



NUFFIELD CENTRE FOR EXPERIMENTAL SOCIAL SCIENCES  
DISCUSSION PAPER SERIES

2010 - 005

The Consistency of Fairness Rules: An Experimental Study

Paloma Ubeda (University of Valencia)

November 2010

# The Consistency of Fairness Rules: An Experimental Study\*

Paloma Ubeda<sup>†</sup>

LINEEX, University of Valencia

November 22, 2010

## Abstract

In the last two decades, experimental papers on distributive justice have abounded. Two main results have been replicated. Firstly, there is a multiplicity of fairness rules. Secondly, fairness decisions differ depending on the context. This paper studies individual consistency in the use of fairness rules, as well as the structural factors that lead people to be inconsistent. We use a within-subject design, which allows us to compare individual behavior when the context changes. In line with the literature, we find a multiplicity of fairness rules. However, when we control for consistency, the set of fairness rules is considerably smaller. Only selfishness and strict egalitarianism seem to survive the stricter requirement of consistency. We observe that this result is mainly explained by a self-serving bias. Participants select the rule that is individually optimal in each situation.

Keywords: Distributive Justice, Fairness, Laboratory Experiments, Self-serving bias, Consistency. *JEL classification*: D63, C91.

---

\*I am indebted to Abigail Barr, Dirk Engelmann and Luis Miller for their valuable comments. I also thank Michelle Belot and Enrique Fatas for very useful remarks. Financial support from Conselleria d'Educació de la Generalitat Valenciana (BFPI/2007/260) is gratefully acknowledged.

<sup>†</sup>Contact Information: Paloma Ubeda; LINEEX, ERI-CES, University of Valencia, Campus Tarongers 46022 Valencia, Spain. E-mail: paloma.ubeda@uv.es, Phone:+34-961-625-409, Fax:+34-961-625-415

# 1 Introduction

The concept of fairness is at the very core of many developments in Behavioral Economics and the Psychology of Judgment and Decision-Making. It was first studied under the analytical framework of equity theory (Homans 1961; Adams 1965; Selten 1978; Guth 1988). Subsequently, several one-size-fits-all theories were proposed (Rabin 1993; Fehr and Schmidt 1999; Bolton and Ockenfels 2000; Charness and Rabin 2002). More recently, several studies have claimed that individual fairness concerns can better be captured by a multiplicity of fairness rules (Frohlich et al. 2004; Cappelen et al. 2007).

The objectives of this paper are twofold. Firstly, the application of different fairness rules is experimentally studied in the context of a two-person distribution game preceded by a real-effort task. Secondly and more importantly, individual consistency in the application of fairness rules is scrutinized.

The plurality of fairness ideals hypothesis assumes that different individuals rely on different normative standards when they decide in a distribution problem. In a real effort experiment, Frohlich et al (2004) find that, besides selfishness, there are two other main behavioral patterns that classify individual decisions. Both, *just-desert* and *egalitarian* rules explain a higher proportion of decisions than selfish behavior. Cappelen et al (2007) report similar results. They propose a pluralism of fairness ideals, where the concept of *just desert* is decomposed in the *liberal egalitarian ideal* (*accountability principle* in Konow 2000) and the *libertarian ideal*. The different rules are reviewed in more detail in Section 2.

The finding of several well-defined behavioral patterns in experimental distributions problems raises two important questions. First, what proportion of individuals follows each of the proposed rules? And second, do individuals behave consistently with one single ideal or do they change their ideal depending on the context? The answer to these two questions is important because it would be difficult to accept the existence of a rule that is hardly ever chosen, or that is abandoned when the context slightly changes. There are reasons, however, to expect individuals to deviate from a previously applied rule.

Acting in accordance with a particular behavioral rule does not necessarily imply that a subject always adheres to the fairness ideal this behavioral regularity is based on. For instance, it has been shown in both psychological and economic experiments that individual decisions in fairness contexts display a self-serving bias (Messik and Sentis 1979; Babcock et al. 1995; Konow 2000). If the self-serving bias hypothesis is true, subjects will find strong reasons for switching between different fairness rules as their decision-making context changes. To put it differently, is the self-serving bias a good predictor of individual inconsistencies in a repeated distribution problem?

To address this question, we use a within-subjects design. The experiment consists of twenty periods. Each period contains two sub-periods, the real effort task in the former and the distribution decision task in the latter. The total amount of money to be distributed depends on both, discretionary (effort) and exogenous (random shock) variables. This design allows us to study individual (in)consistent behavior when the context changes, and to answer our main question concerning the role of self-serving bias.

Participants' decisions in the first period of our experiment are very similar to results reported in previous studies. We find a multiplicity of fairness rules. However, when we explore individual consistency across periods, we find that selfish and strict egalitarian participants are strikingly consistent. In contrast, no participant consistently follows the liberal egalitarian nor the libertarian ideals suggested in the previous literature. Moreover, we are reasonably confident to conclude that inconsistencies in our experiment are driven by a self-serving bias. Our statistical analysis confirms that inconsistent people are highly and significantly more likely to choose the rule that benefits them the most.

The rest of the paper proceeds as follows. Section 2.2 reviews previous studies on the multiplicity of fairness rules and self-serving bias. Section 2.3 describes the experimental design and the fairness rules we use in this work. Section 2.4 derives some predictions. In section 2.5 the results are presented. Finally, section 2.6 discusses the results and concludes.

## 2 Multiplicity of Fairness Rules and Self-serving Bias

Although the experimental study of equity goes back to Homans (1961) and Adams (1965), only recently has there been an explosion of economic papers that explore how individuals distribute the benefits of joint production (see, among others, Konow 2000 and Cherry et al. 2002). One of the main findings in this literature is that participants' behavior is guided by a multiplicity of normative standards.

Frohlich et al. (2004) extends the inequity aversion model of Fehr and Schmidt (1999) incorporating a social context. Participants perform a real-effort task and subsequently distribute the joint production. They find that more than half of the individuals share the joint earnings according to the *just desert* (equity) rule. They also find a non-negligible number of egalitarian participants (27.5%), and only 10% of selfish individuals. Frohlich et al. (2004) stress the role of entitlement in the interpretation of their results. Furthermore, they find evidence of a self-serving bias.

In Cappelen et al. (2007), participants have to decide how much they want to invest in a group account, keeping the part of the endowment they do not invest. Afterwards, they have to distribute the total income produced by the group. Players only differ in the rates of return assigned to their investment. The authors find a pluralism of fairness ideals. They also find that participants' decisions are not biased, in the sense that they do not endorse the fairness ideal that most benefit them. This result is not in line with previous literature, which predominantly finds different types of self-serving bias.

Self-serving bias seems to be an important phenomenon when individuals can rely on several rules when they have to make fairness decisions. In a hypothetical experiment, Messick and Sentis (1979) find an egocentric bias in fairness judgments. They show that non-strict equality subjects consider it fair to be paid more than the other participant, when they both work the same amount of hours. Similarly, they think that being paid equally is fair when they work fewer hours than their partner.

In economic experiments, Konow (2000) models and tests cognitive dissonance in a standard distribution problem. He uses a joint production task where discretionary as well as exogenous variables are present. The conflict between self-interest and fairness leads subjects to self-deception, and hence to behave in a self serving bias fashion.

