NHH

INSTITUTT FOR SAMFUNNSØKONOMI

DEPARTMENT OF ECONOMICS

SAM 40 2007

ISSN: 0804-6824

DECEMBER 2007

Discussion paper

The importance of moral reflection and self-reported data in a dictator game with production

BY

ALEXANDER W. CAPPELEN, ASTRI DRANGE HOLE, ERIK Ø. SØRENSEN, AND **BERTIL TUNGODDEN**

Norges Handelshøyskole

NORWEGIAN SCHOOL OF ECONOMICS AND BUSINESS ADMINISTRATION

The importance of moral reflection and self-reported data in a dictator game with production*

Alexander W. Cappelen† Astri Drange Hole‡
Erik Ø. Sørensen§ Bertil Tungodden¶

January 9, 2008

Abstract

This paper studies how individual behavior is affected by moral reflection in a dictator game with production, and the informational value of self-reported data on fairness. We find that making individuals reflect on fairness before they play the dictator game has a moderate effect on the weight attached to fairness in distributive choices, and a strong effect on what people consider fair. Furthermore, we find that self-reported data have substantial informational value, but still do not add explanatory power to a random utility model estimated on purely behavioral data. Finally, by studying the behavior of individuals who deviate from their self-reported fairness ideal, we do not find much support for the hypothesis that people are self-serving in their choice of fairness ideal.

JEL codes: C91, D63. **Keywords:** dictator game, distributive justice, experimental data, fairness, moral reflection, self-serving bias, survey data.

*We are grateful for excellent research assistance from Silje Sandstad, Inga Søreide and Jan Tore Ynnesdal.

†Corresponding author. Department of Economics, Norwegian School of Economics and Business Administration, Helleveien 30, N-5045 Bergen, Norway. Fax: +47 55 95 95 43, phone: +47 55 95 95 77, email: `alexander.cappelen@nhh.no`.

‡Department of Economics, Bergen University College, P.O. Box 7030, 5020 Bergen, email: `astri.drange.hole@hib.no`.

§Department of Economics, VU University Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands, email: `sameos@gmail.com`.

¶Department of Economics, Norwegian School of Economics and Business Administration, Helleveien 30, N-5045 Bergen, Norway, email: `bertil.tungodden@nhh.no`.

1 Introduction

The dictator game has been an important device for studying social preferences, and it is well documented that, on average, people give away a substantial amount of money when they act as dictators in distributive situations (Camerer, 2003). But there is considerable heterogeneity in the extent to which people care about fairness; some give away half of the money whereas others do not give away anything. Furthermore, by studying dictator games where the distribution phase is preceded by a production phase, it has been shown that there is substantial heterogeneity in what people consider as fair in distributive situations (Cappelen, Hole, Sørensen, and Tungodden, 2007). Some find it fair to share equally, whereas others find it fair to share in proportion to individual effort or individual production.

Traditionally, these experiments have been conducted in a neutral manner, in the sense that the fairness issue involved has not been stated explicitly for the participants. This is contrary to many real life-situations, where distributive issues are presented and discussed as questions of fairness. Does this difference matter? Would people's choices in a dictator game be affected by introducing the issue of fairness explicitly, and moreover by making individuals reflect on what is fair in such situations? Moral reflection could potentially have two different effects. First, it might affect the weight that individuals attach to fairness considerations relative to their self-interest, i.e., their willingness to distribute in the way they think is fair. Second, it might affect what they consider to be the fair distribution.

In this paper, we study the effect of moral reflection in a dictator game with production. We consider two treatments that are identical, except that in one treatment we made the participants reflect on what they consider a fair distribution. More specifically, we presented the participants with distributive situations of the same kind that they would face later in the experiment, and we asked them a hypothetical question regarding what they considered to be a fair distribution in such situations. A comparison of the two treatments allows us to study the extent to which moral reflection causes people to attach more weight to fairness when they make distributive decisions, and also, importantly, to study the extent to which moral reflection changes people's perceptions of what is fair.

Our design also allows us to address the question of how we can best obtain information about people's social preferences. There are two important strands of literature on this issue. The experimental literature has typically studied this question by focusing on behavioral data generated in the lab, where real stakes are involved for the participants (see for example Cherry, Frykblom, and Shogren, 2002; Frohlich, Oppenheimer, and Kurki, 2004; Konow, 2000). A parallel literature has generated data by asking hypothetical questions about what people find to be fair (see for example Amiel and Cowell, 1992; Gaertner, 1994; Gaertner and Schwettmann, 2007; Schokkaert and Lagrou, 1983; Schokkaert and Devooght, 2003). Both literatures have obvious attractions. The real stakes in experiments ensure that the participants take the issue seriously, whereas surveys often allow questions that map more cleanly into the normative literature. To what extent do the two approaches arrive at the same conclusions regarding people's social preferences? And to what extent does the questionnaire approach generate data that are of value in explaining individual behavior? We study these questions by comparing the individuals' actual behavior with the answers they give when asked what they think is fair in hypothetical situations.

