

Karni, Edi; Salmon, Timothy C.; Sopher, Barry

Working Paper

Individual sense of fairness: An experimental study

Working Paper, No. 2001-07

Provided in Cooperation with:

Department of Economics, Rutgers University

Suggested Citation: Karni, Edi; Salmon, Timothy C.; Sopher, Barry (2001) : Individual sense of fairness: An experimental study, Working Paper, No. 2001-07, Rutgers University, Department of Economics, New Brunswick, NJ

This Version is available at:

<https://hdl.handle.net/10419/79147>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

INDIVIDUAL SENSE OF FAIRNESS: AN EXPERIMENTAL STUDY

Edi Karni
Johns Hopkins University

Timothy C. Salmon
Florida State University

Barry Sopher
Rutgers University

January 30, 2001

Abstract

1 Introduction

By and large neo-classical economics is founded on the notion of self-interest seeking behavior which means that individual choice behavior is motivated solely by the desire to attain higher levels of *material well-being*. This presumption stands in stark contrast to long held views in philosophy and psychology maintaining that human behavior is motivated in part by emotions and, in particular, by moral sentiments¹. Recently, however, there is growing interest among economists in the potential implications of broadening the psychological base of the model of individual behavior by incorporating emotions into the theory of choice. (See, for example, a survey by Elster [1998] and discussions by Loewenstein [2000], Romer [2000].) This interest is partly due to mounting experimental evidence of cooperative behavior in situations in which maximization of material self-interest alone would imply non-cooperative behavior. (See, for instance, Camerer (1997), Berg, Dickhaut, and McCabe (1995).)

In this paper we explore, via experiments, some issues pertaining to presence of intrinsic sense of fairness as a motive force in individual choice behavior. Specifically, confronting subjects with choices among allocation procedures that involves random selection of a winner of predetermined prize, we test the willingness of these subjects to sacrifice their chance of winning in order to attain what they perceive to be fairer allocation procedure. This work was inspired by two recent papers of Karni and Safra (2000, 2000a), the first of which contains an axiomatic model of choice behavior that is motivated, in part, by concern

¹See Hume (1740), Smith (1756), Rowls (1963, 1971) for philosophical discussions. (Barry and Tim: We need to add some references to the psychology literature. Any suggestions?)

for fairness and the second introduces measures of the intensity of the sense of fairness and derive their behavioral characterizations. Our experimental design is based on the analytical framework of Karni and Safra (2000). Our results, therefore, constitute a direct test of their contention concerning the manifestation inherent sense of fairness in individual behavior.

To set the stage we summarize briefly those elements of the theory of Karni and Safra that are relevant for our work. We focus on aspects of the axiomatic theory that underlie the experimental design and on aspects of the measurement of the intensity of the sense of fairness that are relevant for the interpretation of our findings.

Let $N = \{1, \dots, n\}$, $n > 2$, be a set of individuals who must decide on a procedure by which to allocate, among themselves, one unit of an indivisible good. Clearly, the ex-post allocation is necessarily unfair. One individual is awarded the good and the others are not. The issue, therefore, is what allocation procedure may be implemented to attain higher level of fairness ex-ante. Karni and Safra (2000) restrict attention to procedures that allocate the good by lot. Formally, denote by e^i , the unit vector in \mathbb{R}^n , the *ex-post* allocation in which individual i is assigned the good. Let $X = \{e^i \mid 1 \leq i \leq n\}$ be the set of *ex-post* allocations and let $P = \Delta(X)$ be the $n-1$ dimensional simplex representing the set of all probability distributions on X . In this context P has the interpretation of the set of *random allocation procedures*.

Each individual is represented by two binary relations on P , the relation \succsim , representing his actual choice behavior, and the relation \succsim_F , representing his concept of fairness. The relation \succsim has the usual interpretation:, namely, for any pair of allocation procedures p and q in P , $p \succsim q$ means that, if he were to choose between p and q , the individual would choose p or would be indifferent between the two. The fairness relation, \succsim_F , has the interpretation of “fairer than.” In other words, $p \succsim_F q$ means that the allocation procedure p is regarded by the individual as being at least as fair as the allocation procedure q . In general, the notion of fairness may be subjective or objective. In either case, it is intrinsic and, jointly with concern for self-interest, governs the individual’s choice behavior among allocation procedures.

Taking the preference and the fairness relations as primitives, Karni and Safra (2000) show conditions that permit the derivation of the self-interest motive implicit in the individual choice behavior. Loosely speaking, an allocation procedure p is preferred over another allocation procedure q from a self-interest point of view if the two allocation procedures are equally fair and p is preferred over q . Moreover, Karni and Safra introduce axioms that are equivalent to the existence of an *affine* function $\kappa : P \rightarrow \mathbb{R}$ representing \succsim_S , the derived binary relation representing the *self-interest component* of the preference relation \succsim , a strictly quasi-concave function $\sigma : P \rightarrow \mathbb{R}$ that represents the fairness relation \succsim_F , and a utility function V representing the preference relation \succsim as a function of its self-interest and fairness components, i.e., for all allocation procedures, $p, q \in P$,

$$p \succ q \Leftrightarrow V((\kappa \cdot p, \sigma(p))) \geq V((\kappa \cdot q, \sigma(q))).$$

In addition, Karni and Safra (2000) introduce additional conditions under which the function V is additively separable in the self-interest and fairness components. Formally,

$$V((\kappa \cdot p, \sigma(p))) = h(\kappa \cdot p) + \sigma(p),$$

where h is monotonic increasing function.

Observing that individuals that have the same concept of fairness may differ in the intensity of their moral convictions Karni and Safra (2000a) quantify the intensity of the sense of fairness. In particular, analogously to the measurement of risk aversion in the theory of individual decision-making under risk, the intensity of the sense of fairness may be measure by a fairness premium which has the interpretation of the maximal sacrifice of self-interest an individual would be willing to make to attain a given increment in the level of fairness. Formally, $E(p) \subset \mathbb{R}^n$ be the set of ε satisfying $\kappa \cdot \varepsilon = 0$ and $\sigma(p + \varepsilon) < \sigma(p)$ for which the equation $V(\kappa \cdot q, \sigma(q)) = V(\kappa \cdot (p + \varepsilon), \sigma(p + \varepsilon))$ is satisfied for some q that is as fair as p . Define the fairness premium $\pi(p, \varepsilon)$.by the equation

$$V(\kappa \cdot p - \pi(p, \varepsilon), \sigma(p)) = V(\kappa \cdot (p + \varepsilon), \sigma(p + \varepsilon)).$$

Broadly speaking, one individual is said to display more intense sense of fairness than another if his fairness premium function is always larger. Furthermore, among individuals that share the idea of fairness, individuals display stronger sense of fairness if and only if, for any feasible convex set of allocation procedures, they will always select a fairer allocation procedure.

The rest of the paper is organized as follows: In Section 2 we describe the design and the experiments. In Section 3 we present and analyze the findings. The main conclusions and issues raised by this work are summarized in Section 4.