Babcock et al. (1995)<sup>1</sup> ask subjects to judge real trial cases. Participants play either the role of plaintiff or defendant, and they have to reach a settlement and to provide arguments supporting their decision. Introducing a richer context in the experiment allows subjects to focus on the arguments that favor themselves. The authors show that settlements are more difficult because the arguments of the different players are biased towards their own interests.

Previous studies on the self-serving bias phenomenon have two main limitations. First, some of them use non-incentivized methods to elicit fairness ideals. This is the case of survey and hypothetical experiments (Konow 1996; Messick and Sentis 1979). But, as Konow (2005: 359) puts it, ‘Although studies with and without real stakes lead to similar conclusions, some results are affected by the presence of real stakes, such as the average level of unfairness.’

The second limitation of the self-serving bias literature is related to the type of empirical analysis typically used in this literature. Studies on social preferences have mostly tested their hypothesis with an aggregate level data technique. If we are interested, however, in the study of behavioral (in)consistencies or in the self-serving bias problem, this type of analysis could lead to a wrong interpretation of the data. To understand individual behavioral inconsistencies, a within subject analysis is essential. Andreoni and Miller (2002), Fisman et al. (2007) and Blanco et al. (2010) among others, have used individual-level analyses to test the consistency of models of other-regarding preferences.

Using a design that shares some features with the one presented in this paper, Brosig et al. (2007) analyze the consistency and stability of individual behavior across games. They use a within-subject design where participants face the same group of games repeatedly. They find that selfish subjects are consistent across

---

<sup>1</sup>See also Babcock et al (1993, 1996, 1997).

games. They also show that only selfish decisions are stable over time.

We build on this existing literature by performing a within-subjects analysis in the context of a real-effort experiment. By the manipulation of exogenous and endogenous variables, we study whether participants rely on different normative standards and whether these normative standards are selected in a self-serving manner.

## 3 Experimental Design and Procedures

### 3.1 Experimental Design

The experiment consists of 20 one-shot pure distribution games with production. In each one-shot, participants play two phases. At the beginning of the first phase random pairs are formed. The same two subjects take part in the first and the second phase. In the first phase participants perform a real effort task. In the second phase, subjects face a pure distribution problem.

#### *Production phase*

The goal of the production phase is to induce a feeling of entitlement by using a real-effort task. The real-effort task consists of a series of puzzles in which the letters of a word have been scrambled. Subjects have to unscramble as many puzzles as they can out of ten.<sup>2</sup> Individuals are endowed with initial endowments corresponding to their effort in this phase. For each word they correctly unscramble they get 4 tokens. Individual and group endowments are common knowledge.

After the production phase, a random shock is introduced. Each individual outcome has an independent 50% probability to be affected. The shock halves participants' endowment, thus the group endowment is also reduced. In this experiment, both discretionary and non-discretionary variables can potentially differ among subjects.

---

<sup>2</sup>The real effort task is taken from Carpenter et al (2010).

### *Distribution phase*

In the second phase, participants face a pure distribution problem. Both members of the group have to decide anonymously how to distribute, between them, the joint benefits after the shock.<sup>3</sup> They do not receive any feedback until the end of the experiment, preventing them from forming expectations about others' behavior, so trying to rule out reciprocity.

In every one-shot, participants are randomly paired with another participant in the room. A random stranger mechanism is used. At the end of the experiment, the computer randomly chooses one period and one decision for each pair to be paid. The exchange rate is 3 tokens=1 pound.

## **3.2 Fairness Rules**

After the experiment, participants' decisions are classified according to several fairness rules. In this paper, we use the definition of fairness ideals proposed by Cappelen et al (2007) to evaluate subjects' decisions. The three fairness ideals are then defined as:

Strict egalitarian:

$$m^{SE}(a, q) = \frac{X(a, q)}{2} \quad (1)$$

Liberal egalitarianism:

$$m^{LE}(a, q) = \frac{q_1}{(q_1 + q_2)} X(a, q) \quad (2)$$

Libertarian:<sup>4</sup>

---

<sup>3</sup>If a shock has occurred at all.

<sup>4</sup>We use the definition of libertarian as output equity as proposed by Cappelen et al. (2007).



$$m^L(a, q) = a_1 q_1 \tag{3}$$

$m^k \in \{m^{SE}, m^L, m^{LE}\}$  represents the different fairness ideals, and it corresponds to the amount a particular subject keeps for herself.

$X(a, q)$  represents the final outcome to be distributed, and it comes from the amount produced by both subjects  $i$ , where  $i \in \{1, 2\}$ . In this case the outcome could be affected by two variables: discretionary ( $q$ ) and exogenous ( $a$ ) variables.

$$X(a, q) = x_1(a_1, q_1) + x_2(a_2, q_2) \tag{4}$$

$$x_i = a_i q_i \tag{5}$$

In our experiment,  $a$  is determined by an external shock, which subjects have no control over. In contrast,  $q$  corresponds to subjects' effort in the production phase, which of course can be discretionally affected by subjects.  $a$  and  $q$  take the following values:

$$a \in \{0.5, 1\} \tag{6}$$

$$q \in [0, 40] \tag{7}$$

Variables,  $q$  and  $a$ , change across the 20 one-shot games producing heterogeneity in the final outcomes. This heterogeneity allows us to study the (in)consistency of subjects' decisions.

The rationale for modeling the non-discretionary component of the model as an

---

This is not of course the only definition of libertarian that one can find in the social science and philosophical literatures.

external shock is the real-world observation that production is often affected by exogenous shocks, e.g. meteorological conditions.

Imagine, for instance, two tomato-field workers working for the same company. They farm tomatoes in two smallholdings situated in different fields. They are equally productive, and hence earn the same wage. But, sometimes at the beginning of the harvest, strong storms occur in one of the fields, causing the crop to fail. The boss receives less money for this incident, and also has less money to pay his employees. On the one hand, both peasants have put the same effort into their work, but on the other hand only one production has been successful. In view of this situation the principal has to decide how to distribute the money produced to pay the wages of the two workers. The question is: *which is the fair distribution for each worker?* The same question arises when the exogenous shock does not happen, but the productions of the two peasants differ because one of them cultivates the tomatoes better and produces a higher amount as a consequence. The two situations described, and any combination of the two, creates a dilemma: *how should the benefits be distributed?*

### 3.3 Experimental Procedures

The experiment took place at CESS at Nuffield College, University of Oxford. Subjects were 60 students<sup>5</sup> from 26 different fields at the University of Oxford<sup>6</sup> who were recruited using ORSEE (Greiner 2004). Also, we balanced the data to have half of them male and half female. The experiment was run using z-tree (Fischbacher, 2007) and lasted for about 90 minutes. Subjects were randomly assigned to one computer terminal. We gave a copy of the instructions to subjects and then a research assistant read the instructions aloud. To assure that everyone understood the instructions we asked them to complete a control questionnaire. After the experiment, participants filled in a post-experimental questionnaire in which we asked them some socio-demographic questions and questions related to the rules and strategies they

---

<sup>5</sup>Subjects were originally from English speaker countries (England, Ireland, Canada, and US) only one participant was Swiss.

