risk and charity risk differ. In order to account for the possibility that a dollar for oneself is not the same, in utility terms, as a dollar to charity, Exley (2016) uses a normalization task before eliciting risk preferences. This task calibrates the dollar amount $Y such that the individual is indifferent between receiving $10 for themselves or donating $Y to charity. Then, $Y is used in the risk preference elicitation. Exley (2016) jointly estimates both curvature parameters, but does so in a stepwise approach that first calibrates a value for $Y and then estimates risk preferences rather than jointly estimating all of the parameters of the model.

Our data come from a laboratory experiment in which subjects make a series of decisions between keeping money for themselves and/or donating money to Hope for Children, an Australian based charity working to help vulnerable children in Ethiopia.[5] The experiment is designed to elicit the curvature of the utility function over own payoffs, the curvature of the utility function over a donation to charity and the social preference parameter over a series of four decision tasks. In the first two tasks, we elicit aversion to self risk and aversion to charity risk by asking subjects to make binary choices between lotteries for themselves and for the charity, respectively. In the third task, we eliminate risk, but introduce a trade-off between self and charity by asking subjects to make a series of decisions in a standard dictator games in which they decide how much of their endowment to give to the charity. These first three decision tasks allow us to estimate altruism and diminishing marginal utility of money for self and the diminishing marginal utility of money for the charity. In the fourth task, we introduce risk to the dictator game, which allows us to estimate altruism and aversion to self risk and aversion to charity risk.

[4] On the other hand, as noted by DellaVigna, List, and Malmendier (2012) and Null (2011), pure altruism is best described by linear utility over the donation as an individual with pure motives receives a constant increase in utility from each dollar donated provided that there is a need for the additional dollar.

[5] See Hope for Children

We provide empirical evidence, consistent with Exley (2016), that there is significantly more concavity in the utility over charity's payoff than in the utility over self payoffs. Second, assuming (incorrectly) that the curvature over self and charity payoffs are equal leads to significant underestimation of altruism, consistent with the simulation results presented in Gauriot, Heger, and Slonim (2020). Further, while our main results specify a constant relative risk aversion (CRRA) utility function, we show that they are robust to additional utility specifications, including a CRRA utility function with probability weighting and a power utility function specification. Third, we show that comparisons of altruism across relevant subgroups, such as gender and wealth, are significantly biased by the single-parameter assumption.[6] Specifically, under the assumption of a single curvature parameter over self and charity payoffs, one would conclude that men are more altruistic than women and that the less wealthy individuals are more altruistic than wealthier individuals. However, when we allow the curvature over self payoffs and the curvature over charity payoffs to differ, we conclude that women are more altruistic than men and that there are no significant differences in altruism across wealth.

Our experimental design, which elicits curvature and altruism both in environments with and without risk, allows us to speak to the literature on the potential differences in preferences stemming from risk versus diminishing marginal utility, a topic on which there is mixed evidence (see Andersen et al. (2008), Andreoni and Sprenger (2012a), Andreoni and Sprenger (2012b), Cheung (2015), Harrison, Lau, and Rutström (2013) and Andreoni and Sprenger (2015)). When we (incorrectly) assume there is a single curvature parameter over self and charity payoffs, we find that introducing risk significantly *decreases* the curvature of the utility function relative to when there is no risk and that individuals are also less altruistic when there is risk versus when there is not (Exley, 2016). However, in the unrestricted model, which allows the curvature parameters to differ, we find that the introduction of risk has different effects on each of the curvature parameters: the utility function over self payoffs is significantly less concave with risk while the utility function over charity payoffs is significantly more concave with risk. Further, there is no longer a significant

[6] Gauriot, Heger, and Slonim (2020) also show that the commonly used restrictive assumptions on the shape of the utility function may lead to incorrect comparisons of altruism across relevant subgroups, including gender, wealth and giving motives. To test these hypothesis, we collected data on gender but also exogenously vary the level of background wealth of the subject ($5 show-up fee versus $20 show-up fee) and manipulate the motive for giving (pure versus warm glow).

difference in altruism in environments with and without risk.

2 Design and Procedures

In this section, we describe the design of the experiment, the model on which our experiment is based, our main hypotheses and the estimation procedure used to generate our main results.

2.1 Design

Our experimental design is akin to the design proposed by Andersen et al. (2008) to estimate risk and time preferences. Andersen et al. (2008)'s design includes two distinct tasks, one to identify the curvature of utility and one to identify the discount rate. The data generated by these two tasks are then estimated jointly. In a similar spirit, our experimental design consists of four tasks. The first task identifies risk preferences of self payoffs (r_s), the second task identifies risk preferences over charity payoffs (r_c), the third task identifies altruism (α) as well as diminishing marginal utility over self and charity payoffs (r_s and r_c, respectively) and the fourth task identifies altruism (α) as well as risk preferences over self and charity payoffs (r_s and r_c, respectively). We describe each of the four tasks below.

2.1.1 Experimental tasks

Subjects were asked to complete four tasks in which each decision follows a similar structure. In each decision, the subject chooses between two lotteries, lottery A and B. In each lottery, the subject can earn two possible payoffs: a high payoff with probability p and a low payoff with probability with probability $1-p$. The probability to earn the high payoff is the same in both lottery A and lottery B. Following Brown and Healy (2018), within each task we show each decision on a separate screen and in a random order. Whether option A and B are shown on the left or right of the screen is also determined randomly.

Task 1: Identify r_s only In the first task, subjects make choices over 36 lotteries listed in Table A5, which is the standard task proposed by Holt and Laury (2002) to elicit risk aversion. We use the same lotteries that are used in Andersen et al. (2008) but increased, proportionally, the size of the payoffs. Figure A1 in the Appendix shows how these decisions are displayed to the subjects. Our graphical representation contributes to the literature by showing both the height of the payoffs (y-axis) and the probabilities to get those payoffs (x-axis).

Task 2: Identify r_c only Task 2, used to identify r_c, is identical to task 1 (see Table A5) except that now the lotteries are over the charity's payoff rather than self's payoff. Figure A2 in the Appendix displays how those decisions are shown to the subjects.

Task 3: Identify r_s, r_c and α without risk The third task does not involve risk and asks subjects to divide money between themselves and the charity. In this task, subjects faced 46 allocation decisions, which are listed in Table A6. The payoffs in this Task come from points on budget lines similar to the budget lines used in Fisman, Kariv, and Markovits (2007).[7] Figure A3 displays a screenshot of the experimental interface as seen by the subject.

Task 4: Identify r_s, r_c and α with risk The fourth and final task is a risky dictator game (Andersen et al., 2018), which includes risk over self payoffs and risk over the charity's payoffs as well as tradeoff between self and charity's payoffs. In this task, subjects make 45 decisions between lotteries, which are displayed in Table A7. Figure A4 displays the experimental interface for this task.

The four tasks were presented in the same order to all subjects, but the decisions within each task were presented randomly. At the end of the sessions, we did manual randomization to determine which choice was pay-off relevant and to resolve any lotteries, if necessary.

[7] In our task, subjects choose from a binary option whereas in Fisman, Kariv, and Markovits (2007) they choose from a continuous budget line. This type of task has also been used in Fisman et al. (2015), Fisman, Jakiela, and Kariv (2015), and Fisman, Jakiela, and Kariv (2017). This literature often uses a CES utility framework to motivate their estimation. For our purposes—studying the effect of imposing the restriction of $r_s = r_c$ on estimates of altruism—the CES function is not appropriate as there is a single parameter that governs the curvature. However, there is a direct mapping between the CES utility function estimated in Fisman, Kariv, and Markovits (2007) and our CRRA functional form when we impose the curvature to be identical over self and charity payoffs (see Gauriot, Heger, and Slonim (2020).)