2 The Experiments: Design and Implementation

The experiments described below are intended to test whether subjects are willing to trade-off their own chance of winning so as to attain a fairer lottery by which the winner of an indivisible good is to be determined. The design of these experiments is a modified version of the interface used in Sopher and Narramore(1996). The mechanism for making a choice in the two settings is the same, but the contexts of the choices are much different.

The experiments involved bringing groups of subjects in multiples of three into a computer lab. The subjects were given a verbal introduction to the experiment including an overview of the rules and they were then lead through an interactive help program to make sure that they understood the interface and understood the rules of the experiment. When the subjects had completed the instructions, they were randomly assigned a player type of either A, B or C and assigned into three person groups with one player of each type in each

group. The player A in each group was asked to choose the allocation that will be used to run an actual lottery for a \$15 prize. The question to the players A was phrased as “Please choose the allocation of chances to be used in deciding who among A, B, and C wins the prize.”

The B and C players in each group are asked to make similar choices but their choices did not effect their payoffs. The players B were asked to respond to the question: “Please select the allocation that you would choose if you were the decision maker, Player A.” The players C were asked to respond to the question: “Please select the allocation that you believe is fair.” The main purpose of doing this is to give the other players a choice task such that no player can identify who is a player A based upon observing some players making a choice and others not. As a secondary consideration, though, it may be interesting to examine the beliefs of other players concerning what the decision maker will do and to elicit beliefs on what the fair allocation is.

The A players were informed that they were the only ones making a choice affecting the payoffs of everyone in their group and the B and C players were fully informed that their choices do not effect their payoffs. The A players were further informed that the B and C players will be responding to other questions, but they were not told what those other questions were. For precise information on what the subjects were told, there is a complete record of the help screens that the subjects are lead through to explain the experiment contained in appendix A.

This choice set induced by this design is a 2-dimensional simplex depicted in figure 1. The top vertex of the triangle represents the point that gives the entire probability of winning to the subject of type A. The lower left vertex gives the entire probability to subject of type B and the one at the lower right to subject of type C with the probabilities of winning being denoted by the triple (p_A, p_B, p_C) . The degree to which someone prefers fairness can then be represented in this context by the curvature in their indifference curves. If the subject of type A exhibits no sense of fairness then his indifference curves parallel linear lines, such as line 1, along which p_A is constant. If, however, he is concerned about the fairness of the lottery and moreover, if he believes that the more equal are these chances the fairer is the lottery, then the indifference curves may be convex as shown by line 2. This represents a willingness to sacrifice own probability of winning to bring about a fairer allocation procedure.

The method used to test the convexity of the preference relation is quite straight forward. A subject is presented with a chord in this simplex along which the subject own probability of winning varies and is asked to choose a point along it. If there is no curvature to their indifference curves (i.e. no preference for fairness) then the subject will select the endpoint that gives him the maximum probability to win the prize. If, however, the indifference curves are convex then the optimal lottery is represented by a point in the interior of the chord.

The subjects were asked to make a total of six choices along different chords. These are shown graphically in figure 2 and the actual numbers are shown in table 1. After each choice, the groups were reshuffled randomly, but the subjects

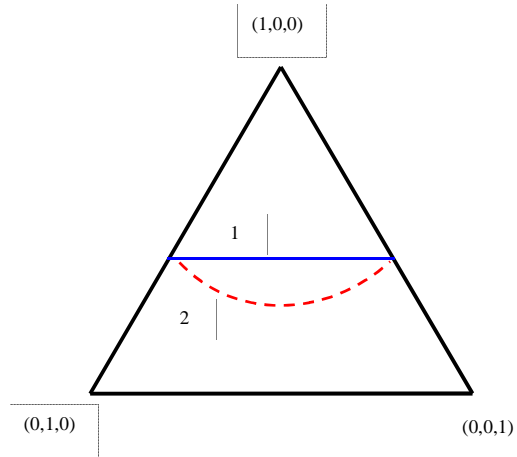


Figure 1: Characterization of indifference curves. Line 1 characterizes a person with no preference for fairness. Line 2 characterizes a person who exhibits a preference for fairness.

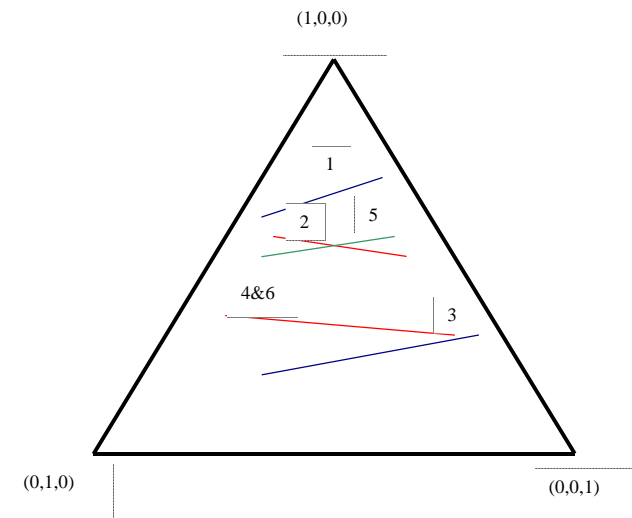


Figure 2: Graphical representation of the chords used in the experiment.

	Endpoint 1			Endpoint 2		
	A	B	C	A	B	C
1	70	5	25	60	35	5
2	55	35	10	50	10	40
3	30	5	65	20	55	25
4 & 6	35	55	10	30	10	60
5	55	10	35	50	40	10

Table 1: Chords used in experiment.

remained the same type throughout the experiment. Thus, subject who was assigned type A at the outset remains type A for all six choices. This raises the possibility that subjects engage in behavior based upon compounding the lotteries across choices. To address this concern only the first lottery generated an actual payoff. The choices made by the A types on this question were used to run actual lotteries to determine the winner of the \$15 prize. To ensure that the subjects believed the lotteries were run fairly, an extra subject was recruited in each session to run the lotteries with a pair of 10-sided dice and then observe that the proper amount of money was inserted into each envelope.

The chords chosen to be used possessed some specifically designed similarities to allow the investigation of particular issues. As seen in figure 2, chords 1 and 3 possess the same relative slope as do the chords for questions 2 and 4. The reason for this was to determine whether or not there are wealth effects in the preferences of the subjects. That is, if the player A begins with a larger probability of winning, are they willing to give up a larger amount of probability to make the allocation more fair than if they begin with a smaller probability. Questions 2 and 5 are inverses of each other, or rather the chord used for question 5 can be obtained by simply switching the probabilities for players B and C. This was intended to serve as a consistency check or test for an “equal treatment of equals” property in the choice behavior. The chords used for questions 4 and 6 were identical. The difference between these questions is that the hypothetical prize was \$45 for question 6 and only \$15 for question 4. The point being investigated here is whether or not the subjects preferences for fairness were effected by the size of the prize.

Since only one chord was used to run an actual lottery, this raises a question concerning the reliability of the answers given for the other five questions. To determine the degree to which this is important, half of the subjects were asked question 1 first and the other half question 2. By checking the degree of consistency between the choices of the two groups on the paid and unpaid question, we can characterize the degree to which subjects are willing to evidence a higher degree of a preference for fairness when the decision is hypothetical versus when it is paid.