We estimate a random utility model on the behavioral data, and then, at the aggregate level we compare the population estimates of this model with the self-reported data. In order to study whether self-reported data add explanatory power, we consider the extent to which the fit of the random utility model can be improved by incorporating self-reported data. In addition, we study the match between self-reported fairness ideals and actual choices at the individual level, where we present a ternary diagram representing the a posteriori probabilities that each individual holds a specific fairness ideal. Finally, we test the hypothesis that the participants adopt their fairness ideal in a self-serving manner.

Our main findings are the following: First, moral reflection increases the weight people attach to fairness moderately, and it has a strong effect on what people consider to be fair. In particular, we find a large drop in the share of individuals who hold a strict egalitarian fairness ideal. Second, we find that self-reported data have substantial informational value, but at the same time, they do not add explanatory power to a random utility model estimated purely on behavioral data. Third, we do not find much support for the hypothesis that people are self-serving in their choice of fairness ideal.

Section 2 provides a discussion of the experimental design, and Section 3 analyzes the effect of moral reflection on people's distributive choices. Section 4 studies how much information is yielded from self-reported data compared with behavioral data. Section 5 addresses the extent of self-serving bias in the adoption of fairness ideal. Section 6 contains some concluding comments.

2 Design

The experiment is a version of a one-shot dictator game with production. In the production phase, each participant was given credits equal to 300 Norwegian kroner (NOK), approximately 50 US dollars. Each participant was also randomly assigned a low or a high rate of return. Participants with a low rate of return would double the value of any investment they made, whereas those who were assigned a high rate of return would quadruple their investment. The participants were asked to determine how much they wanted to invest in two different one-shot games. Production thus depended on factors both within and beyond individual control; the investment was clearly within individual control, and the rate of return on the investment was clearly beyond individual control.

Before they made their investment choice, all participants were informed about the rules of the game, and they were given a complete description of how the game would proceed. Their choice alternatives were limited to 0 NOK, 100 NOK, and 200 NOK, and the total amount invested in the two games could not exceed the initial credit they received. Any money they chose not to invest they kept. Thus they faced a genuine investment choice.

In the distribution phase, each participant was given information about the other participant's rate of return, investment, and production before they were asked to propose a distribution of the total income produced by the two participants. The participants were not informed about the outcome of the first game before the second game was completed, i.e., they considered the two one-shot games simultaneously. For each participant one of the two proposals (the participant's own or that of the opponent) in one of the two games was randomly selected to determine the actual outcome. An individual's total earnings from the experiment were given by the actual outcome plus the amount of money not invested.

The base treatment (B) in our study contained only the production phase and the distribution phase, where the presentation was neutral and the fairness issue never mentioned. We have reported the results from this treatment independently in Cappelen et al. (2007). In order to study the role of moral reflection and the informational value of self-reported data, we conducted a new elicitation treatment (E), where the participants were asked a question about fairness before they entered the production phase. More specifically, they were given a description of a hypothetical setting that was identical to the actual setting that they would face later in the experiment. They were presented with three fairness ideals: to share equally, to share in proportion to individual investment, and to share in proportion to individual production. In the following we refer to these fairness ideals as strict egalitarianism, liberal egalitarianism and libertarianism, respectively, but the participants were never confronted with these technical terms. To illustrate the fairness ideals, the participants were shown how each of them worked in three different distributive situations, and then they were asked to state which ideal they considered to be most fair. They were told explicitly that their answer would not in any way constrain their choices later in the experiment. Furthermore, at this stage, they had no detailed information about the production phase and the distribution phase, and thus no basis for strategic self-reporting.