2.2 Utility Framework & Hypotheses

In order to structurally estimate parameters of a utility function, we first must make some assumptions on the parametric form of the utility function. We make a number of standard assumptions in choosing our parametric form. First, we assume that utility over self payoffs are independent from the utility to charity payoffs, which allows for the subutility functions to enter additively.[8] Second, we allow for preference over self risk (r_s) to differ from preferences over charity risk (r_c). This second assumption makes the CRRA utility function a good candidate and thus we follow a large literature in structural estimation that uses the CRRA functional form Andersen et al. (2008, 2018). Thus the latent model that we will use as the basis for our statistical estimation is given by

$$U(s,c) = \frac{s^{1-r_s}}{1-r_s} + \alpha\frac{c^{1-r_c}}{1-r_c}, \tag{1}$$

where r_s represents the curvature over self payoffs (s), r_c represents the curvature over the charity's payoffs (c) and $\alpha > 0$ represents altruism (and $\alpha < 0$ represents spite). Thus, $r = 0$ corresponds to a linear utility function (i.e., risk neutrality), while $r > 0$ corresponds to a concave utility function (i.e., risk aversion).

The first order conditions from maximizing equation 1 are given by

$$\alpha = \frac{c^{r_c}}{s^{r_s}} \tag{2}$$

In this paper, we are interested in how inferences about α are affected by assumptions on r_s and r_c. Consider two cases: in case 1, we assume that $r_s = r_c = r$ and in case 2, we assume that $r = r_s < r_c$ (i.e., individuals are more averse to charity risk than self risk) and let α_1 and α_2 denote the α generated by case 1 and case 2, respectively. Suppose we know that the data-generating process is given by case 2, but we incorrectly assume case 1. In this case, it is straightforward to show that we will under-estimate α. To see this, note that the comparison of case 1 and case 2 is given by

$$\alpha_1 = (\frac{c}{s})^r \lesseqgtr A_\alpha = \frac{c^{r_c}}{s^r} \tag{3}$$

[8]This has been the standard in the literature. See, for example, Null (2011), DellaVigna, List, and Malmendier (2012), and Lilley and Slonim (2014a)

Since $0 \leq r_c, r_s, r \leq 1$, implies $\alpha_1 < \alpha_2$ and thus, altruism is under-estimated.

Gauriot, Heger, and Slonim (2020) show by simulation that not accounting for differences in r_s and r_c will lead to biased estimates of α. In particular, their simulations show that if $r_s > r_c$, but $r_s = r_c$ is assumed, then altruism is over-estimated. On the other hand, if $r_s < r_c$, but $r_s = r_c$ is assumed, then altruism is under-estimated. Our main results, presented in Section 3 tests whether the restricted model ($r_s = r_c$) leads to a biased estimate of α compared to our unrestricted model ($r_s \neq r_c$).

Our first hypothesis builds from the findings in Exley (2016) who finds that when subjects are making trade-offs between keeping money for themselves or giving to charity that they are significantly more averse to charity risk than self risk. Thus, we hypothesize that $r_s < r_c$. Following from the predictions in Gauriot, Heger, and Slonim (2020) described above, we subsequently hypothesize that α is significantly underestimated in the restricted model relative to the unrestricted model.

Hypothesis 1. *The curvature over the utility for self pay-offs and the curvature over the utility to charity pay-offs are significantly different and individuals are more averse to charity risk than self risk; that is, under CRRA utility, $r_s < r_c$.*

Hypothesis 2. *Restricting $r_s = r_c$ will lead to an underestimation of altruism: $\alpha_{restricted} \leq \alpha_{unrestricted}$.*

2.2.1 The Importance of Heterogeneity

In addition to identifying these three preference parameters, we also examine how heterogeneity in curvature affects estimates of altruism across relevant sources of heterogeneity: gender, wealth, donor motives and risk. For each of these relevant subgroups, there are theoretical and conceptual reasons why the curvature might vary between subgroups and thus not accounting for these differences will lead to a biased estimate of altruism. We discuss each below.

Gender To examine heterogeneity from gender we collected demographic information from subjects at the end of the experiment. Building on Croson and Gneezy (2009), we hypothesize that women will be more averse to risk than men, leading to more curvature in the utility function of women relative to men. Not accounting for this increased curvature may lead to incorrect comparisons of altruism between genders.

Wealth To test whether wealth differentially affects our parameter estimates we include a treatment in which the show-up fee is \$20 instead of the typical \$5. Motivated by the decreasing absolute risk aversion of the CRRA utility function, we hypothesize that wealthier individuals are willing to take on more self risk than less wealthy individuals, but not necessarily more charity risk. Further, not accounting for the differences in risk preferences due to differences in wealth will lead to inflated estimates of altruism among the wealthier individuals.

Donor Motives To test whether donor motives, specifically, pure altruism versus warm glow, affect the curvature of the utility function, we frame the decisions as motivated by Warm Glow (WG) or by Pure Altruism (PA) (Lilley and Slonim, 2014b). In the WG framing we emphasize the sacrifice made to self in order to give to the charity, using words such as *"you gave up \$Y of your own money to donate"* and *"you donated \$X"*. In the PA treatment we emphasize the benefit provided to the recipient and use words such as *"Hope for Children receives \$X"*. These treatments are motivated by the discussion in DellaVigna, List, and Malmendier (2012) and Null (2011) where the curvature of a utility function that captures pure altruism may be less concave than a utility function that captures warm glow. We hypothesize that subjects in the Pure Altruism treatment will display less aversion to charity risk than subjects in the Warm Glow treatment and thus not accounting for motives will lead to an inflated estimate of altruism for subjects in the Pure Altruism treatment relative to the Warm Glow treatment.

Hypothesis 3. *Ignoring heterogeneity in curvature of utility functions within subgroups (i.e., male versus female, low versus high wealth and pure versus impure motives), will lead to biased inferences about differences in altruism across subgroups.*

Risk Finally, we investigate how curvature generated by risk versus diminishing marginal utility affects the degree of altruism. By comparing estimates obtained with the data from Task 3 (no risk) versus those obtained with the data from Task 4 (with risk) we can ask how the introduction of risk affects altruism. Exley (2016) finds that subjects are significantly less altruistic in the presence of risk, when $r_s = r_c$ is imposed on the data.

2.3 Data

The data for our experiment come from 186 University of Sydney's students recruited through ORSEE (Greiner, 2015) between May and August 2018 and the experiment was coded in Ztree (Fischbacher, 2007). Table 1 shows the subjects' assignment by treatment and gender. Subjects made 165 decisions during the approximately 90 minute sessions. The average earnings were \$25.7 AUD for self and \$10.40 AUD for Hope for Children.

TABLE 1: SAMPLE SIZES

	Male	Female	Total
Warm Glow \$5	31	31	62
Pure Altruism \$5	31	31	62
Warm Glow \$20	31	31	62
Total	93	93	186

Number of subjects per treatment, by gender.

2.4 Econometric Specification

Our econometric specification closely follows the specification used in Andersen et al. (2018).

Let m denote the show up fee and let EU_j $(j \in \{A, B\})$ denote the expected utility the individual receives from choosing option j. EU_j is given by:

$$EU_j = p * U(m + s_h, c_h) + (1 - p) * U(m + s_l, c_l), j \in \{A, B\} \tag{4}$$

where m is the background wealth of the individual and $U(s, c)$ is defined by equation (1).

To estimate the parameters of the utility function, we use the following random utility model where the latent index is:

$$\nabla EU = \frac{\frac{EU_B - EU_A}{\nu}}{\mu_{task}} \tag{5}$$

where μ_{task} is the standard fechner error parameter associated with each task (Hey and Orme, 1994) and ν is the "contextual utility" term proposed by Wilcox (2011).