<sup>6</sup>We want to have a heterogeneity of students in order to avoid to have a large mass of ‘selfish economists’ (Frank et al, 1993), and hence have a better representation of the society.

had followed. Subjects earned 11.5 GBP on average, ranged from a minimum of 6 GBP to a maximum of 26 GBP, including the show-up fee of 4 GBP and 2 GBP for the post experimental questionnaire (announced after the experiment was over).

## 4 Predictions

The experiment describe above is designed to answer three questions: Are participants heterogeneous with respect to the rule they follow in the distribution phase? Are participants consistent across slightly different situations? Are participants' preferences for fairness self-serving? Relying on prior literature, we tentatively derive the following predictions.

1. Participants' behavior can be better described by a multiplicity of fairness rules. Specifically, we expect to find participants whose decisions can be classified by either a purely selfish behavioral rule or the three fairness rules explained above.
2. Participants that display purely selfish behavior are consistent across contexts. By consistency we mean that a particular individual follows always the same behavioral rule, independently of the context. This will be more common among selfish subjects because they do not face any trade-off between selfish consumption and other-regarding motives. Once they act selfishly, they would not find reasons to deviate from their maximizing strategy. Hence, subjects will take consistently selfish decisions over time and across contexts. This is indeed what Brosig et al (2007) find.
3. Non-selfish decisions are context-dependent (List 2007; Bardsley 2008). Hence, those individuals who are motivated by notions of fairness behave more inconsistently. Deviating from fairness rules will be more likely to occur than deviations from a purely selfish behavior. Deviations can be due to the self-serving bias phenomena, (see among others, Miller and Ross 1975; Arkin, et al 2007) leading participants to make inconsistent choices. We predict that subjects endowed with preferences for fairness will chose the fairness rule that benefits

them the most in any given moment. The amount participants will take for themselves in a given period,  $m_{it}$ , will be given by the following options:

$$\begin{aligned} m_{it} &= m^{SE} \text{ if } m^{SE} > m^{LE} \text{ and } m^{SE} > m^L \\ m_{it} &= m^{LE} \text{ if } m^{LE} > m^{SE} \text{ and } m^{LE} > m^L \\ m_{it} &= m^L \text{ if } m^L > m^{SE} \text{ and } m^L > m^{LE} \end{aligned}$$

In our experiment subjects will face different contexts. The heterogeneity of situations in the experiment guarantees that one specific rule is more beneficial in one context but not in all the contexts. Thus, unselfish subjects will consider as fair different distributions according to the most profitable rule in each situation.

## 5 Results

The data comprise 5 experimental sessions involving a total of 60 subjects. Each session lasted for 20 periods. Given that each subject makes a decision in each of the 20 periods, we have a total of  $20 \times 60 = 1200$  distribution decisions.

In this section, we first present some descriptive statistics to show a general picture of the data. We then classify decisions using the three fairness rules described in section 3. After that, we explore individual consistency across periods. Finally, we analyze the factors that lead participants to take inconsistent decisions.

### 5.1 Frequency of different allocation rules

Figure 1 shows the relative amount participants allocate to their partner. As it is commonly found in standard dictator games (Camerer 2003), the distribution is bimodal with peaks at 0% and 50%, corresponding to the typical allocation of a purely selfish and a strict egalitarian person, respectively. Allocations between these two points, as well as those above the 50-50 distribution, indicate the presence of additional behavioral rules.

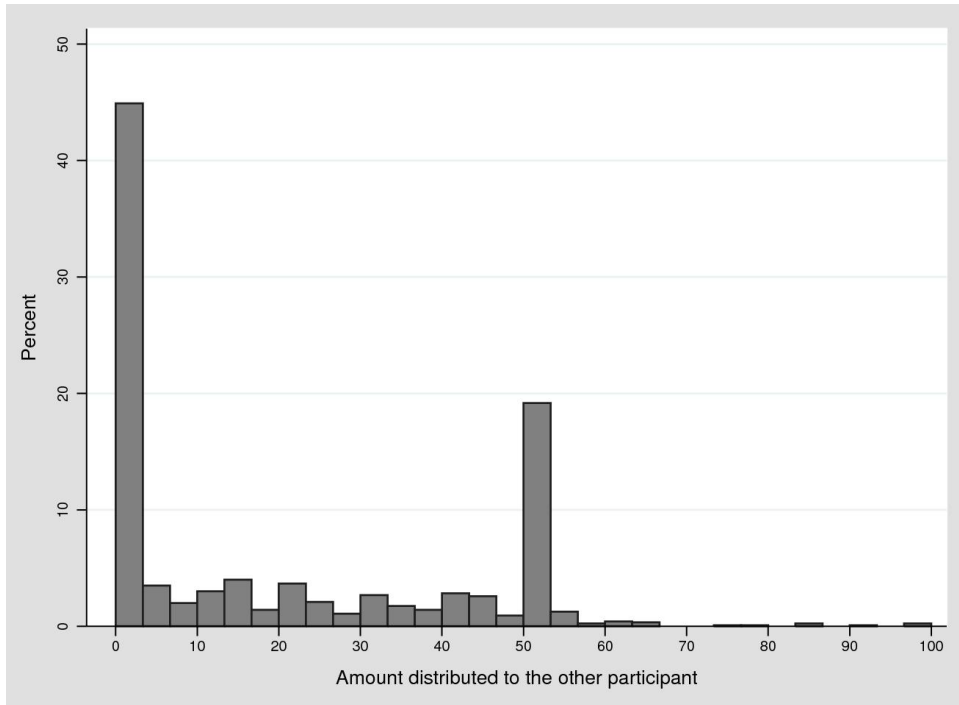


Figure 1: Relative amount allocated to the other participant

Table 1 reports the percentage and the frequency of decisions taken according to each of the fairness rule we consider in this paper. We classify decisions according to the exact prediction of each rule. The large majority of decisions are either purely selfish or can be classified using one of the three fairness ideals described in section 3, and this is true for period one (70%) as well as for the twenty periods (69%). An additional rule, defined as Charity, explains an extra 10% of behavior.<sup>7</sup> All other decisions that are not able to classify are in the ‘Other group’. Given that 80% of decisions exactly coincide with one of the five behavioral rules described so far, in the following analyses we will not put any additional structure on the data in order to avoid an over-fitting of the data.<sup>8</sup>

<sup>7</sup>The group defined as Charity consists of subjects that give less than 4 tokens and more than nothing. Note that the exchange rate is 3 tokens 1 GBP. Frohlich et al (2004) and Becker and Miller (2009) find a similar behavior rule. This behavior is, moreover, confirmed by the answers subjects gave in the post-experimental questionnaire. For instance, subject 16 answered the question ‘Have you followed any particular rule to distribute the total number of tokens?’ in the following way: ‘Largest number divisible by three’.