The interface used for this experiment can be found in figure 3. The subjects were presented with an initial allocation indicating the chances out of 100 for

Figure 3: Sample screen of experimental interface.

each subject in the group to win the lottery. This initial allocation corresponded to Endpoint 1 in table 1. The subject can then use the slider bar to move along the chord between this point and the other endpoint. With each movement of the slider bar, both the chances of the subjects to win and the pie chart were updated accordingly. Their final choice can be represented by a number $0 \leq \lambda \leq 1$ such that λ is the weight used to create the convex combination of the endpoints resulting in the chosen allocation. A λ chosen to equal 1 indicates that the player A chose the point that maximized their probability to win while a λ of 0 indicates that the player A's chances to win were minimized.

One complication to this analysis was introduced due to the desire to list the probabilities of each subject winning the lottery as integers between 0 and 100. This caused the actual chords the subjects were choosing along to be jagged instead of smooth. The formulae used to generate these probabilities were :

$$\begin{aligned}
 p_a^* &= \text{Round}(\lambda * p_a^1 + (1 - \lambda) * p_a^2) \\
 p_b^* &= \text{Round}(\lambda * p_b^1 + (1 - \lambda) * p_b^2) \\
 p_c^* &= 100 - p_a^* - p_b^*
 \end{aligned}
 \tag{1}$$

p_i^* represents the probability allocated to player i by the choice of λ while p_i^j is the probability that player i would win at endpoint j . The slider bar used had 30 discrete “click” points along it that the subjects could choose. λ was then calculated by taking the “click” point along the slider bar chosen and dividing it by the number of discrete clicks that were made available.

The subjects used in these experiments were drawn from two separate subject pools. One group of subjects consisted largely of undergraduate students at The California Institute of Technology (CIT) and the other consisted of students from Pasadena City College (PCC). In total there were 135 subjects used in these experiments with 69 coming from the CIT population and 66 coming from the PCC population. Each session run contained subjects from only one group or the other.

Earnings from these sessions consisted of 1 out of every 3 subjects winning a \$15 prize in addition to their show-up fee and the other 2 out of 3 subjects receiving only their show-up fee. For CIT subjects, the show-up fee was \$5 and for PCC subjects the show-up fee was \$10². The sessions for these experiments lasted as few as 20 minutes and one up to 40. Most finished between 20 and 30 minutes.

3 Results and Analysis

The empirical analysis is organized as follows. First, we summarize the results via graphs showing the distribution of choices made by subjects of each type (the decider, the hypothetical decider, and the fair decider) for each of the six questions asked in the experiment. We then present results of statistical tests among these different distributions designed to answer a number of questions pertaining to the predictions of the theory, as well as other questions not motivated by the experimental design. We then summarize the results via an estimated ordered discrete choice model, which allows further statistical inference on the theoretical and design-specific questions.

Figures 1 to 6 contain histograms of the choices for each of the questions asked in the experiment. Each figure contains 3 histograms, one for each type of player in the experiment. Recall that since subjects were assigned to be one type of player (A, B or C) for the entire experiment, the 3 histograms in each figure represent independent samples. Before proceeding, one, unanticipated, feature of the data must be noted. It was our intent in the formulation of the instructions that all types of players make their choices from the point of view of player A. That is, player A, the decider, made his or her allocation of the chances of winning the prize knowing his or her position, and thus what chances of winning he would have for any given allocation chosen. Player B was to make his choice of allocation as if he were in the position of player A—i.e., as if he were player A. Similarly, player C was to choose what he believed to be the "fair" allocation as if he were player A. It is evident, however, that some type B and C subjects made their choices as if they were in the position of having the power to decide the allocation, but with themselves still occupying the place of player B or C and getting the chances of winning for a player B or C from any given

²The reason for the differential is simply to encourage PCC students to travel the extra distance to Caltech where the experiments were run. In addition, some subjects redeemed recruitment coupons worth \$10 that are given to PCC students when they sign-up to be on the recruitment list to be used in their first experiment.

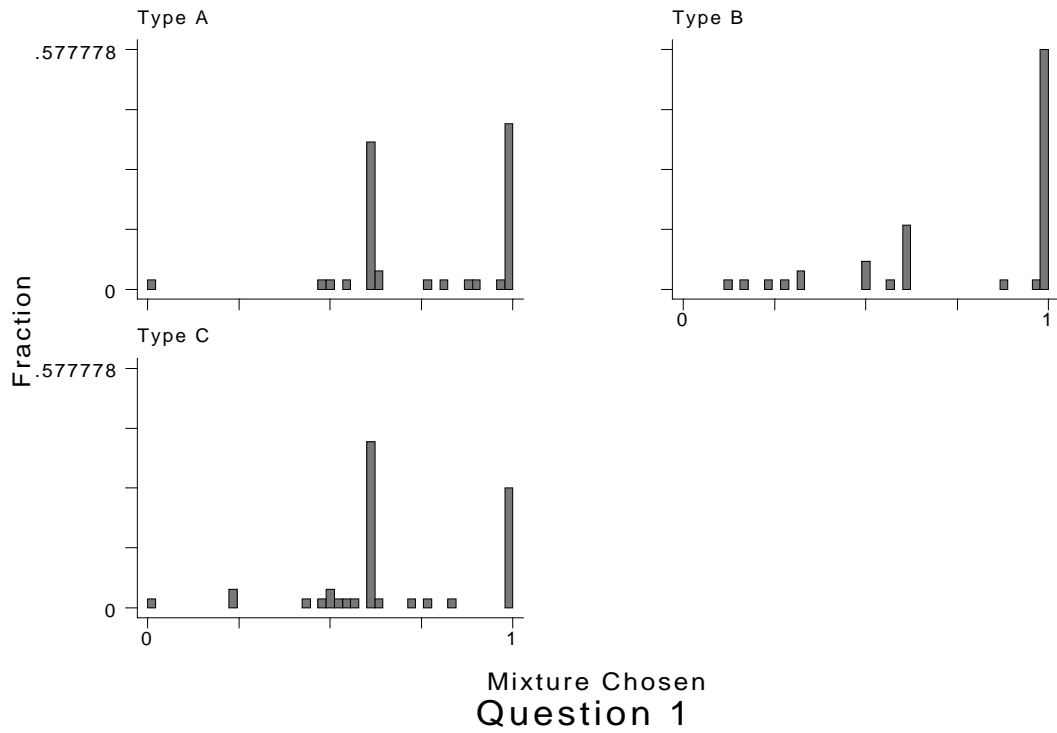


Figure 4:

allocation. In particular, on chord choices that slope down to the left (1, 3 and 5), some type B subjects (12% of them) would choose the lowest point on the chord, and on chord choices that slope down to the right (2, 4 and 6), some type C players (11% of them) would also choose the lowest point of the chord. Such choices maximize the chances of the chooser (B or C, as the case may be) winning the prize but, would minimize the available chances of player A winning the prize. One might argue that these are completely altruistic people who wish for others to share all of their own chances of winning. We take the position, however, that this is a point-of-view issue, as already outlined. We did occasionally see type A subjects choosing in a way that minimized their chances of winning, but this was rare, less than 1% of choices. In comparing the three histograms in each figure, then, it is more informative to shift the mass on the choice of 0 in the relevant histogram to 1. We have done this in the figures by shifting the mass from 0 to 1, and indicating the amount of mass that was shifted in outline at a choice of 0. With this shift, less than 1% of remaining type B and type C choices are at 0.