When $\mu_{task} \rightarrow 0$ the decision becomes deterministic; that is, the option with the highest expected utility is chosen with certainty. And as $\mu_{task} \rightarrow \infty$ the decision is best described as randomly chosen.[9] ν, the contextual utility term, is the difference between the maximum and minimum utility over all possible prizes and ensures that the term $\frac{EU_B - EU_A}{\nu}$ is between 0 and 1.[10] The contextual utility term is used widely throughout the literature (Harrison and Rutström, 2008b; Blavatskyy, 2011; Cheung, 2020).

The likelihood of observing that A is chosen in decision i is:

$$L(A|\alpha, r_s, r_c, \mu_{task}) = \Phi(\nabla EU) \tag{6}$$

where $\Phi(\cdot)$ is the cumulative normal distribution. The log-likelihood over all observations is given by

$$L(i|\alpha, r_s, r_c, \mu_{task}) = \sum_{i=1}^{N} \left[ln(\Phi(\nabla EU)) * \mathbb{1}_{[j_i=1]} + ln(\Phi(1 - \nabla EU)) * \mathbb{1}_{[j_i=0]} \right] \tag{7}$$

where j_i equals to 1 if option A is chosen and 0 otherwise. We maximize the log-likelihood using Stata's modified Newton-Raphson (NR) algorithm. The standard errors are clustered by subject.

3 Results

Throughout this section, we will compare the results from the unrestricted model (i.e., where we allow r_s and r_c to differ) versus the restricted model (i.e., where we constrain $r_s = r_c = r$).

In Table 2 we pool together all the collected data and test Hypotheses 1 and 2. First, consistent with Hypothesis 1 and Exley (2016), results from the unrestricted

[9] Because the different experimental tasks do not have the same level of difficulty we allow μ_{task} to differ in each of the four tasks and we therefore estimate four noise parameters. For example, Tasks 1, 2, and 4 contain risk, whereas Task 3 does not; Tasks 1 and 2 do not contain trade-offs between self and charity, while Tasks 3 and 4 do.

[10] To be clear, ν takes the following form

$$\nu = max\left(U(s_h^A, o_h^A), U(s_l^A, o_l^A), U(s_h^B, o_h^B), U(s_l^B, o_l^B)\right) - min\left(U(s_h^A, o_h^A), U(s_l^A, o_l^A), U(s_h^B, o_h^B), U(s_l^B, o_l^B)\right)$$

. However, in Task 3, $\nu = 1$ because there is no risk and hence only two possible outcomes.

TABLE 2: Main Result: Parameter Estimates from the Unrestricted and Restricted Model

Parameter	Estimate	Standard Error	Lower 95% Confidence Interval	Upper 95% Confidence Interval
A. Unrestricted Estimation (allowing r_s and r_c to differ)				
α	0.251	0.039	0.175	0.327
r_s	0.633	0.059	0.517	0.748
r_c	0.802	0.056	0.692	0.913
μ_1	0.203	0.012	0.180	0.226
μ_2	0.276	0.024	0.228	0.324
μ_3	1.181	0.191	0.806	1.556
μ_4	0.640	0.041	0.558	0.721
B. Restricted Estimation (assuming $r_s = r_c$)				
α	0.164	0.013	0.139	0.188
r	0.718	0.043	0.634	0.802
μ_1	0.199	0.011	0.178	0.220
μ_2	0.262	0.021	0.222	0.302
μ_3	0.917	0.113	0.695	1.139
μ_4	0.652	0.043	0.567	0.736

All data pooled. Observations: 30,690; # clusters: 186

Hypothesis Testing

Test	χ^2 test statistic	p-value
$r_s = r_c$	6.28	0.01
α (unrestricted) = α (restricted)		$< .01$
Additional Statistics		
Log-likelihood (restricted)		-15,523
Log-likelihood (unrestricted)		-15,494

model presented in Panel A show that the curvature over self payoffs and the curvature over charity payoffs are significantly different and that subjects display significantly more aversion to charity risk than to self risk (p-value=0.01).

Result 1. *Consistent with Hypothesis 1, we reject the hypothesis that $r_s = r_c$. Further, we find that subjects are more averse to charity risk than to self risk, i.e., $r_c < r_c$.*

Second, consistent with Hypothesis 2 and the theoretical prediction in Gauriot, Heger, and Slonim (2020), the altruism parameter, α, is significantly underestimated in the restricted model (Panel B) relative to the unrestricted model (Panel A): the unrestricted estimate is 0.251 in Panel A and 0.164 in Panel B. We find that this difference is statistically significant (p-value$< .01$) and thus ignoring differences in aversion to charity risk and self risk results in an significant under-estimation of altruism.[11]

Result 2. *Consistent with Hypothesis 2, we find that assuming a single parameter over curvature, i.e., $r_s = r_c = r$, leads to a significant underestimation of altruism.*

In equation 4, the background wealth, m, is included in the expected utility of the subject. In the domain of the experiment, we have considered that the show-up fee the subject earns serves as the relevant background wealth (i.e., either \$5 or \$20). To ensure that our results are not driven by inclusion of the show-up fee as background wealth in our estimation, we re-estimate our main results excluding the show-up fee from the estimating equation and display these results in Table A1. These results are also consistent with Hypotheses 1 and 2: we reject that $r_s = r_c$ in the unrestricted model ($\chi^2 = 47.12$, p-value< 0.001) and we reject the hypothesis that α (unrestricted) $= \alpha$ (restricted) (p-value< 0.01).

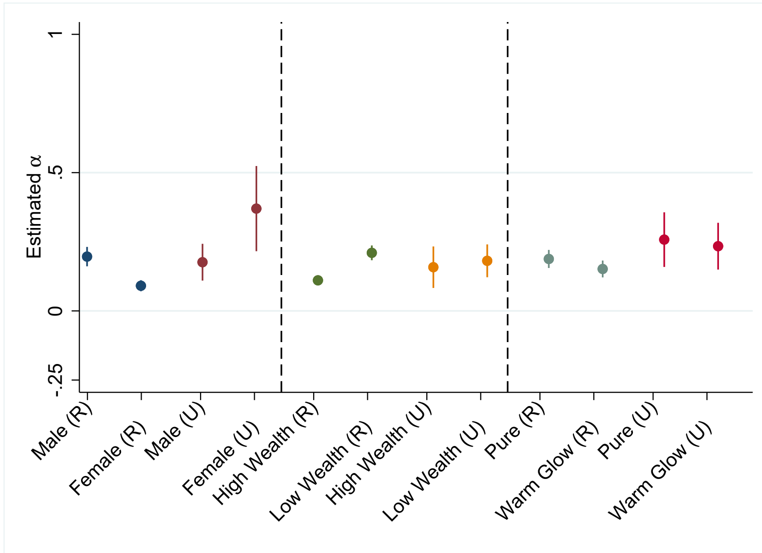
3.1 Heterogeneity: Comparisons of Altruism Across Subgroups

In this subsection, we examine whether imposing strict assumptions on risk preferences will lead to biased estimates of altruism across gender, wealth and giving motives. We focus the presentation of the results on comparing the conclusions drawn about altruism in the restricted (i.e., $r_s = r_c$) model versus those drawn from the unrestricted model (i.e., $r_s \neq r_c$) to understand whether restricting (incorrectly) to a single parameter for curvature leads to incorrect conclusions about the differences in altruism across gender, wealth and altruism motives.