<sup>8</sup>Notice that the sum of the percentage in Table 1 is higher than 100%. That is because of in some cases the prediction of the different rules coincided.

| <b>Rules</b>               | 1 period     | 20 periods      |
|----------------------------|--------------|-----------------|
| <i>Strict Egalitarian</i>  | 13/60(21.7%) | 222/1200(18.5%) |
| <i>Liberal Egalitarian</i> | 5/60(8.3%)   | 105/1200(8.8%)  |
| <i>Libertarian</i>         | 7/60(11.7%)  | 107/1200(8.9%)  |
| <i>Selfish</i>             | 24/60(40%)   | 513/1200(42.8%) |
| <i>Charity</i>             | 7/60(11.7%)  | 109/1200(9.1%)  |
| <i>Others</i>              | 11/60(18.3%) | 262/1200(21.8%) |
| <i>Total explained</i>     | 81.7%        | 78.2%           |
| <i>Total</i>               | 111.7%       | 109.9%          |

Table 1: Percentage and frequency of decisions according to one rule.

In this kind of situation it is realistic to assume some behavioral noise around the point predictions. Treating an one-token deviation above or below the prediction for each rule as noise, we do not classify much more decisions (82%) (See Table 5 in the appendix). Doing a k-means cluster analysis,<sup>9</sup> which classify decision in different categories (selfish, strict egalitarian, liberal egalitarian, libertarian, charity and other) we obtain similar results (see Table 6 in the Appendix).

Given the nature of our data, assuming some deviations (errors) poses a trade-off. In some cases, the predictions of several rules are very close to one another. Therefore, allowing for some deviations may increase the number of classified decisions, but it also increases the risk of classifying the same decision in more than one category. If we allow for a one-token mistake, we are not able to disentangle whether this subject is acting in a strict egalitarian or in a libertarian way.

The percentage of coincidence increases as we increase the size of the permitted deviation. Table 5 shows that this percentage goes from 9.8% to 41.9%. In addition, small deviations from the theoretical rule’s prediction are not always mistakes. When we consider fairness decisions, small deviations can also reflect a self-serving

---

<sup>9</sup>This method of cluster analysis classifies observations into k number of clusters, in which each observation belongs to the cluster with the nearest mean. Alternative clustering methods provide similar results.

behavior.<sup>10</sup> This is shown by Figures 4 and 5 in the Appendix, where most of the deviations in allocations to self are above what one fairness rule predicts. After all these considerations and with 80% of the individual decisions precisely classified, we will not assume any deviation, and the results can be interpreted as a lower bound in rule-following behavior.

Focusing now on the frequency of different behavioral rules here considered, we find that selfish behavior is considerably above the level reported in previous real-effort experiments. In our experiment, more than 40% of decisions can be classified as selfish, in contrast to 30% reported in Cappelen et al. (2007) and 10% reported in Frohlich et al. (2004). About 20% of decisions coincide with a strict egalitarian rule, and slightly less than 20% with liberal egalitarian (8.8%) or a libertarian (8.9%) rules. These levels are below the ones reported in previous studies.<sup>11</sup>

At an aggregate level, the frequencies of decisions in the first period (left column, table 1) and in the complete experiment (right column, table 1) are strikingly similar. We will explore the consistency of decisions at the individual level in the next section. Before that, we will study whether participants' behavior evolve over time as well as other behavioral determinants.

## 5.2 Behavioral determinants

Figure 2 plots the period-average relative amount participants keep in the distribution phase. We observe that subjects keep on average 75% of the total joint amount being the trend slightly increasing.

Table 2 displays the results of a fixed effect model.<sup>12</sup> Our dependent variable is the

---

<sup>10</sup>This statement is supported by the answer of subject 35 to the question Q5: Have you followed any particular rule to distribute the total number of tokens?. A: Try to be fair, while obviously favouring myself slightly.

<sup>11</sup>Papers that study the multiplicity of fairness rules have predominantly focused on one-shot games. The repeated setting that we study may lead to an increase of selfish behavior.

<sup>12</sup>We have the same results using a random effects GLS model with clustering at the individual level.

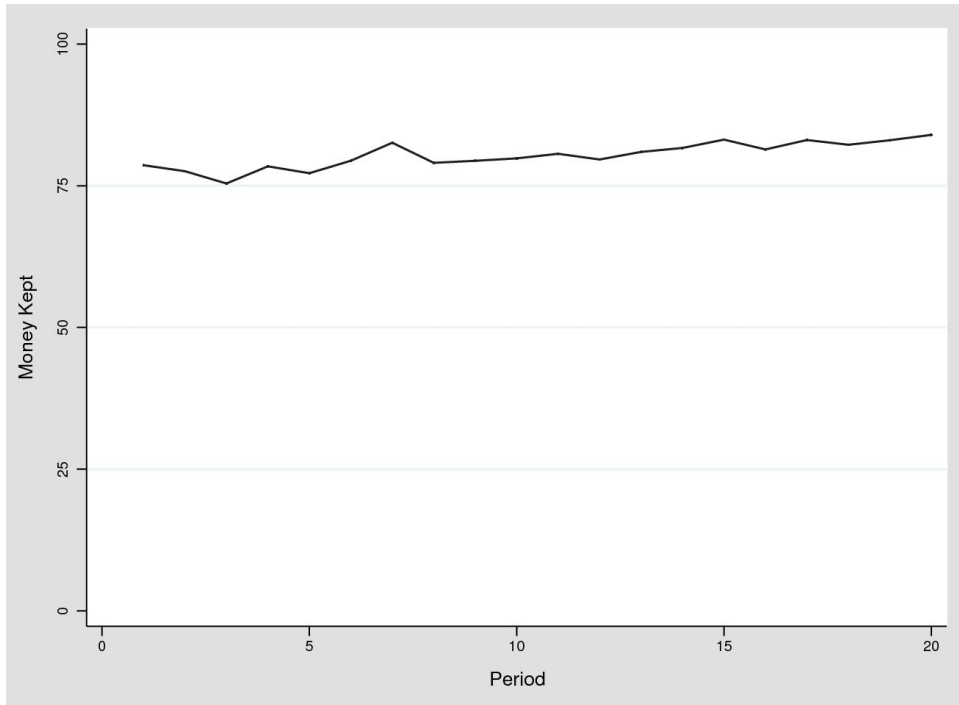


Figure 2: Distribution across periods

relative amount subjects keep for themselves. The variables we consider relevant for explaining distribution behavior are the following: *Period* that takes values from 1 to 20. *Shock* is a dummy that takes value 1 if the participant has suffered a shock in this period and 0 otherwise. *Final Endowment* is the subject's relative contribution to the total amount after the shock.

We observe that *Period* and *Final Endowment* affect individual decisions. As we have observed in Figure 4, subjects tend to keep for themselves a higher proportion of the total joint amount as time goes on. It seems that subjects learn to behave selfishly, that is in line with Brosig et al. (2007). Also, *Final Endowment* has a positive effect in subject's decisions. They keep a higher proportion of the total amount as their contribution increase.

In the next subsection we will show whether some subjects are immune to these effects and they take consistent decisions.



|                                    | Marginal Effects | Standard errors |
|------------------------------------|------------------|-----------------|
| <i>Cons</i>                        | 70.630***        | (1.460)         |
| <i>Period</i>                      | 0.334***         | (0.051)         |
| <i>Shock</i>                       | 0.911            | (0.691)         |
| <i>Final endowment</i>             | 0.115***         | (0.022)         |
| <hr/>                              |                  |                 |
| <i>Number of observations=1200</i> | F(59,1137)=72.90 | Prob>F=0.0000   |

Table 2: Behavioral Determinants

Note: We use a fixed-effects model. The dependent variable is the relative amount of money subjects keep for themselves. \*\*\* denotes significance at the 1% level.