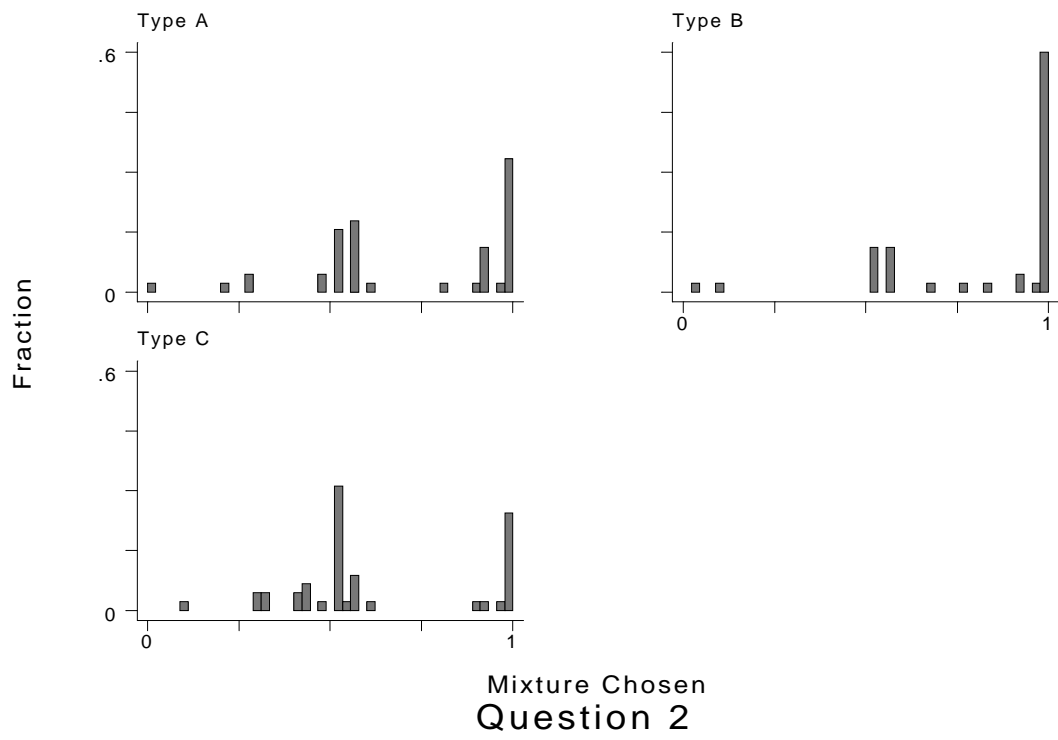


Figure 5:

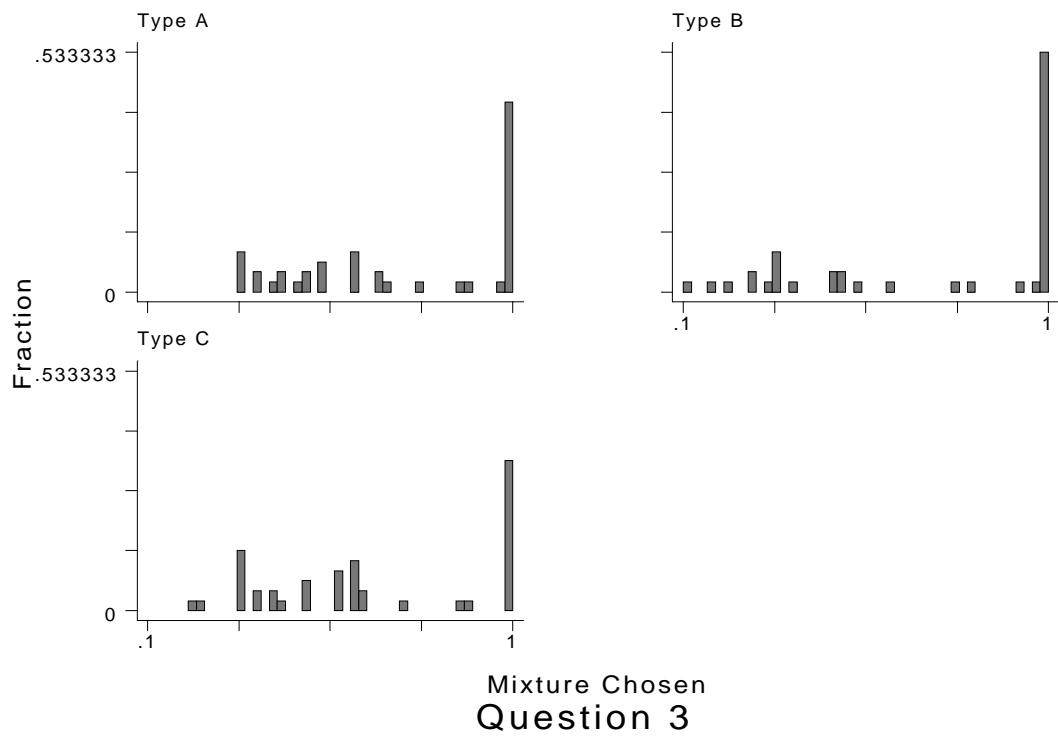


Figure 6:

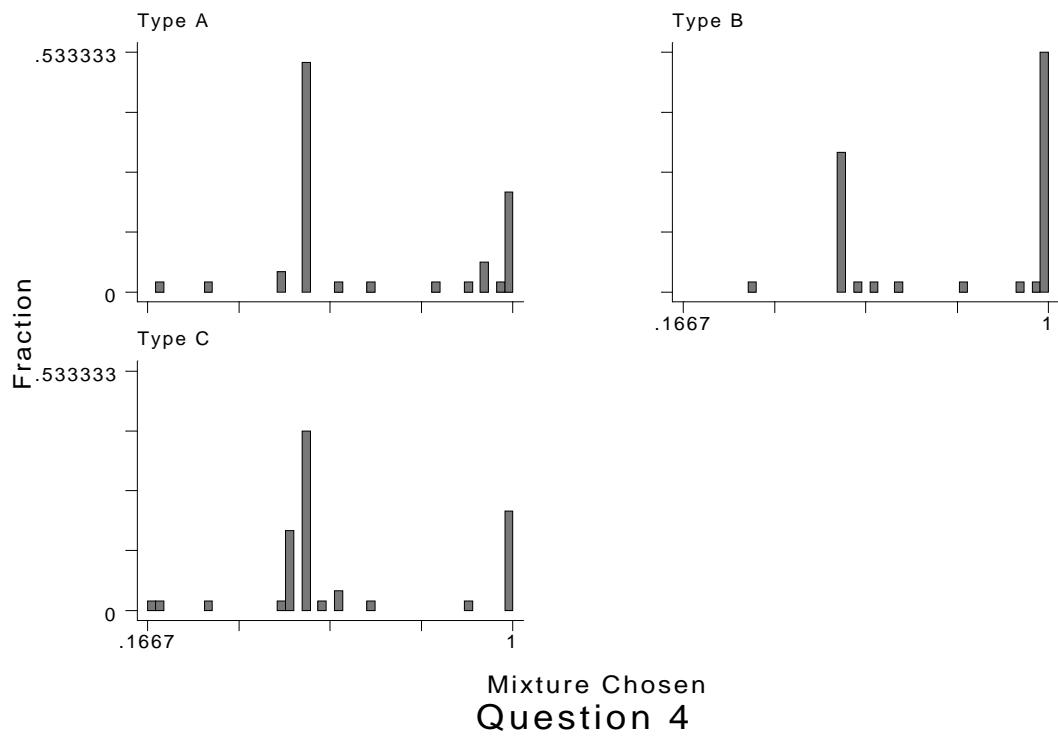


Figure 7:

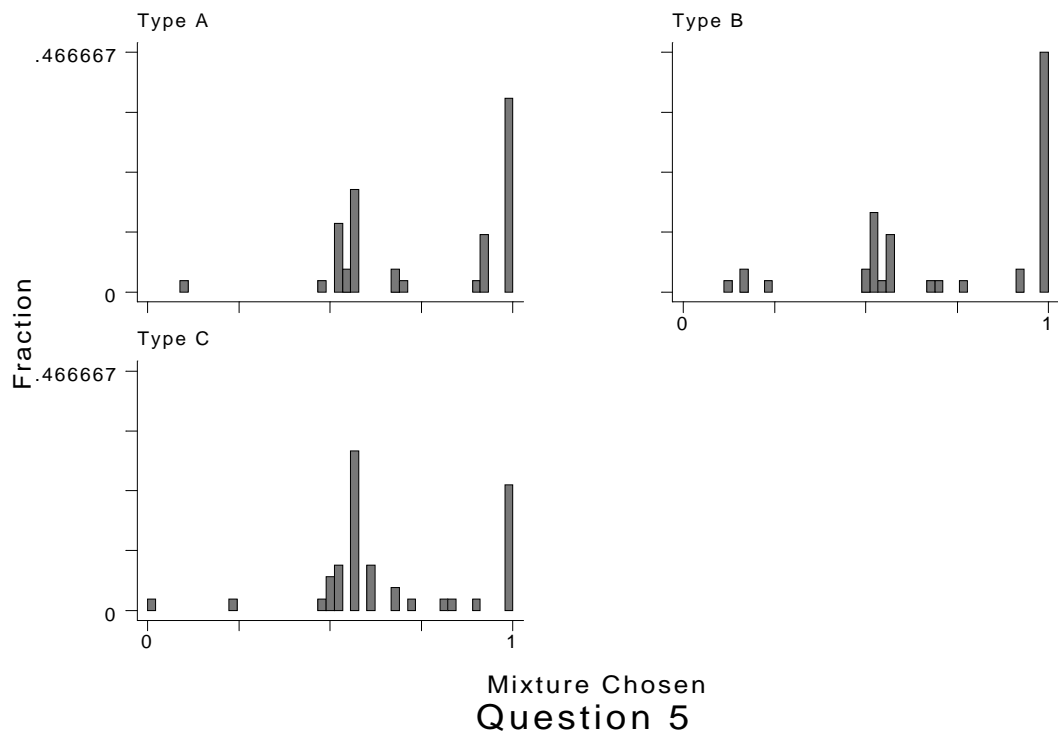


Figure 8:

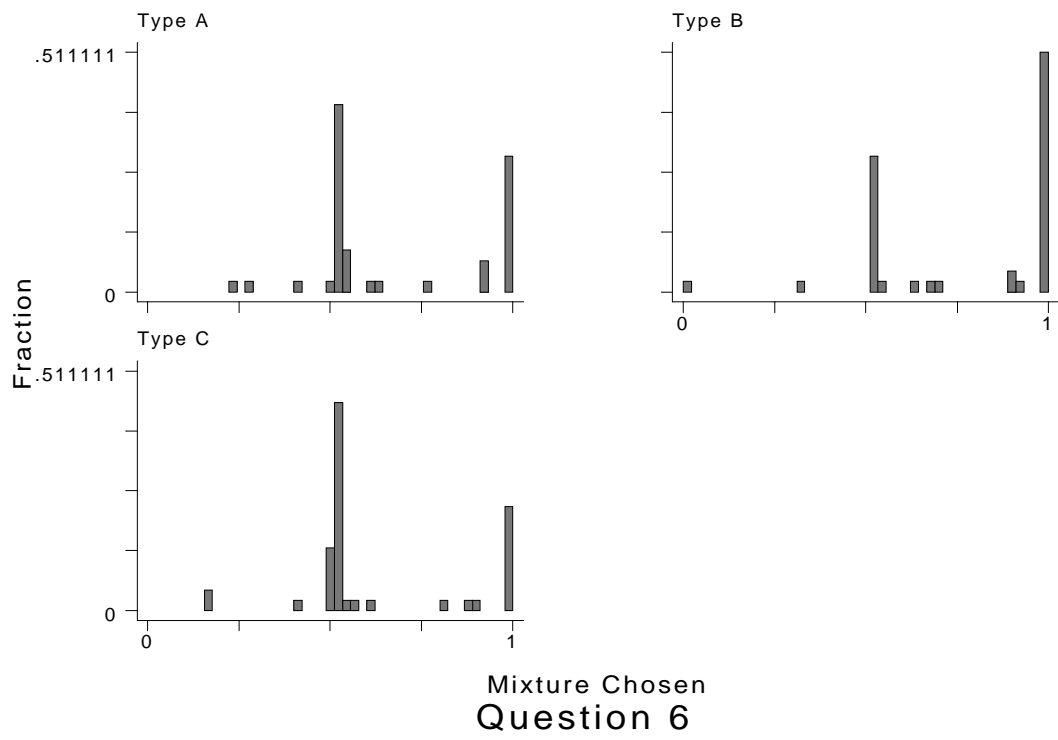


Figure 9:

With the foregoing adjustments made, the distributions are rather easily summarized. On all of the questions, the distribution of type A and type C choices appear to be remarkably similar. The typical pattern for all types of players is bimodal, with a mode in the neighborhood of the choice that would lead to an equitable distribution for players B and C, and a second mode at or near a (perfectly selfish) choice of 1. The distributions for type B players are also, typically, bimodal, but with a larger mode at 1 than is the case for type A or C players. Thus, the distribution of choices by type A players corresponds rather closely to the distribution of "fair" choices, while the distribution of type B player choices appears to be significantly more selfish. (Edi and Tim: This suggests, to me the notion that "fair" choices are something that have to do with the social context in which you are choosing—it is another layer on top of your personal, selfish preferences. In particular, in the experiment, the context for type A players is that people's earnings in the experiment are going to be determined by their choices, and that matters to them—they want to do the right thing—even though only the first question counts, they seem to feel this responsibility. The type B players, on the other hand, don't have to worry about that—their choices don't count anyway. Neither do the type C player choices, but then they were explicitly asked for the fair allocation, which is a different question. I think these thoughts don't belong here, necessarily, but somewhere in the paper—in a more general discussion of the results, perhaps, a little later).