We summarize our findings in Figure 1 where we plot the coefficient estimate of α and the corresponding 95% confidence interval for each relevant subgroup that come from a full estimation of the restricted (R) model and the unrestricted (U) model. While we focus only on the estimate of α in this section, we present the full regression

[11] To conduct the test $X = \alpha$ (unrestricted) $= \alpha$ (restricted), we compute the standard error of $\alpha_u - \alpha_r$ by bootstrap. We re-sample, with replacement, our sample 1000 times and compute $X^* = \alpha_u - \alpha_r$ for each bootstrapped sample. The standard deviation of the 1000 X^*'s is the estimate of the standard error.

Figure 1: Altruism Across Subgroups: Restricted versus Unrestricted Model



This figure plots the coefficient of altruism (α) estimated from the restricted (R) and unrestricted (U) models with 95% confidence intervals. The restricted model predicts that men are significantly more altruistic than women and that the low wealth individuals are more altruistic than the high wealth individuals. By contrast, the unrestricted model predicts that women are significantly more altruistic than men and that there are no differences in altruism across wealth.

results in Tables A2, A3, and A4. We find that across each of the relevant subgroups, gender, wealth and motives, the restricted model and the unrestricted model lead to different conclusions.

Gender In terms of gender, we find that the conclusion of the restricted model is that women are less altruistic than men (p-value$<$ 0.01), while the conclusion of the unrestricted model is that women are more altruistic than men (p-value=0.04). This is driven by the larger divergence of r_s and r_c for women than for men in the unrestricted model; that is, women and men do not significantly differ in their value of r_s, but women have utility functions over charity payoffs that are significantly more concave than those of men. Thus, the restricted model underestimates the altruism parameter *more* for women than for men, leading to the incorrect conclusion that women are less altruistic than men.

Wealth Next, we compare differences that may arise due to unobserved wealth by comparing subjects who receive a \$5 show-up fee (baseline) to those who receive a \$20 show-up fee (high wealth). The CRRA utility function displays decreasing absolute risk aversion, meaning that, *ceteris paribus*, wealthier individuals are willing to take on more risk. Our results on wealth show that the estimates from the restricted model would lead to the conclusion that high wealth individuals (i.e., \$20 show-up fee treatment) are less altruistic than the low wealth individuals (i.e., \$5 show-up fee treatment) (p-value<0.01). However, in the unrestricted model there is no difference in estimated altruism (p-value=0.58). Again, this is explained by the shortcomings of the restricted model. In the restricted model, the single parameter for risk does not differ between the low and high wealth subjects. However, when we separate self risk from charity risk, we find that charity risk does not vary across wealth, but self risk does—the low wealth subjects are more risk averse over self payoffs than high wealth subjects. Thus, the difference between r_s and r_c is greater for high wealth individuals than for the low wealth subjects implying that the restricted model underestimates altruism among high wealth individuals more than the low wealth subjects.

Donor Motives Finally, we consider the effect of different motives for giving on the curvature of the utility function. DellaVigna, List, and Malmendier (2012) hypothesize, but do not test, that the utility function of a donor who gives for warm glow motives would be more concave than a donor giving for pure altruism motives. Consistent with this notion, Null (2011) models utility from pure altruism as linear and utility from warm glow as concave. Our estimates in Table A4 are not consistent with the assumptions of their models (i.e., we find no differences in curvature between the Warm Glow and Pure Altruism treatments), but it could be that our priming of warm glow and pure altruism were not successful. However, we compare subjects that are (potentially) motivated by warm glow versus pure altruism differ in their estimated level of altruism and find no significant differences in the restricted model (p-value=0.34) or the unrestricted model (p-value=0.54).

Result 3. *Ignoring heterogeneity in the curvature of utility across subgroups, resulted in biased estimates of altruism across gender and wealth.*

Table 3: Parameter Estimates with Risk and without Risk

Parameter	Estimate	Standard Error	Lower 95% Confidence Interval	Upper 95% Confidence Interval
A. Unrestricted Estimation (allowing r_s and r_c to differ)				
α	0.349	0.192	-0.027	0.726
$\alpha \times$ Risk	2.390	1.661	-0.865	5.645
r_s	0.264	0.088	0.091	0.437
$r_s \times$ Risk	-0.306	0.105	-0.513	-0.100
r_c	0.338	0.127	0.090	0.587
$r_c \times$ Risk	0.629	0.125	0.384	0.875
μ_3	2.940	0.869	1.236	4.644
μ_4	0.551	0.028	0.495	0.607
B. Restricted Estimation (assuming $r_s = r_c$)				
α	0.284	0.034	0.216	0.351
$\alpha \times$ Risk	-0.181	0.034	-0.247	-0.115
r	0.286	0.053	0.183	0.389
$r \times$ Risk	-0.244	0.056	-0.352	-0.135
μ_3	2.734	0.465	1.823	3.646
μ_4	0.697	0.043	0.612	0.782

Only data from Task 3 and 4 are used for these estimates. Observations: 17,298; # clusters: 186.

Hypothesis Testing

Test	χ^2 test statistic	p-value
$r_s = r_c$	0.15	0.70
$r_s + r_s \times$ Risk $= r_c + r_c \times$ Risk	18.92	< 0.001
α (unrestricted) $= \alpha$ (restricted)		< 0.01
α_{Risk} (unrestricted) $= \alpha_{\text{Risk}}$ (restricted)		< 0.01
Additional Statistics		
Log-likelihood (restricted)		-9,112
Log-likelihood (unrestricted)		-9,042

3.2 The Role of Risk

In Table 3 we display results from our estimation of the unrestricted model and the restricted model that account for whether there was risk involved in the decision task or not. There are three key findings. First, we reject the hypothesis that $r_s = r_c$ (p-value< 0.001) in environments with risk but we cannot reject the hypothesis that $r_s = r_c$ (p-value=0.70) in environments without risk.

Second, in the restricted model and consistent with Exley (2016), we find that subjects are less altruistic in risky environments than in environments without risk. In the unrestricted model, our estimate of altruism is large and imprecise, thus we do not think we can draw any concrete conclusions about the difference in altruism under risk in the unrestricted model.

Third, our results allow us to speak to the question of whether preferences, specifically the curvature parameter of utility, are different when elicited in environments with and without risk, which has been a topic of much debate in the time preference literature (Andreoni and Sprenger, 2012b; Miao and Zhong, 2015; Abdellaoui et al., 2013; Cheung, 2015; Andreoni and Sprenger, 2015).[12] In the restricted model we find utility is *less* concave when measured under risk (i.e., task 4) than when measured in environments without risk (i.e., task 3). However, the restricted model hides important differences between the curvature parameter over own payoffs (r_s) and the curvature parameter over charity payoffs (r_c): similar to the restricted model, curvature over self payoffs is *less* concave when measured in environments with risk than without risk, while the curvature over charity payoffs is more concave when measured in environments with risk than without risk.

3.3 Robustness to Additional Utility Specifications

In this section, we ask whether our results are robust to two additional utility specifications. In particular, we explore whether the results in Section 3 extend to a (1) power utility and (2) CRRA specification with probability weighting.

Power utility is given by the following equation

$$U(s,c) = s^{r_s} + \alpha c^{r_c} \tag{8}$$

where, as before, the restricted model assumes $r_s = r_c$ while the unrestricted model does not.

Table 4 presents the results of the power utility specification given in equation

[12] Andreoni and Sprenger (2012b) use the convex budget approach to elicit curvature and time preferences and find that utility is more concave in the presence of risk than without risk. Cheung (2015), on the other hand, uses a multiple price list aproach and finds no such differences and claims that the difference found in Andreoni and Sprenger (2012b) is an artefact of their elicitation procedure. preferences depend on whether they are measured in an environment of risk or not.