### 5.3 Consistency of individual decisions

Table 3 reports the proportion of subjects that choose consistently with one of the rules 17 times or more.<sup>13</sup> Only Selfish, Strict Egalitarian and Charity rules are consistently chosen. No individual distributed consistently according to Liberal Egalitarian or Libertarian rules, even allowing for 3 inconsistent decisions.

|                            | 20 times<br>(in every round) | 19 times  | 18 times  | 17 times<br>(in 85% of rounds) |
|----------------------------|------------------------------|-----------|-----------|--------------------------------|
| <i>Strict Egalitarian</i>  | 3 (5%)                       | 4 (6.6%)  | 6(10%)    | 7(11.6%)                       |
| <i>Liberal Egalitarian</i> | 0                            | 0         | 0         | 0                              |
| <i>Libertarian</i>         | 0                            | 0         | 0         | 0                              |
| <i>Selfish</i>             | 18(30%)                      | 19(31.6%) | 20(33.3%) | 21(35%)                        |
| <i>Charity</i>             | 1(1.7%)                      | 1(1.7%)   | 1(1.7%)   | 1(1.7%)                        |

Table 3: Numbers and percentages of subjects applying each rule most of the time.

Figures 4 and 5 show graphically the level of individual consistency. Both figures provide individual plots of the relation between the amount participants keep for them and the endowment before the shock (Figure 4) and after the shock (Figure 5). More specifically, figure 4 shows whether subjects follow the Strict egalitarian

<sup>13</sup>We use three-times mistakes because no subject takes decisions according to one of the three main fairness rules more than 12 times and less than 17 times.

rule (middle horizontal line), the Liberal egalitarian rule (decision on the  $45^\circ$  line) or they are behaving purely selfishly (top horizontal line). In contrast, figure 5 does the same for the Libertarian rule (decisions in the  $45^\circ$  line).

Approximately half of the subjects make consistent choices, i.e., all their decision fall along the same line. They follow the same rule across the twenty periods. In line with our second prediction, a large number of participants (35%) are consistently selfish. However, there is a non-negligible proportion of subjects that consistently apply the Strict Egalitarian rule (11.7%). There is only one participant that consistently chooses distributions that fit the Charity rule. Finally, we find no single participant that consistently applies the Liberal Egalitarian or the Libertarian ideal.

Brosig et al. (2007) find selfish subjects to be highly consistent. We additionally show that strictly egalitarian individuals also seem to be consistent. Although this result has not been shown in an economic experiment before, Messick and Sentis (1979) find that strict egalitarians are consistent with their choices using hypothetical questionnaires.

In the next section we address the issue of the large number of inconsistencies.

## 5.4 Self-serving bias

As described in Section 4, one of the potential explanations for the inconsistency of some participants might be self-serving bias. To test this hypothesis, we analyze the behavior of subjects that follow some of the fairness rules presented in section 3. We then define whether the chosen rule in this period gives the highest possible individual payoff among the three rules. We say that a rule is optimal when its associated payoff is the highest among the three fairness rules. If a participant mostly uses the optimal rule, we will consider her behavior self-serving.

Figure 3 shows the proportion of choices that coincides with a strict egalitarian, a liberal egalitarian and a libertarian rule. In each column, the black area shows the proportion of decisions that are optimal and the white area the proportion of non-optimal decisions. The percentage of decisions taken according to one rule when

this rule is the most advantageous is strikingly high (77.1%). Sixty-one percent of strict egalitarian decisions were taken when this rule was the optimal. In the Liberal egalitarian and Libertarian cases this percentage is higher, being 82.7% and 87.9%, respectively. The percentage of strict egalitarian decisions is lower due to a proportion of subjects that choose consistently this rule. Considering only inconsistent subjects, the percentage of decisions taken according to the strict egalitarian rule when this is the optimal rule increases considerably (76.7%). This clearly shows that decisions from inconsistent subjects are biased.

A one-side binomial test of proportions confirms our hypothesis. We reject the null hypothesis of equal proportion of optimal and non-optimal decisions ( $p < 0.001$ ). It seems that inconsistent participants are biased in a self-serving manner.<sup>14</sup>

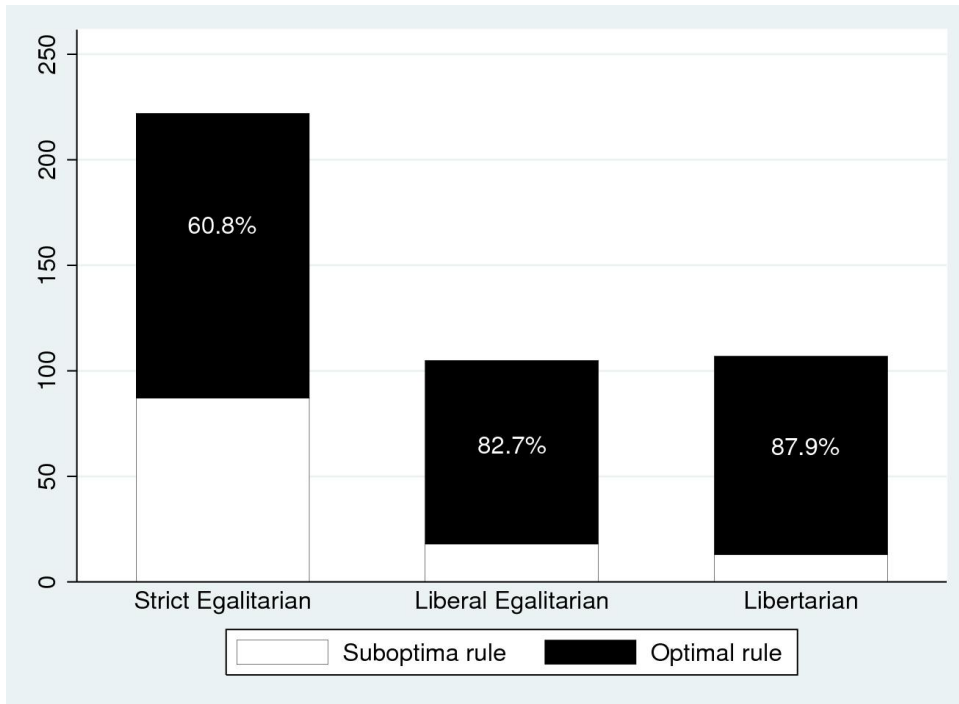


Figure 3: Self-serving bias

For a robustness check, we now replicate the self-serving result at the individual level, using a multinomial logit model (see Table 4). Our dependent variable has

<sup>14</sup>Subject number 2 confirms that she is acting in a self-serving bias manner. She answered the question Q8: Have you distributed: equally, proportionally before the shock, proportionally after the shock, other; as following: ‘did all three depending on the situation’.

four categories: decisions taken according to the Strict Egalitarian rule, decisions consistent with the Liberal Egalitarian rule, decisions in line with the Libertarian prediction and all other decisions. This latter category will be used as the base outcome for comparison.