At the end of the experiment, the participants were assigned codes and instructed to mail their codes and bank account numbers to the accounting division of the Norwegian School of Economics and Business Administration. Independently, the research team mailed a list including the codes and the total payment to the accounting division, which then disbursed the earnings directly to each participant's bank account. This procedure ensured that neither the participants nor the research team were in a position to identify how much each participant earned in the experiment.[1]

All participants recruited were first-year students at the Norwegian School of Economics and Business Administration, and no participant was allowed to participate more than once. In the invitation, they were told that they would initially receive 300 NOK (approximately 50 USD) to use in an experiment that would last about 40 minutes and that their total earnings from the experiment would depend on their choices. They were not informed about the purpose of

[1]Complete instructions are available upon request.

the experiment. The average hourly opportunity cost for these students would be about 100 NOK, whereas the average payment from the experiment was 476 NOK. We had 96 students in the B-treatment and 92 students in the E-treatment. All communication was anonymous and conducted through a web-based interface.

3 Is moral reflection important?

In principle, making the participants reflect on fairness could affect behavior both in the production phase and in the distribution phase of the experiment. However, we only observe a large difference between the two treatments in the distribution phase.

The investment pattern was almost the same in the two treatments: 83 out of 96 participants in the B-treatment and 89 out of 92 participants in the E-treatment invested their full endowment of 300 NOK. Of the 13 participants not investing the full endowment in the B-treatment, 12 participants invested 200 NOK.[2] We therefore focus on how moral reflection affected the choices in the distribution phase. In doing so, we distinguish between the effect on the participants' weight attached to fairness and the effect on the participants' perception of fairness.

3.1 The effect on the weight attached to fairness

Table 1 reports statistics from the distribution phase. We observe that the average share given to the other participant increases from 27.1% in the B-treatment to 34.6% in the E-treatment. This difference is statistically significant (a one-sided t-test has p-value of 0.008). We also note that the median share given increases from 29.2% in the B-treatment to 50% in the E-treatment, and that the share of participants demanding everything decreases from 31% in the B-treatment to 26% in the E-treatment.

[Table 1 about here.]

These results suggest that introducing moral reflection explicitly in the experiment, not surprisingly, increases the weight that participants attach to fairness

[2] In Cappelen et al. (2007), we showed that removing individuals who did not invest the full endowment from the analysis had no important impact on the results.

considerations. However, the effect is moderate, which probably reflects that the fairness issue is salient to most people even in the neutral version of the dictator game.

3.2 The effect on perception of fairness

In order to study whether moral reflection changes people's perception of fairness, we estimate a random utility model in which people make a trade-off between income and fairness. To do so we need to introduce some notation. Let a participant i's production be given by $x_i = a_i q_i$, where q_i is the investment and a_i is the rate of return on investment for participant i. The total production to be distributed between the two participants is then $X(\boldsymbol{a}, \boldsymbol{q}) = x_1 + x_2$, where $\boldsymbol{a} = (a_1, a_2)$ and $\boldsymbol{q} = (q_1, q_2)$. Each participant proposes an amount of income y for himself and $X - y$ for his opponent. Since almost all participants chose numbers that were multiples of 50, we restrict the choice of y to the set $\mathcal{Y}(\boldsymbol{a}, \boldsymbol{q}) = \{0, 50, 100, \dots, X(\boldsymbol{a}, \boldsymbol{q})\}$.[3]

Consider now the following simple random utility model:

$$U_i(y; \cdot) = \gamma y - \beta_i \frac{(y - m^{k(i)}(\boldsymbol{a}, \boldsymbol{q}))^2}{2X(\boldsymbol{a}, \boldsymbol{q})} + \varepsilon_{iy}. \tag{1}$$

We assume that ε_{iy} is i.i.d. extreme value distributed. The parameters $\gamma > 0$ and $\beta_i \geq 0$ determine the weight individual i attaches to income and to fairness respectively, and $m^{k(i)}(\boldsymbol{a}, \boldsymbol{q})$ specifies the amount that individual i holds to be his fair income. There are three prominent fairness ideals in this situation, all presented explicitly for the participants in the E- treatment: strict egalitarianism ($m^{SE}(\boldsymbol{a}, \boldsymbol{q}) = X(\boldsymbol{a}, \boldsymbol{q})/2$), liberal egalitarianism ($m^{LE}(\boldsymbol{a}, \boldsymbol{q}) = q_1 X(\boldsymbol{a}, \boldsymbol{q})/(q_1 + q_2)$), and libertarianism ($m^{L}(\boldsymbol{a}, \boldsymbol{q}) = a_1 q_1$).

Each person is characterized by $k(i)$ and β_i. In this section, we estimate only the distribution of $(k(i), \beta_i)$ in the sample population, and we approximate the distribution of β by a log-normal distribution, such that $\log \beta \sim N(\zeta, \sigma^2)$.[4]

Table 2 reports the estimates from this model for both treatments, and we

[3]In the two treatments only 8 offers were not multiples of 50, and in the estimation of our model we rounded these numbers to the nearest 50.