8. The unrestricted model shows that the curvature parameter over the charity's payoffs, r_c, is approaching 0, indicating an extremely high degree of concavity. This high degree of concavity means that for any donation amount c the utility value is quite small. To counteract this, α is estimated to be extremely large. However, the model converges and, as in Table 2, we find that we can reject the hypothesis that $r_s = r_c$ and that altruism is significantly underestimated in the restricted model.

Table 4: Robustness II: Power Utility

Parameter	Estimate	Standard Error	Lower 95% Confidence Interval	Upper 95% Confidence Interval
A. Unrestricted Estimation (allowing r_s and r_c to differ)				
α	$1.8e+08$	$1.1e+07$	$1.6e+08$	$2.1e+08$
r_s	0.356	0.022	0.314	0.399
r_c	0.000	0.000	0.000	0.000
μ_1	0.202	0.011	0.181	0.223
μ_2	0.326	0.027	0.273	0.379
μ_3	0.424	0.055	0.317	0.531
μ_4	0.682	0.036	0.611	0.752
B. Restricted Estimation (assuming $r_s = r_c$)				
α	0.164	0.013	0.139	0.188
r	0.282	0.043	0.198	0.366
μ_1	0.199	0.011	0.178	0.220
μ_2	0.262	0.021	0.222	0.302
μ_3	0.258	0.068	0.126	0.391
μ_4	0.652	0.043	0.567	0.736

All data pooled. Observations: 30,690; # clusters: 186

Hypothesis Testing

Test	χ^2 test statistic	p-value
$r_s = r_c$	272.59	$< .01$
α (unrestricted) $= \alpha$ (restricted)		$< .01$

Next, we explore whether our main CRRA specification is robust to the inclusion of probability weighting given by

$$w(p) = \frac{p^\gamma}{[p^\gamma + (1-p)^\gamma]^{\frac{1}{\gamma}}} \tag{9}$$

Table 5 displays the estimates from our CRRA specification with probability weighting from equation 9. Consistent with the results in Table 2, we reject the hypothesis that $r_s = r_c$ and specifically, find that the utility function over the charity's payoffs is more concave than the utility function over self payoffs and that incorrectly assuming they are equal leads to significant underestimation of altruism. Further, in both the restricted and unrestricted model, we do not find evidence consistent with probability weighting, specifically, we cannot reject the null hypothesis that $\gamma = 1$.

Table 5: Robustness I: CRRA Utility with Probability Weighting

Parameter	Estimate	Standard Error	Lower 95% Confidence Interval	Upper 95% Confidence Interval
A. Unrestricted Estimation (allowing r_s and r_c to differ)				
α	0.255	0.049	0.159	0.351
r_s	0.632	0.061	0.513	0.752
r_c	0.812	0.067	0.681	0.942
γ	1.029	0.093	0.846	1.212
μ_1	0.207	0.020	0.168	0.246
μ_2	0.281	0.031	0.221	0.341
μ_3	1.187	0.206	0.783	1.590
μ_4	0.644	0.044	0.557	0.731
B. Restricted Estimation (assuming $r_s = r_c$)				
α	0.168	0.012	0.144	0.193
r	0.700	0.039	0.623	0.777
γ	0.886	0.061	0.766	1.005
μ_1	0.182	0.014	0.155	0.209
μ_2	0.242	0.023	0.198	0.286
μ_3	0.957	0.106	0.750	1.165
μ_4	0.624	0.045	0.536	0.711

All data pooled. Observations: 30,690; # clusters: 186

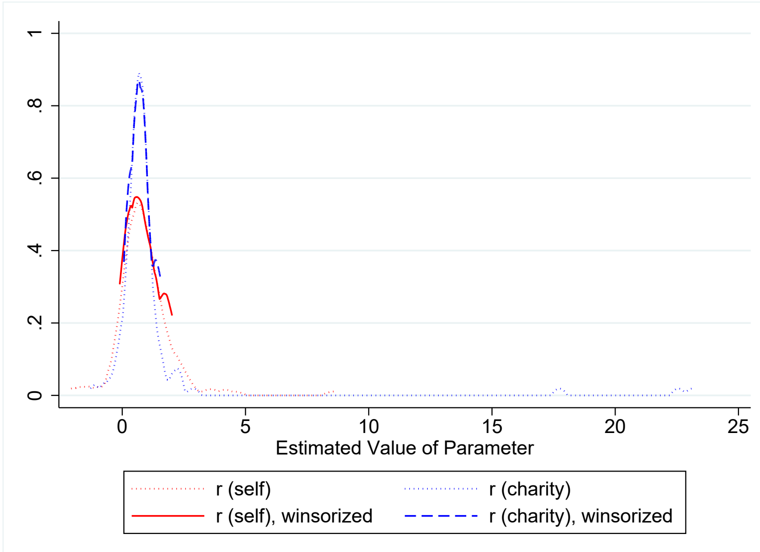
Hypothesis Testing

Test	χ^2 test statistic	p-value
$r_s = r_c$	4.08	0.04
α (unrestricted) = α (restricted)		$< .01$

3.4 Robustness to Unobserved Heterogeneity

In this section, we exploit our within subject design feature, which asked each subject to make 165 choices, to estimate the restricted and unrestricted model at the individual level. By doing this, we are able to account for unobserved heterogeneity across

FIGURE 2: Distribution of Estimated Curvature Parameters



The distribution of estimates of r_s and r_c from the unrestricted model.

individuals. Of the 186 individuals, our maximum likelihood estimation converged for 144 subjects in the restricted model and 124 in the unrestricted model.[13]

Figure 2 shows the distribution of estimates of the curvature parameters from the unrestricted model (r_s, r_c). However, as Figure 2 shows clearly the large outliers present in the individual estimates, particularly in the estimates of r_c. To deal with the outliers, we "winsorize" the distributions of the r_s and r_c to reduce the effect that spurious outliers have on the data (Hastings Jr et al., 1947). We winsorize at the 80% level, which transforms the largest and smallest 10% of estimates of r_s and r_c and assigns them the value of the next largest or smallest estimate, respectively. We show the distributions of the original and the winsorized estimates in Figure 2. A matched pairs test of r_s and r_c does not reject the hypothesis that $r_s = r_c$ (p-value=0.24); however, a matched pairs test of the winsorized values of r_s and r_c *rejects* the hypothesis that $r_s = r_c$ (p-value=0.07).

Next, we compare the distribution of altruism resulting from the restricted and unrestricted model across the same subgroups as in Figure 1 and report these tests in Table 6. In sum, we find that the restricted model leads us to different conclusions

[13]These convergence rates are consistent with similar studies. For example, Andersen et al. (2014b) obtains individual estimates for 44%-78% of subjects.

Table 6: Testing α Across Subgroups

Test	test statistic	p-value
A. Gender: Male versus Female		
Mann-Whitney test		
α_f (restricted) = α_m (restricted)	1.81	0.07
α_f (unrestricted) = α_m (unrestricted)	0.744	0.46
Test of Equality of Variances		
α_f (restricted) = α_m (restricted)	3.88	0.05
α_f (unrestricted) = α_m (unrestricted)	0.18	0.67
B. Wealth: $5 versus $20		
Test	test statistic	p-value
Mann-Whitney test		
α_5 (restricted) = α_{20} (restricted)	1.38	0.17
α_5 (unrestricted) = α_{20} (unrestricted)	1.83	0.07
B. Motives: Warm Glow versus Pure		
Test	test statistic	p-value
Mann-Whitney test		
α_P (restricted) = α_{WG} (restricted)	1.48	0.13
α_P (unrestricted) = α_{WG} (unrestricted)	0.41	0.69

about the distribution of altruism across gender and wealth than we would arrive at under the unrestricted model. Specifically, we find gender differences in altruism from the estimates in the restricted model, but which disappear in the unrestricted model. Further, we find gender differences in the standard deviation of the altruism distribution in the restricted model, but not in the unrestricted model. We find no wealth differences in the distribution of altruism in the restricted model, but we do find evidence of wealth differences in the unrestricted model.