We study the probability of choosing one of the three fairness rules. We regress the rule chosen by subject on a dummy indicating whether this rule is optimal or not and controlling for other control variables. *Shock* is a dummy variable that takes value 1 when the participant has suffered a shock in the current period and 0 otherwise. *Final Endowment* represents the relative amount of money each participant has contributed to the group account after the shock. Finally, *Period* takes values from 1 to 20 and it will indicate a possible time effect.

Consistent with the non-parametric test, we find that the optimal rule variable predicts subjects' behavior in all the three cases. The probability of choosing according to one of the three fairness rules increases when this specific rule gives the highest payoff to the subject.

*Shock* and *Final Endowment* variables also explain subjects' behavior and reinforce the self-serving explanation. *Shock* only matters in the second case, the Liberal Egalitarian, increasing the probability of behaving according to what this rule predicts. However, *Final Endowment* has an impact in the probability to choose any of the three fairness rules, although in different directions. The probability of behaving as a strict egalitarian person decreases as the participant's *Final Endowment* increases. On the contrary, this variable has a positive effect in the probability of choosing according to the Liberal egalitarian or Libertarian rule.

The results from the multinomial logit regression are consistent with the non-parametric results, showing a very strong association between choosing a rule and the fact that this rule is optimal. This supports the idea that inconsistencies are due to self-serving bias.

|   | Strict egalitarian   | Liberal egalitarian | Libertarian         |
|---|----------------------|---------------------|---------------------|
| <i>Optimal rule</i>   | 0.323***<br>(0.042)  | 0.288***<br>(0.032) | 0.290***<br>(0.020) |
| <i>Shock</i>  | -0.053<br>(0.033)    | 0.063**<br>(0.026)  | -0.001<br>(0.013)   |
| <i>Final endowment</i>  | -0.009***<br>(0.003) | 0.005***<br>(0.002) | 0.005***<br>(0.002) |
| <i>Period</i>   | 0.0001<br>(0.003)    | -0.0002<br>(0.002)  | -0.001<br>(0.001)   |
| Log Likelihood=-744.6237    Observations=1337    Pseudo $R^2$ =0.4229 |                      |                     |                     |

Table 4: Multinomial Logit

\*\*\* $p < 0.01$  \*\* $p < 0.05$  \* $p < 0.1$ .

## 5.5 Summary of the results

We find that the large majority of the decisions made in our experiment can be explained with reference to a limited number of behavioral rules. 80% of choices, both in the first period and the whole experiment, conform to one of the following rules: selfishness, strict egalitarian, liberal egalitarian, libertarian, and charity. We observe a slight tendency but not significant towards more selfish allocations, but fairness rules do not vanish even in the last periods of the experiment. Therefore, we confirm our first prediction.

Although we find a multiplicity of behavioral rules, participants do not use different rules in the same way. Specifically, whereas people that choose either selfish or strict egalitarian rules are highly consistent across periods, we find no consistent use of either the liberal egalitarian or the libertarian rule. This confirms our second prediction regarding consistency of selfish individuals. The consistency of strict egalitarian participants is a result that we had not predicted, but that it is in line with at least one previous study (Messick and Sentis, 1979).

Finally, we provide a reason why approximately half of the participants behave inconsistently. We show that they apply the fairness rule that is most beneficial to

them in a given context. This means that they do not stick to a rule, but switch between fairness rules in order to maximize their payoff at the same time that they behave fairly.

## 6 Discussion

The main goal of this paper is to test the assumption of individual consistency in fairness models (Camerer and Fehr, 2004). To achieve this end, we use a within-subjects design where each subject makes twenty distribution decisions. Exogenous and endogenous factors change across decisions. The manipulation of these two sets of factors allows us to derive different point predictions for each rule and period, and therefore we are able to study individual consistency across periods. This makes our design unique to study consistent and inconsistent behavior in an incentive compatible manner and in a repeated setting.

Interestingly, we find that although a small set of rules can explain most of our data, the majority of participants do not behave consistently with one rule, but switch among different rules. We provide a plausible explanation about individual inconsistencies. The self-serving bias, broadly studied in the psychological literature, seems to be the main reason for inconsistent behavior.

We find, however, a new result in the experimental economics literature. We predicted that only selfish subjects would make consistent decisions, but we find experimental evidence that suggests that a non-negligible number of participants consistently split the pie equally. In the following, we provide an intuition for this result.

The 50-50 norm is probably the best known and accepted rule of distributive justice. Previous experimental results on the Ultimatum Game and related experimental paradigms confirm so (Güth et al. 2001). In our experiment as well as in real life, the definition of alternative rules such as equity may be ambiguous. For instance, participants may consider it fair to distribute the money according to effort or to outcomes. In contrast, the definition of the equality norm is always the same.

An alternative explanation of the higher frequency and consistency rate of the strict egalitarian rule is the one proposed by Andreoni and Bernheim (2009). They claim that social image may help to explain why the equal split is a widespread norm in many social contexts. Dividing the pie equally is a clear signal of fair behavior. On the contrary, people may feel that an alternative fairness rule, e.g., liberal egalitarian or libertarian, do not convey as clear signal of fairness as the egalitarian rule.

## 7 References

Adams, J. Stacy., 1965. Inequality in Social Exchange, in *Advances in Experimental Social Psychology*, vol. 2. Leonard Berkowitz, ed. NY, London: Academic Press Inc, pp. 267-99.

Andreoni, J., and Bernheim, B.D., 2009. Social Image and the 50-50 norm: A theoretical and experimental analysis of audience effects. *Econometrica*, 77(5), 1607-1636.

Andreoni, J., and Miller, J., 2002. Giving According to GARP: An Experimental Test of the Consistency of Preferences for Altruism. *Econometrica*, 70(2), 737-53.

Arkin, R., Cooper, H., and Kolditz, Thomas., 2007. A statistical review of the literature concerning the self-serving attribution bias in interpersonal influence situations. *Journal of Personality*, 48(4), 435-448.

Blanco, M., Engelmann, D., Kock, A., 2010. A Within-Subject Analysis of Other-Regarding Preferences. *Games and Economic Behavior*, forthcoming.

Babcock, L., Loewenstein, G., Issacharoff, S., and Camerer, C., 1995. Biased Judgments of Fairness in Bargaining. *American Economic Review*, 85(5), 1337-43.

Babcock, L., Wang, X., and Loewenstein, G., 1996. Chossing the Wrong Pond: Social Comparisons in Negotiations that Reflect a Self-Serving Bias. *Quarterly Journal of Economics*, 111(1), 1-19.

Babcock, L., and Loewenstein, G., 1997. Explaining Bargaining Impasse : The Role of Self-Serving Biases. *Journal of Economic Perspectives*, 11(1), 109-26.

Bardsley, N., 2008. Dictator game giving: altruism or artefact? *Experimental Economics*, 11, 122-133.

Becker, A., and Miller, L., 2009. Promoting justice by treating people unequally: an experimental study. *Experimental Economics*, 12(4), 437-449.



Bolton, G.E., and Ockenfels, A., 2000. ERC: A theory of equity, reciprocity, and competition. *American Economic Review*, 90(1), 166-193.