The apparent differences and similarities in the distributions for different subjects types noted above are substantiated for many of the questions by formal statistical tests. The following table contains the results of tests among the different subject types for each of the six questions. Column 1 contains results of Kruskal-Wallis tests for the hypothesis that there is no difference between the distribution of choices for Types A, B and C. The table shows the test statistic, which is Chi-squared with 2 degrees of freedom, and the associated p-value for the test in parentheses. Columns 2 to 4 contain results of Wilcoxon two-sample tests for the hypothesis that there is no difference in the distribution of choices for each possible pair of subject types. The table shows the test statistic, which is a standard normal, with the associated p-value in parentheses. Questions 2, 4 and 6 all show strong evidence of a difference overall. For each of these questions there turns out to be a strong pairwise difference between types A and B and between types B and C. For question 2 there is also a significant difference between types A and C as well. There is weaker evidence of an overall difference in type choices for Question 1 (11% significance level), apparently driven by differences between types B and C and between types A and C but not, as in the other cases, by a difference between types A and B. To summarize, there are more significant differences between types A and B (3 at the 5 % level), and between Types B and C (4 at the 5% level) than between types A and C (1 at the 5 % level), consistent with the observation suggested by the histograms that Types A and C are more similar.

Table: Tests for Differences in Choice Distributions Among Types (within Question)

Question	All types	Types A and B	Types B and C	Types A and C
1	4.47 (.11)	-0.66 (.51)	1.93 (.05)	1.85 (.07)
2	15.82 (.00)	-2.42 (.02)	4.00 (.00)	1.99 (.05)
3	1.78 (.41)	-0.18 (.86)	1.13 (.26)	1.26 (.21)
4	13.83 (.00)	-2.79 (.01)	3.64 (.00)	1.31 (.19)
5	2.10 (.35)	0.09 (.93)	1.12 (.26)	1.44 (.15)
6	10.43 (.01)	-2.06 (.04)	3.28 (.00)	1.53 (.13)

All types: Kruskal Wallace test, Chi-squared, 2 d.o.f., (p-value).

Pairwise tests: Wilcoxon 2-sample test, standard normal z, (p-value)

Table: Characteristics of Choice Questions

Questions	Choice Leading to Equity for B and C	Implied Distribution (A, B, C)	Minimum Attainable Chances for A	Maximum Attainable Chances for A
1	.6	(.66,.17,.17)	.6	.7
2 & 5	.53	(.53,.23,.23)	.5	.55
3	.33	(.24,.38,.38)	.2	.3
4 & 6	.53	(.33,.33,.33)	.3	.35

Questions 2 and 5 slope indifferent directions, while questions 4 and 6 involved a different size of hypothetical prize.

The next table contains results of statistical questions. The table above is useful in motivating our choice of which questions to test between. The table above shows, for each question, the choice by player A that would lead to an equitable distribution for the other two players, what that distribution would be, the absolute slope of the chord (indicating the rate at which one can trade off one's own chances of winning with the chances for the other two players, and the minimum and maximum attainable chances of winning for player A. Note that questions 2 and 5 are identical with respect to the information in the table. They slope in difference directions in the simplex, and thus it is natural to check that there is no difference in the responses to them (a difference would suggest that players, systematically, wish to treat player B different from player C). Questions 4 and 6 are also identical in the table. They had different size monetary prizes associated with them (larger for 6), and thus we are checking whether the scale of payoffs affects the distribution of choices. Questions 1 and 3 each involve the same rate of tradeoff between one's own chances of winning and the chances for other subjects to win (.1), but in Question 1 the Type A subject has a larger "endowment". Thus, we are testing for whether there is an endowment effect in choice, in the sense that one is more or less likely to treat other subjects equitably depending upon one's own range of possible allocation of the chances of winning. The indicate essentially no difference in any of the comparisons, either with all types pooled or for each type separately. The one exception is that type C subjects are significantly different (at the 6% level) in Questions 2 and 5.

Table: Test for Differences in Choice Distributions Between Selected Questions

Questions	All types	Type A	Type B	Type C
2 vs. 5	0.11 (.74)	0.49 (.48)	2.18 (.14)	3.64 (.06)
4 vs. 6	0.01 (.93)	0.10 (.75)	0.06 (.80)	0.03 (.85)
1 vs. 3	1.39 (.24)	0.54 (.46)	0.47 (.50)	0.63 (.43)

All tests: Wilcoxon 2-sample test, standard normal z, (p-value)

Finally, we present results of ordered qualitative choice regression models (one for each question) which help us to efficiently summarize the within-question differences noted above, as well as some other factors not yet discussed. Since choices are ordered along each chord in the simplex, we can treat each mixture choice with positive mass in the distribution of choices as a discrete choice, and form a likelihood function based on the interpretation of subject choices as reflecting some underlying latent variable summarizing subjects' preferences over possible allocations of chances of winning a monetary prize. In particular, if we can think of choice mixtures ranging from "1" (perfect selfishness) to "e" (the choice that implies equity for the other players) to "0" (perfect selflessness) as being generated by some sort of monotonic propensity to be more or less selfish in allocating the chances of winning amongst the three subjects in any given choice situation in our experiment, then the ordered model is sensible. Since, in fact, the relevant range of choices really only ranges from "1" to "e", with very few choice allocations giving player A less than what he would get in the allocation that is equitable for B and C, we seem to be on fairly safe ground with this interpretation. (it is a bit tricky for choices between "e" and "0"—one is giving up some payoff and, at the same time, making the allocation less equitable for the other two). (I can, and should, dress this up with a little detail about the nature of the likelihood, etc.).

The next table summarizes results for ordered logit regressions for each of the six questions. The dependent variable is CHOICE, the mixture chosen, while the regressors are MALE (1 if male, 0 if female), CIT (1 if Caltech student, 0 if PCC student), TYPEB and TYPEC (type B and C dummies), and PAY (1 if a paid question, 0 otherwise). Since the regressors are all discrete 0-1 variables, individual tests of significance for coefficients on these regressors can be interpreted as tests for whether the distribution of choices by one group of subjects is significantly different (in terms of a shift one way or the other in the distribution over the 0-1 interval) is different than that of the rest of the subjects. Since coefficient estimates themselves, aside from their signs, are not too informative in this type of regression model, we will summarize the implied probabilities of choice for the different groups in a subsequent table. Here, we report either a '+', a '0' or a '-', indicating whether the regressor in question was 'positively significant', 'insignificant' or 'negatively significant' at the 1, 5 or 10% level, as indicated. The table also reports N (number of independent observations), LL (log-likelihood for the estimated model), and PRsq (pseudo-R-squared) for each regression.