4 Conclusion

Using data from a laboratory experiment we extend the insights gained in Andersen et al. (2008) to the literature on social preferences and provide strong empirical support for the theoretical results in Gauriot, Heger, and Slonim (2020). First, subjects are significantly more averse to charity risk than to self risk. Second, assuming that there is a single risk preference parameter leads to significant underestimation of altruism and to incorrect inferences about altruism across subgroups.

The lingering question is how might these results inform future experimental design and preference elicitation? While Gauriot, Heger, and Slonim (2020) provides some suggestions on how the literature might proceed in light of their theoretical findings, our empirical results shed an important new light on a path forward. Specifically, we find no differences between curvature over own payoffs and curvature over charity payoffs *in the absence of risk*. Thus, environments without risk may be less prone to the bias induced by the "single parameter" assumption than environments with risk.

References

Abdellaoui, Mohammed, Han Bleichrodt, Olivier l'Haridon, and Corina Paraschiv. 2013. "Is there one unifying concept of utility? An experimental comparison of utility under risk and utility over time." *Management Science* 59 (9):2153–2169.

Andersen, Steffen, John Fountain, Glenn W Harrison, and E Elisabet Rutström. 2014a. "Estimating subjective probabilities." *Journal of Risk and Uncertainty* 48 (3):207–229.

Andersen, Steffen, Glenn W Harrison, Morten I Lau, and E Elisabet Rutström. 2008. "Eliciting risk and time preferences." *Econometrica* 76 (3):583–618.

———. 2014b. "Discounting behavior: A reconsideration." *European Economic Review* 71:15–33.

———. 2018. "Multiattribute Utility Theory, Intertemporal Utility, and Correlation Aversion." *International Economic Review* 59 (2):537–555.

Andreoni, James and Charles Sprenger. 2012a. "Estimating time preferences from convex budgets." *American Economic Review* 102 (7):3333–56.

———. 2012b. "Risk preferences are not time preferences." *American Economic Review* 102 (7):3357–76.

———. 2015. "Risk preferences are not time preferences: reply." *American Economic Review* 105 (7):2287–93.

Blavatskyy, Pavlo R. 2011. "A model of probabilistic choice satisfying first-order stochastic dominance." *Management Science* 57 (3):542–548.

Brown, Alexander L and Paul J Healy. 2018. "Separated decisions." *European Economic Review* 101:20–34.

Cheung, Stephen L. 2015. "Risk preferences are not time preferences: on the elicitation of time preference under conditions of risk: comment." *American Economic Review* 105 (7):2242–60.

———. 2020. "Eliciting utility curvature in time preference." *Experimental Economics* 23 (2):493–525.

Croson, Rachel and Uri Gneezy. 2009. "Gender differences in preferences." *Journal of Economic Literature* :448–474.

DellaVigna, S, JA List, and U Malmendier. 2012. "Testing for altruism and social pressure in charitable giving." *The Quarterly Journal of Economics* 127 (1):1.

Exley, Christine L. 2016. "Excusing selfishness in charitable giving: The role of risk." *The Review of Economic Studies* 83 (2):587–628.

Fischbacher, Urs. 2007. "z-Tree: Zurich toolbox for ready-made economic experiments." *Experimental Economics* 10 (2):171–178.

Fisman, Raymond, Pamela Jakiela, and Shachar Kariv. 2015. "How did distributional preferences change during the great recession?" *Journal of Public Economics* 128:84–95.

———. 2017. "Distributional preferences and political behavior." *Journal of Public Economics* .

Fisman, Raymond, Pamela Jakiela, Shachar Kariv, and Daniel Markovits. 2015. "The distributional preferences of an elite." *Science* 349 (6254):aab0096.

Fisman, Raymond, Shachar Kariv, and Daniel Markovits. 2007. "Individual preferences for giving." *The American Economic Review* 97 (5):1858–1876.

Gauriot, Romain, Stephanie A Heger, and Robert Slonim. 2020. "Altruism or diminishing marginal utility?" *Journal of Economic Behavior & Organization* 180:24–48.

Greiner, Ben. 2015. "Subject pool recruitment procedures: organizing experiments with ORSEE." *Journal of the Economic Science Association* 1 (1):114–125.

Harrison, Glenn W. 2018. "The methodologies of behavioral econometrics." .

Harrison, Glenn W, Morten I Lau, and E Elisabet Rutström. 2013. "Identifying time preferences with experiments: Comment." *Center for the Economic Analysis of Risk, Working Paper* 9.

Harrison, Glenn W and E Rutström. 2008a. "Risk aversion in the laboratory." In *Risk aversion in experiments*. Emerald Group Publishing Limited, 41–196.

Harrison, Glenn W and E Elisabet Rutström. 2008b. "Risk aversion in the laboratory." In *Risk aversion in experiments*. Emerald Group Publishing Limited.

Hastings Jr, Cecil, Frederick Mosteller, John W Tukey, and Charles P Winsor. 1947. "Low moments for small samples: a comparative study of order statistics." *The Annals of Mathematical Statistics* 18 (3):413–426.

Hey, John D and Chris Orme. 1994. "Investigating generalizations of expected utility theory using experimental data." *Econometrica: Journal of the Econometric Society* :1291–1326.

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

Lilley, Andrew and Robert Slonim. 2014a. "The price of warm glow." *Journal of Public Economics* 114:58–74.

———. 2014b. "The price of warm glow." *Journal of Public Economics* 114:58–74.

Miao, Bin and Songfa Zhong. 2015. "Risk preferences are not time preferences: separating risk and time preference: comment." *American Economic Review* 105 (7):2272–86.

Null, Clair. 2011. "Warm glow, information, and inefficient charitable giving." *Journal of Public Economics* 95 (5):455–465.

Wilcox, Nathaniel T. 2011. "Stochastically more risk averse: A contextual theory of stochastic discrete choice under risk." *Journal of Econometrics* 162 (1):89–104.

Appendix

Appendix.1 Additional Results

Table A1: Main Result: Parameter Estimates from the Unrestricted and Restricted Model Without Showup Fee Included in Estimation

Parameter	Estimate	Standard Error	Lower 95% Confidence Interval	Upper 95% Confidence Interval
		A. Unrestricted Estimation (allowing r_s and r_c to differ)		
α	0.587	0.058	0.474	0.701
r_s	0.435	0.029	0.378	0.492
r_c	0.739	0.049	0.644	0.834
μ_1	0.198	0.010	0.178	0.218
μ_2	0.265	0.021	0.223	0.306
μ_3	2.640	0.201	2.246	3.033
μ_4	0.584	0.031	0.522	0.646
		B. Restricted Estimation (assuming $r_s = r_c$)		
α	0.297	0.015	0.269	0.326
r	0.527	0.028	0.473	0.581
μ_1	0.198	0.011	0.177	0.219
μ_2	0.256	0.019	0.218	0.293
μ_3	2.075	0.152	1.776	2.374
μ_4	0.567	0.030	0.509	0.625

All data pooled. Observations: 30,690; # clusters: 186

Hypothesis Testing

Test	χ^2 test statistic	p-value
$r_s = r_c$	47.12	< 0.001
α (unrestricted) = α (restricted)		$< .01$