Brosig, J., Riechmann, T., and Weimann, J., 2007. Selfish in the End?: An Investigation of Consistency and Stability of individual Behaviour. *MPRA Paper* No. 2035.

Camerer, C., 2003. Behavioral Game Theory: Experiments in Strategic Interaction. *Princeton University Press, New York*.

Camerer, C.F., Fehr, E., 2004. Measuring Social Norms and Preferences Using Experimental Games: A Guide for Social Scientists. In *Foundations of Human Sociality: Economic Experiments and Ethnographic Evidence from Fifteen Small-Scale Societies*, eds. Joseph Henrich et al., 55-95. Oxford: *Oxford University Press*.

Cappelen, A., Astri, D.H, Sorensen, E., Tungodden, B., 2007. The Pluralism of Fairness Ideals: An Experimental Approach. *American Economic Review*, 97(3), 818-827.

Carpenter, J., Holmes, J., and Matthews, P., 2010. Jumping, Squatting, Sniping, Soaring and Stumbling at the Silents: Does it matter for Charities? *Working paper*.

Charness, G., and Matthew, R., 2002. Understanding Social Preferences with Simple Tests. *Quarterly Journal of Economics*, 117(3), 817-69.

Cherry, T.D., Frykblom, P., and Jason F. Shogren, J.F., 2002. Hardnose the Dictator. *The American Economic Review*, 92 (4), 1218-1221.

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

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

Fisman, R., Kariv, S., and Markovits, D., 2007. Individual Preferences for Giving. *American Economic Review*, 97(5), 1858-1876.

Frank, R.H., Gilovich, T., and Regan, D.T., 1993. Does Studying Economics Inhibit Cooperation? *The Journal of Economic Perspectives*, 7(2), 159-171.

Frohlich, N., Oppenheimer, J., A.Kurki., 2004. Modeling other-regarding preferences and an experimental test. *Public Choice*, 119: 91-117.

Greiner, B., 2004. An online recruitment system for economic experiments. In K. Kremer., V. Macho. *Forschung und wissenschaftliches Rechnen*, 79-93.

Guth, Werner., 1988. On the Behavioral Approach to Distributive Justice: A Theoretical and Experimental Investigation. In *Aspects of Distributive Justice*. Shlomo Maital, ed. NY: NYU Press, pp. 703-17.

Guth, W., Huck, S., and Müller.,W., 2001. The Relevance of Equal Split in Ultimatum Games. *Games and Economic Behavior*, 37, 161-169.

Homans., 1961. Social Behavior: Its Elementary Forms. *Harcourt, Brace & World, New York*.

Konow, James., 1996. A Positive Theory of Economic Fairness. *Journal of Economic Behavior and Organization*, 31(1), 13-35.

Konow, J., 2000. Fair Shares: Accountability and Cognitive Dissonance in Allocation Decisions. *American Economic Review*, 90 (4), 1072-1091.

Konow, J., 2003. Which is the Fairest One of All? A Positive Analysis of Justice Theories. *Journal of Economic Literature*, XLI, 1188-1239.

Konow, J., 2005. Blind Spots: The Effects of Information and Stakes on Fairness Bias and Dispersion. *Social Justice Research*, 18(4), 349-390.

List, J.A., 2007. On the Interpretation of Giving in Dictator Games. *Journal of Political Economy*, 115, 482-493.

Loewenstein, G., Issacharof, S., Camerer, C and Babcock, L., 1993. Self-Serving Assessments of Fairness and Pretrial Bargaining. *Journal of Legal Studies*, 22(1),

135-59.

Messick, D.M., and Sentis, K.P., 1979. Fairness and Preferences. *Journal of Experimental Social Psychology*, 15, 418-434.

Miller, D.T., and Ross, M., 1975. Self-serving Biases in the Attribution of Causality: Fact or Fiction? *Psychological Bulletin*, 82 (2), 213-225.

Rabin, Matthew., 1993. Incorporating Fairness into Game Theory and Economics. *American Economic Review*, 83(5), 1281-302.

Selten, Reinhard., 1978. The Equity Principle in Economic Behavior. In Decision Theory and Social Ethics; *Issues in Social Choice*. H. Gottinger and W. Leinfellner, eds. Dordrecht: Reifel Pub., pp. 289-301.

## 8 Appendix

|                            | 0 tokens   | 1 tokens   | 2 tokens    | 3 tokens    |
|----------------------------|------------|------------|-------------|-------------|
| <i>Strict Egalitarian</i>  | 222(18.5%) | 253(21.1%) | 296(24.7%)  | 328(27.3%)  |
| <i>Liberal Egalitarian</i> | 105(8.8%)  | 131(10.9%) | 175(14.6%)  | 194(16.1%)  |
| <i>Libertarian</i>         | 107(8.9%)  | 143(11.9%) | 205(17.1%)  | 239(19.9%)  |
| <i>Selfish</i>             | 513(42.8%) | 558(46.5%) | 598(49.8%)  | 622(51.8%)  |
| <i>Charity</i>             | 109(9.1%)  | 136(11.3%) | 159(13.3%)  | 233(19.4%)  |
| <i>Others</i>              | 262(21.8%) | 215(17.9%) | 111(9.3%)   | 89(7.4%)    |
| <i>Total explained</i>     | 938(78.2%) | 985(82.1%) | 1089(90.8%) | 1111(92.6%) |
| <i>Total</i>               | 109.9%     | 119.6%     | 128.8%      | 141.9%      |

Table 5: Tokens deviations

| Possible rules             | 4 Categories                                | 5 Categories                          | 6 Categories                       |
|----------------------------|---|---------------------------------------|------------------------------------|
| <i>Strict Egalitarian</i>  | 306(25.5%)                                  | 295(24.6%)                            | 269(22.4%)                         |
| <i>Liberal Egalitarian</i> | 119(9.9%)                                   | 104(8.7%)                             | 84(7.0%)                           |
| <i>Libertarian</i>         | 181(15.1%)                                  | 117(9.8%)                             | 87(7.3%)                           |
| <i>Selfish</i>             | 594(49.5%)                                  | 567(47.3%)                            | 580(48.3%)                         |
| <i>Charity</i>             | —   | 117(9.8%)                             | 113(9.4%)                          |
| <i>Others</i>              | —   | —                                     | 67(5.6%)                           |
|                            | Observations=1200                           | Observations=1200                     | Observations=1200                  |
|                            | Pearson $\chi^2(177) = 1.8e + 0.3$          | Pearson $\chi^2(236) = 2.1e + 0.3$    | Pearson $\chi^2(295) = 2.2e + 0.3$ |
|                            | Pr=0.000                                    | Pr=0.000                              | Pr=0.000                           |
|                            | Likelihood-ratio $\chi^2(177) = 1.8e + 0.3$ | Likelihood $\chi^2(236) = 2.0e + 0.3$ | Pearson $\chi^2(295) = 2.1e + 0.3$ |
|                            | Pr=0.000                                    | Pr=0.000                              | Pr=0.000                           |