Table: Ordered Logit Results for Choice Mixture Chosen

Independent Variable	Q1	Q2	Q3	Q4	Q5	Q6	Q4 & Q6
MALE	+ (.01)	0	0	0	+ (.01)	0	0
CIT	+ (.05)	0	+ (.05)	0	+ (.05)	0	0
TYPEB	- (.01)	+ (.01)	- (.01)	+ (.01)	- (.01)	+ (.05)	+ (.01)
TYPEC	- (.10)	- (.01)	0	- (.01)	0	- (.01)	- (.01)
PAY	0	+ (.10)	NA	NA	NA	NA	NA
N	135	135	135	135	135	135	270
LL	-264.19	-287.1	-340.27	-259.94	-310.51	-245.94	-524.39
PRsq	.05	.11	.03	.09	.05	.09	.08

4 Interpretation

To interpret the results of the experiments consider the separately-additive representation in Karni and Safra (2000). Assume that,

(A.1) For all subjects, the fairness relation is symmetric with indifference curves that are concentric circles and a utility representation that is proportional to the distance from the center of the simplex. Formally,

$$\sigma_i(p) = -\theta_i \sqrt{\sum_{j \in \{A, B, C\}} \left(p_j - \frac{1}{3}\right)^2},$$

where i the name of the subject, p_j , $j = A, B, C$ is the probability type j winning, and $\theta_i > 0$ is an individual parameter.

The symmetry assumption seems justifies on the ground of “the principle of insufficient reason.” In each experiment the participating subjects came from that same background and, as far as we can tell, have no other information that would allow them to apply some other *ethical* consideration to differentiate among their claims to the prize. Put differently, it seems reasonable to assume that the claims of all members in each group to the prize are of equal merit and should be treated symmetrically. Thus, from behind a ‘veil of ignorance’ each lottery (p_A, p_B, p_C) is equally fair, or equally unfair, as any of its permutations. The assumption regarding the shape of the indifference curves and the specific functional form are chosen for convenience. The robustness issue this raises is addressed below.

(A.2) For all subjects the self-interest component depends solely on the own probability of winning. Formally,

$$h_i(\kappa \cdot p) = \kappa_i p_i, i = A, B, C.$$

The presumption underlying (A.2) is that there are no externalities in the enjoyment of the prize.

A chord in our experiment with endpoints p and q is a set $T(p, q) = \{\lambda p + (1 - \lambda)q \mid \lambda \in [0, 1]\}$. Assume throughout that $p_A > q_A$ and let $\bar{\alpha}_A(p, q) =$

$\arg \max_{\lambda \in T(p,q)} -\theta_i \left[\sum_{j \in \{A,B,C\}} (\lambda p_j + (1-\lambda) q_j - \frac{1}{3})^2 \right]^{1/2}$. Then, since

$$\sigma_i(\lambda p + (1-\lambda) q) + h_i(\kappa \cdot (\lambda p + (1-\lambda) q)) = \kappa_i \cdot (\lambda p_i + (1-\lambda) q_i) - \theta_i \left[\sum_{j \in \{A,B,C\}} \left(\lambda p_j + (1-\lambda) q_j - \frac{1}{3} \right)^2 \right]^{1/2}$$

is concave in λ , we have:

Proposition: Let $\alpha_A^*(p, q)$ be the optimal choice of type A of an allocation procedure in $T(p, q)$. If assumptions (A.1) and (A.2) hold then $\alpha_A^*(p, q)$ is unique and:

$$(a) \theta_A / \kappa_A \leq \frac{(p_A - q_A) \left[\sum_{j \in \{A,B,C\}} (p_j - \frac{1}{3})^2 \right]^{1/2}}{\sum_{j \in \{A,B,C\}} (p_j - \frac{1}{3})(p_j - q_j)} \text{ implies } \alpha_A^*(p, q) = 1,$$

$$(b) \theta_A / \kappa_A > \frac{(p_A - q_A) \left[\sum_{j \in \{A,B,C\}} (p_j - \frac{1}{3})^2 \right]^{1/2}}{\sum_{j \in \{A,B,C\}} (p_j - \frac{1}{3})(p_j - q_j)} \text{ implies } \alpha_A^*(p, q) \in (\bar{\alpha}_A(p, q), 1).$$

The expression θ_A / κ_A is a measure of the intensity of the sense of fairness. It may be assumed to be distributed over the half-open interval $[0, \infty)$. This induces a distribution on the level of $\alpha_A^*(p, q)$. Note, however, that because the actual range of $\alpha_A^*(p, q)$ is $[\bar{\alpha}_A(p, q), 1]$ and the effect of censoring, the induced distribution tends towards bimodality with concentrations at 1 and around $\bar{\alpha}_A(p, q)$. Clearly, given θ_A and κ_A , the induced distribution of $\alpha_A^*(p, q)$ and the maximal fairness position $\bar{\alpha}_A(p, q)$ depend on the particular chord (see below.)

Denote the critical value $\frac{(p_A - q_A) \left[\sum_{j \in \{A,B,C\}} (p_j - \frac{1}{3})^2 \right]^{1/2}}{\sum_{j \in \{A,B,C\}} (p_j - \frac{1}{3})(p_j - q_j)}$ by ξ then, applying the proposition to the questions in the experiments we get:

	p	q	ξ	$\theta_A / \kappa_A \leq \xi$	$\theta_A / \kappa_A > \xi$
Q1	(0.7, 0.05, 0.25)	(0.60, 0.35, 0.05)	0.45	$\alpha_A^*(p, q) = 1$	$\alpha_A^*(p, q) \in (0.25, 1)$
Q2	(0.55, 0.35, 0.10)	(0.50, 0.10, 0.40)	0.19	$\alpha_A^*(p, q) = 1$	$\alpha_A^*(p, q) \in (0.48, 1)$
Q3	(0.30, 0.05, 0.65)	(0.20, 0.55, 0.25)	0.16	$\alpha_A^*(p, q) = 1$	$\alpha_A^*(p, q) \in (0.37, 1)$
Q4, 6	(0.35, 0.55, 0.10)	(0.30, 0.10, 0.60)	0.07	$\alpha_A^*(p, q) = 1$	$\alpha_A^*(p, q) \in (0.55, 1)$
Q5	(0.55, 0.10, 0.35)	(0.50, 0.40, 0.10)	0.19	$\alpha_A^*(p, q) = 1$	$\alpha_A^*(p, q) \in (0.48, 1)$

Observe that the critical values ξ and $\bar{\alpha}_A(p, q)$ are the same in Q2 and Q5 reflecting the fact that questions 2 and 5 are the inverses of each other (i.e., the probabilities players of type B and C are switched.) Notice also that the range of $\alpha_A^*(p, q)$ is larger for Q1 than it is for Q3. (The same observation and latter observations as well apply to Q2 and Q4.) Recall that chords 1 and 3 have the same slope. This is the "wealth effect" mentioned in section 2. Specifically, the probability of winning of type A is larger in chord 1 than in chord 3 for every $\lambda \in [0, 1]$. We say that Q1 chord-dominates Q3. The range is monotonic increasing with respect to the chord-dominance relation. This is a reflection of a more fundamental principle.