Table A2: Heterogeneity I: Gender

Parameter	Estimate	Standard Error	Lower 95% Confidence Interval	Upper 95% Confidence Interval
A. Unrestricted Estimation (allowing r_s and r_c to differ)				
α	0.176	0.034	0.110	0.243
$\alpha \times$ Female	0.194	0.074	0.049	0.338
r_s	0.633	0.063	0.510	0.757
$r_s \times$ Female	0.017	0.049	-0.079	0.113
r_c	0.587	0.064	0.462	0.712
$r_c \times$ Female	0.456	0.106	0.248	0.664
μ_1	0.202	0.012	0.179	0.225
μ_4	0.274	0.025	0.225	0.322
μ_3	1.128	0.177	0.781	1.476
μ_{f}	0.609	0.036	0.538	0.679
B. Restricted Estimation (assuming $r_s = r_c$)				
α	0.196	0.018	0.161	0.231
$\alpha \times$ Female	-0.105	0.021	-0.146	-0.065
r	0.669	0.048	0.576	0.763
$r \times$ Female	0.125	0.051	0.026	0.224
μ_1	0.197	0.011	0.175	0.218
μ_2	0.258	0.020	0.219	0.298
μ_3	0.936	0.115	0.709	1.162
μ_4	0.657	0.039	0.580	0.735

All data pooled. Observations: 30,690; # clusters: 186.

Additional Statistics

Log-likelihood (restricted)	-15,449
Log-likelihood (unrestricted)	-15,351

Table A3: Heterogeneity II: Wealth

Parameter	Estimate	Standard Error	Lower 95% Confidence Interval	Upper 95% Confidence Interval
A. Unrestricted Estimation (allowing r_s and r_c to differ)				
α	0.158	0.038	0.083	0.233
$\alpha \times$ Low Wealth	0.023	0.043	-0.061	0.107
r_s	0.665	0.056	0.556	0.774
$r_s \times$ Low Wealth	0.112	0.049	0.015	0.208
r_c	0.783	0.068	0.649	0.916
$r_c \times$ Low Wealth	-0.065	0.094	-0.249	0.119
μ_1	0.200	0.011	0.178	0.223
μ_2	0.265	0.021	0.224	0.306
μ_3	0.807	0.120	0.571	1.043
μ_4	0.571	0.030	0.512	0.629
B. Restricted Estimation (assuming $r_s = r_c$)				
α	0.111	0.007	0.097	0.125
$\alpha \times$ Low Wealth	0.099	0.015	0.070	0.128
r	0.686	0.043	0.601	0.771
$r \times$ Low Wealth	0.077	0.048	-0.017	0.172
μ_1	0.200	0.011	0.178	0.221
μ_2	0.265	0.021	0.224	0.307
μ_3	0.816	0.098	0.623	1.009
μ_4	0.575	0.030	0.516	0.633

All data pooled. Observations: 30,690; # clusters: 186

Additional Statistics

Log-likelihood (restricted)	-15,311
Log-likelihood (unrestricted)	-15,286

Table A4: Heterogeneity III: Motives

Parameter	Estimate	Standard Error	Lower 95% Confidence Interval	Upper 95% Confidence Interval
A. Unrestricted Estimation (allowing r_s and r_c to differ)				
α	0.216	0.039	0.140	0.292
$\alpha \times$ Warm Glow	−0.034	0.054	−0.139	0.072
r_s	0.695	0.063	0.572	0.818
$r_s \times$ Warm Glow	0.073	0.059	−0.043	0.188
r_c	0.735	0.081	0.576	0.893
$r_c \times$ Warm Glow	−0.040	0.117	−0.269	0.190
μ_1	0.187	0.013	0.161	0.214
μ_2	0.248	0.022	0.205	0.290
μ_3	0.942	0.146	0.655	1.229
μ_4	0.588	0.037	0.515	0.660
B. Restricted Estimation (assuming $r_s = r_c$)				
α	0.194	0.016	0.162	0.226
$\alpha \times$ Warm Glow	0.026	0.028	−0.028	0.080
r	0.699	0.054	0.593	0.805
$r \times$ Warm Glow	0.048	0.059	−0.067	0.163
μ_1	0.187	0.013	0.161	0.213
μ_2	0.250	0.022	0.206	0.294
μ_3	0.964	0.126	0.716	1.211
μ_4	0.588	0.037	0.516	0.660

In this regression, we exclude subjects who received a $20 show-up fee because there were no such subjects in the Pure Altruism treatment. Observations: 20,460; # clusters: 124

Additional Statistics

Log-likelihood (restricted)	-10,077
Log-likelihood (unrestricted)	-10,072

Appendix.2 Experimental payoffs

Table A5: Task 1 & Task 2 Lotteries

Decision	h_1	l_1	h_2	l_2	p
1-9	45	30	80	10	$\forall p \in \{.1, .2..., .9\}$
10-18	40	35	85	3	$\forall p \in \{.1, .2..., .9\}$
19-27	40	32	77	2	$\forall p \in \{.1, .2..., .9\}$
28-36	50	20	90	1	$\forall p \in \{.1, .2..., .9\}$

Lotteries used in Task 1 and Task 2. The standard Holt and Laury task includes $p = 1$, which we excluded as it is not informative about risk aversion. All amounts are reported in Australian dollars.

Table A6: Task 3 Allocation Choices

Decision	s_A	o_A	s_B	o_B
1	18	5	14	15
2	14	15	8	30
3	8	30	2	45
4	18	4	14	12
5	14	12	8	24
6	8	24	2	36
7	18	3	14	9
8	14	9	8	18
9	8	18	2	27
10	18	2	14	6
11	14	6	8	12
12	8	12	2	18
13	27	5	21	15
14	21	15	12	30
15	12	30	3	45
17	21	12	12	24
18	12	24	3	36
19	27	3	21	9
20	21	9	12	18
21	12	18	3	27
22	27	2	21	6
23	21	6	12	12
24	12	12	3	18
25	36	5	28	15
26	28	15	16	30
27	16	30	4	45
28	36	4	28	12
29	28	12	16	24
30	16	24	4	36
31	36	3	28	9
32	28	9	16	18
33	16	18	4	27
34	36	2	28	6
35	28	6	16	12
36	16	12	4	18
37	45	5	35	15
38	35	15	20	30
39	20	30	5	45
40	45	4	35	12
41	35	12	20	24
42	20	24	5	36
43	9	4	7	12
44	7	12	4	24
45	4	24	1	36
46	9	5	7	15
47	7	15	4	30

Table A7: Task 4 Lottery Choices

	Lottery A				Lottery B				
Choice List	s_A^h	o_A^h	s_A^l	o_A^l	s_B^h	o_B^h	s_B^l	o_B^l	p
1-9	10	5	10	5	5	40	5	5	$\forall p \in \{.1, .2..., .9\}$
10-18	4	10	4	10	10	50	1	5	$\forall p \in \{.1, .2..., .9\}$
19-27	5	10	5	10	1	50	5	1	$\forall p \in \{.1, .2..., .9\}$
28-36	36	1	5	1	7	15	7	15	$\forall p \in \{.1, .2..., .9\}$
37-45	45	3	15	3	10	20	10	20	$\forall p \in \{.1, .2..., .9\}$

Lotteries used in Task 4, the risky dictator game. Payoffs are reported in Australian dollars.

Appendix.3 Graphical representation of experimental task

FIGURE A1: Graphical representation of the decision in Task 1 as shown to the subjects.

FIGURE A2: Graphical representation of the decision in Task 2 as shown to the subjects.