Table 6: Cluster analysis

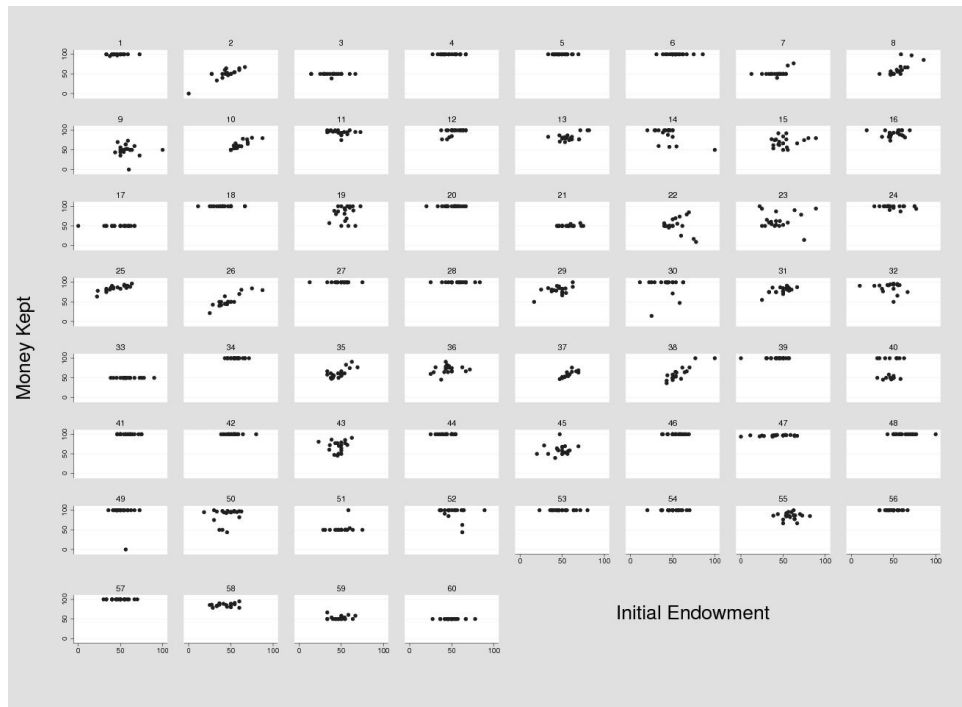


Figure 4: Strict Egalitarian and Liberal Egalitarian rule



Figure 5: Libertarian Rule

## Instructions

Dear participants,

Welcome and thank you for participating in our experiment. The experiment will last for about one and a half hour. Please do remain quiet from now on until the end of the experiment. You will have the opportunity to ask questions in a few minutes, they will be answered privately.

## INSTRUCTIONS

The experiment consists of 20 periods and in each period there will be two phases.

### First phase

- At the beginning of every single period, groups of two people will be formed. You will be randomly paired with another participant in this room. You will remain together until the end of that period. None of you will know with whom you have been paired.
- On the computer screen you will see a series of puzzles in which the letters of a word are scrambled. It is your task to unscramble them. You will see one scrambled word at a time, with a blank below each given letter. In each blank, enter the letter that you think belongs in that space in the correct, unscrambled word. In each blank, please enter **only one letter, with no spaces**, and use only the letters given in the original scramble. None of the words is an acronym. **The words are the same for all the participants in this room and they will follow the same sequence.**
- You may use the mouse or the TAB (on the keyboard) to switch to the next blank.
- You will have a total of **90 seconds** to correctly solve as many scrambles as you can. For each correct answer, you will receive 4 tokens. In each period there are **ten words to unscramble.**



- You may leave a puzzle blank, but once you click on the 'Submit and Continue to Next Puzzle' button, you **will not be able to return to that puzzle**.
- At the end of the first phase, you will know: the number of tokens you receive in that period; the number of tokens the other participant of your group receives in that period; and the total number of tokens the two of you receive.

### First phase

- The second phase starts with a random shock. This shock may affect the number of tokens of each participant with 50% probability. It is independent that the other participant was affected by the shock for you to be affected as well. The shock works in one of the following ways:
  1. By halving your tokens;
  2. By halving the tokens of the other participant;
  3. By halving the tokens of both participants in the group;
  4. By not halving any participant's tokens.
- Both participants in the group have to decide how to distribute the total number of tokens of the group (your tokens + the other participant's tokens) between the two group members. The decision will be taken individually and anonymously. Neither you nor the other participant will know the decision of the other.
- After the 20 periods have been completed, the computer will randomly choose one of the twenty periods to be paid. Only the decision of one participant in each group will be implemented. All earned tokens will be exchanged into pounds at the end of the experiment and paid in cash according to the following exchange rate:

$$3 \text{ tokens} = \text{£}1$$

Finally, each participant will be informed about her/his earnings, which will consist of the profit from the experiment plus the participation fee (£4).

We ask you to remain quiet during the whole experiment. Those who do not respect the silence requirement will be asked to leave the experimental room. Once the experiment is finished, please remain seated. We will need between 10 and 15 minutes to calculate your payment. We will move to another room and you will be called up successively by the number on your table; you will then receive an envelope with your earnings and you will be asked to sign a receipt.

Finally, note that your participation is considered voluntary and you are free to leave the room at any point if you wish to do so. In that case, we will only pay you the participation fee of £4.

**Please leave these instructions on your table when you leave the room.**

You can take notes on these pages if you wish to do so.

If you have any questions, please raise your hand now.

To make sure you understand the instructions, could you please provide an example (of your choice) of the possible outcomes and decisions you and the other participant in your group may make and the earnings you would get at the end of one period:

- Number of words you have unscrambled (out of 10): ...
- Number of words the other participant has unscrambled (out of 10): ...
- Your number of tokens is (number of words you have unscrambled x 4): tokens ...
- The number of tokens of the other participant is (number of words the other participant has unscrambled x 4):tokens ...
- Total number of tokens in the group (Your tokens + the other participant's tokens): tokens ...
- Shock:

**Imagine that the shock has only been affected your number of tokens.**

- Number of tokens after the shock

**Your number of tokens after the shock is: tokens ...**

**The number of tokens of the other participant is: tokens ...**

**Total number of tokens to be distributed: tokens ...**

- How much do you distribute to the other participant? tokens ...
- How much does the other participant distribute to you? tokens ...
- If your decision is chosen, which is your profit? tokens... And the profit of the other participant? tokens ...
- If the other participant's decision is chosen, which is your profit? tokens... And the profit of the other participant? tokens ...

If you have any questions about these instructions, please raise your hand now and wait for the experimenter to come to you. Please return to the computer.

## Questionnaire

- **Q1:** Are you:  
  
    **Male**  
  
    **Female**
- **Q2:** What year were you born?
- **Q3:** Which is your nationality?
- **Q4:** What are you studying?
- **Q5:** Have you followed any particular rule to distribute the total number of tokens?
- **Q6:**
  1. Have you changed the rule depending of the situation?
  2. In case you used different rules, why did you do so?
- **Q7:** Please, could you tell us whether some of the following factors influenced your decision:
  - Your number of tokens in the first phase
  - The number of tokens of the other participant in the first phase
  - The number of tokens of both participants in the first phase
  - Your number of tokens after the shock
  - The number of tokens of the other participant after the shock
  - The number of tokens of both participants after the shock
  - Other factors (Please indicate)
- **Q8:** Have you distributed:
  - Equally?
  - Proportionally before the shock?
  - Proportionally after the shock?
  - Other (Please indicate)