The Principle of Increasing Benevolence: *The willingness of individuals to sacrifice their own probability of winning to bring about fairer allocation procedure tends to increase as a function of this probability.*

The logic underlying this principle is that a subject's own probability contributes to the overall fairness of the lottery. Thus, when this probability is high, giving up some contributes, in itself, to the fairness of the overall outcome while if it is low giving up the own probability detracts from the fairness of the lottery. In the former case, the sacrifice involved in giving up the own probability of winning is mitigated by the contribution it makes to increasing fairness (this is in addition to the gain in fairness attained by reallocation of the probabilities between the other two subjects). In the latter case the sacrifice involved in giving up the own probability of winning is exacerbated by the diminution of fairness it causes. In general, the higher is the own probability the larger is the mitigating and the smaller is the exacerbated force, which together give rise to the principle of increasing benevolence.

5 Conclusions and Further Research

Appendix A: Experiment Instructions

The experiment is begun by recruiting a volunteer from the subjects to serve as a monitor for the experiment. Once a subject has volunteered, the experimenter reads the following message to the rest of the participants.

Thank you for participating in today's experiment. In a few moments, I will ask you to turn to the computer screen in front of you and log-in to the system. Once you have done so, you will be lead through a series of help screens detailing the choice task you will be asked to engage in as well as the interface you will be using. Please practice with the interface so that you understand well how it works.

You will be asked in this experiment to make 6 separate decisions concerning different lottery allocations. The first of these lotteries will be run according to the rules that will be detailed in the instructions. The next 5 will be hypothetical choices and will not be used to determine your earnings in this session. The lotteries will be run by the volunteer monitor using these two ten-sided dice to ensure that they are run fairly. The monitor will also observe and ensure that the proper amount of money is placed in each envelope.

If there are any questions during the experiment, please raise your hand and I will come to assist you. Are there any questions at this point?

Please log-in to the system and begin.

Once the subjects login to the computer system, they are presented with a series of help screens leading them through the experiment. The first is another introduction screen.

You have volunteered to participate in an economic experiment on decision making. If you have any questions during the experiment please raise your hand and ask the proctor.

In this experiment, you will be asked to make a series of 6 choices. For each of these choices, you will be asked to make a decision concerning the chances to win a prize for a group consisting of yourself as well as two other participants in the room. One of you will be designated as player A, one player B and one player C. With each new choice, the group you are in will change but your player type will remain the same throughout the experiment. Player A will be presented with a choice of how to allocate the chances of winning a prize between the three group members. Players B and C will both be making choices that will have no impact on who wins the prize. At the conclusion of the experiment, the probabilities chosen by the player A's for the first choice will be used to award the prize associated with that decision. Once the prizes have been assigned, everyone will be paid their show-up fee and winnings in cash.

If at any point you have a question, please raise your hand and a proctor will help you. Please refrain from talking during the experiment and from looking at the screens of other participants.

After they press a button to continue, a version of the interface, figure ??, is brought up with a box along the left hand side that contains text explaining how the experiment will work. The first block of text orients them to some of the content of the interface screen.

This is how the game screen will appear. When the game begins make sure to look up in the upper left corner of the screen to find your player id and player type. Please remember your player id to assist in paying you after the experiment. Your player type, remember, determines whether you are a decision maker choosing the real probabilities that will be used to award the prize or making a hypothetical choice. The type of choice you are making will be indicated just below your player type.

In each round, you will be presented with an initial allocation of chances to win the prize consisting of the initial chances out of 100 to win the prize for each player. In addition, directly above the graph you will find how much the prize would be worth to the winner.

The next screen explains how players will make choices.

At the beginning of the experiment, you will be randomly assigned a player type that will remain constant throughout the experiment. In each period, however, the group you are in will change.

If you are designated as a player of type A then you will be choosing how to allocate chances to win the prize in each period. At the beginning of each turn, you will be presented with an initial distribution. In this example player A has been allocated 50 chances to win, player B 5 and Player C 45. If you are player A, you will be able to use the slider bar at the bottom of the screen to change these probabilities.

The third screen explain in general terms what players B and C will be doing and has the players practice moving the slider bar. As the text indicates, players could not advance past this screen without moving the slider bar.

If you are designated as a player B or C you will be asked to make a hypothetical choice. You will also do this by moving the slider bar. Your choices will have no impact on anyone's payoffs.

Try moving the slider bar around now to see how it works. Notice that the graph on the right shows a pie chart representation of the possibility of each player winning. As you move the slider bar, the graph updates automatically as do the text boxes indicating the chances for each group member to win.

Note: You must try moving the bar to continue.

The final screen explains how players submit their choices.

Once you have made your choice for the allocation, click on the button labeled “Accept.” You will be asked to confirm your choice before it is sent on to the server. When everyone’s choices have been submitted, the groups will be reshuffled and you will move to the next choice.

Try making a selection with the slider bar clicking on the “Accept” button now to see how it works. Clicking on the “Continue” button now will begin the game.

Once players advance past this screen, they enter into the actual experiment interface. Before all of the other subjects are finished with the instructions, all of the controls are greyed out and inactive and there is a dialog box on the screen asking the subjects to wait patiently. Once all subjects have finished the instructions, the dialog boxes disappear, the controls are enabled and the experiment begins.

REFERENCES

References

- [1] Berg, Joyce, John Dickhaut and Kevin McCabe (1995) “Trust, Reciprocity, and Social History,” *Games and Economic Behavior*, 10, 122-142.
- [2] Camerer, Colin F. (1997) “Progress in Behavioral Game Theory,” *J. of Econon. Perspectives*, 11, 167-188.
- [3] Elster, Jon (1998) “Emotions and Economic Theory,” *J. of Econ. Literature*, 36; 47-74.
- [4] Hume, David (1740) *Treatise on Human Nature*. London, J. M. Dent and Sons, 1939.
- [5] Karni, Edi and Zvi Safra (2000) “Individual Sense of Justice: A Utility Representation,”
- [6] ————— (2000a) “Intensity of the Sense of Fairness: Measurement and Behavioral Characterization,”
- [7] Loewenstein, George (2000) “Emotions in Economic Theory and Economic Behavior,” *The American Economic Review; Papers and Proceedings*. 426-432.
- [8] Rawls, John (1963) “The Sense of Justice,” *Philosophical Review*, 72, 281-305.
- [9] ————— (1971) *A Theory of Justice*. Cambridge, Harvard University Press.

- [10] Romer, Paul M. (2000) "Thinking and Feeling," *The American Economic Review; Papers and Proceedings*. 439-443.
- [11] Smith, Adam (1759) *The Theory of Moral Sentiments*. New Edition, D. D. Raphael and A. L. Macfie (eds.) Oxford, Oxford University Press, 1976.