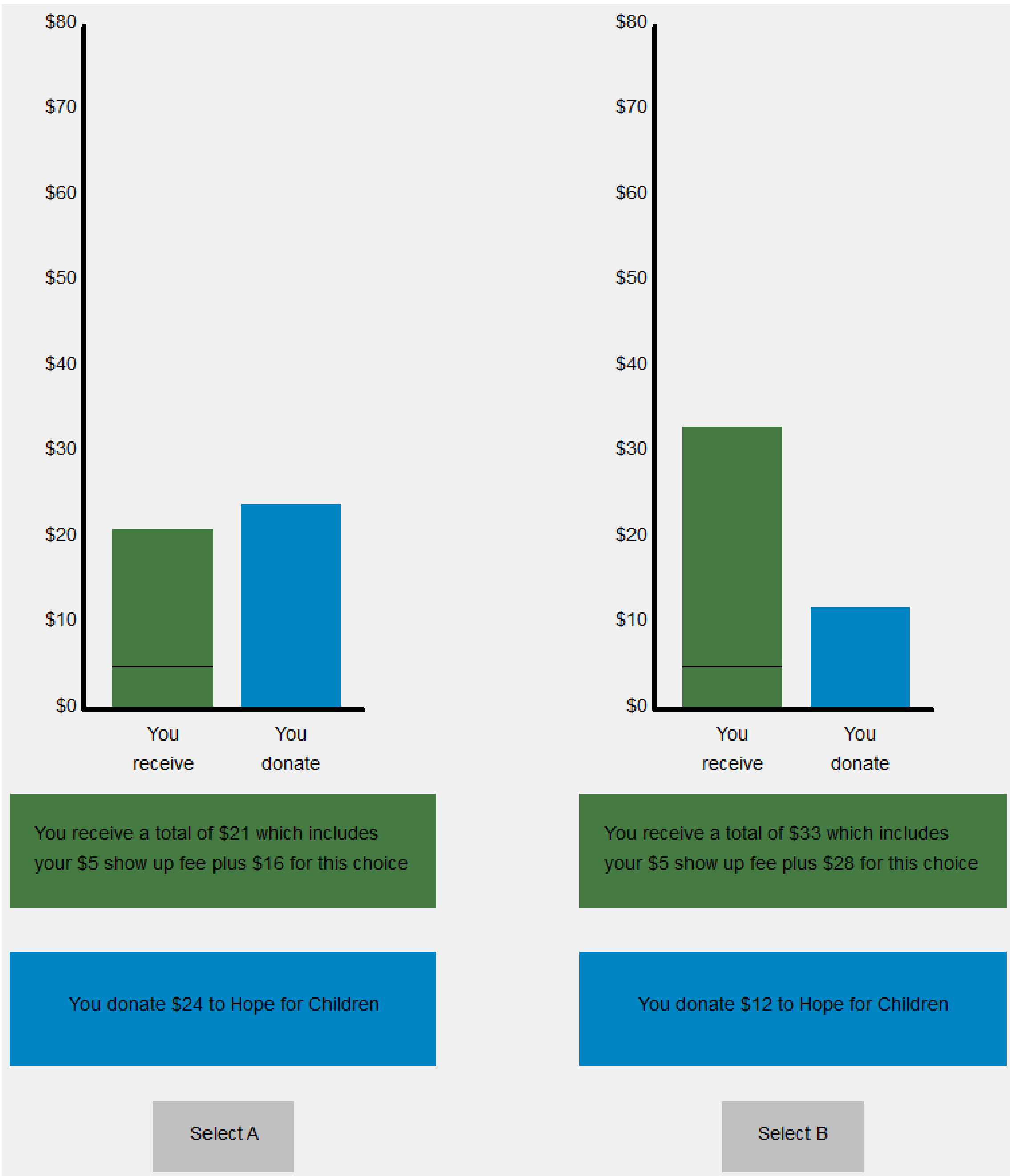


FIGURE A3: Graphical representation of the decision in Task 3 as shown to the subjects.

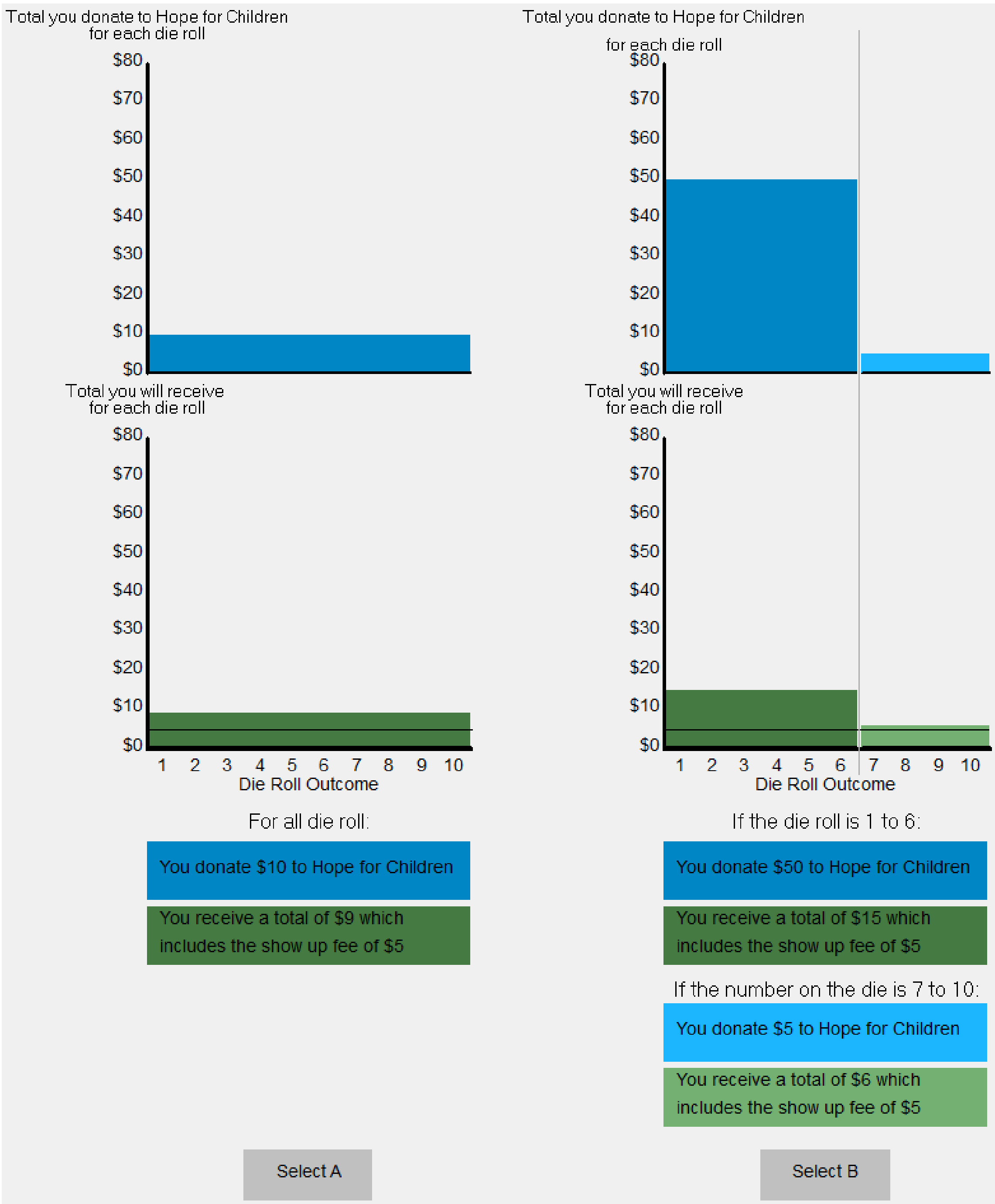


FIGURE A4: Graphical representation of the decision in Task 4 as shown to the subjects.