[4]For further discussion of this model, see Cappelen et al. (2007) and the appendix published on the AER website that includes extensive robustness checks with respect to model variations, `http://www.e-aer.org/data/june07/20050838_app.pdf`.

also include estimates on the pooled data. A likelihood ratio test of equality between the treatments (with $\chi^2_5 = 22.0$) has a p-value of 0.001.

[Table 2 about here.]

Comparing the estimates from the two treatments, we find that there is a substantial difference in the prevalence of fairness ideals. In particular the share of strict egalitarians is much lower in the E-treatment, and the share of libertarians is much higher in the B-treatment. These results suggest that an important effect of introducing moral reflection is that it changes some of the participant's perception of fairness.

4 The informational value of self-reported data

Do people act in accordance with their reported fairness ideal in the distribution phase? At the aggregate level, we can study this question by comparing, for the E-treatment, the estimated population shares with the self-reported population shares.

[Table 3 about here.]

In Table 3, we observe that the self-reported population shares are quite close to the estimated population shares. The self-reported data give a smaller share of strict egalitarians and a larger share of libertarians, but the overall picture is very much the same. We find this similarity striking. Our participants were not given any economic incentives when responding to the hypothetical question, and they knew that they would not be constrained in any way later in the experiment by what they self-reported. In such an environment, one might fear that the participants would not think carefully about the question and would respond more or less at random. On the contrary, our finding suggests that, to a great extent, people truthfully self-reported their fairness ideal. This provides support for the questionnaire approach to the study of social preferences.

4.1 Can self-reported data explain behavior?

To what extent can self-reported data improve our understanding of individual behavior? To address this question we apply the self-reported data to classify individuals by fairness ideal. We then compare the informational value of this classification to Bayesian classification that only relies on behavioral data. To do so, we introduce both classifications into the estimation of the random utility model (1), and we then compare how these two specifications fit the data for the E-treatment.

The classification of individuals according to self-reported data is straightforward. A person self-reporting strict egalitarianism is classified as strict egalitarian, and similarly for a person self-reporting liberal egalitarianism and libertarianism. The Bayesian classification relies on behavioral data. We apply the estimates of the random utility model to calculate a posteriori probabilities for an individual having each of the fairness ideals, and then we apply the a posteriori probabilities to classify individuals.

The a posteriori probability of an individual i having fairness ideal k is given by

$$P(k|Y_i, Z_i) = \frac{P(Y_i|k, Z_i)P(k|Z_i)}{P(Y_i|Z_i)}, \tag{2}$$

where Y_i and Z_i represent the choices and the economic environment of individual i. We assume that the proportion of individuals of each fairness ideal is independent of the environment, such that $P(k|Z_i) = P(k)$ for all Z_i. In order to calculate the a posteriori probabilities $P(k|Y_i, Z_i)$, we can then use the estimated population shares reported in Table 2 for $P(k)$, and calculate $P(Y_i|Z_i)$ and $P(Y_i|k, Z_i)$ by the unconditional and conditional likelihoods of observing this behavior given the estimates.

Given the a posteriori probabilities, there are different ways of classifying individuals. The most straightforward approach is the simple Bayes classification rule, whereby each individual is classified as having the fairness ideal k that maximizes $P(k|Y_i, Z_i)$ (McLachlan and Peel, 2000, p. 30). If we let z_{ik} be a dummy indicator for whether individual i is classified as having fairness ideal k, we can formally write the classification program as follows:

$$\max_{\{z_{ik}\}} \sum_i \sum_k z_{ik} \log P(k|Y_i, Z_i) \tag{3a}$$

$$\text{subject to} \quad \sum_k z_{ik} = 1 \text{ for all } i, \tag{3b}$$

$$\text{and } z_{ik} \in \{0, 1\} \text{ for all } i \text{ and } k. \tag{3c}$$

The problem with this classification procedure in our context is that some individuals in their distributive choices reveal very little information about their fairness ideal. The simple Bayes classification rule assigns all these individuals the fairness ideal that is most prevalent in the population (liberal egalitarianism). However, we have no reason to suspect that it is only liberal egalitarians who behave in a way that is not informative about their fairness ideal, and therefore the simple Bayes classification rule tends to classify too many individuals as liberal egalitarians.

An alternative approach is to require the share of people classified as being of each fairness type to be the same as our estimated population shares. Formally, this means that we add the restriction that

$$\sum_i z_{ik} = \text{round}_1(n\widehat{\lambda}_k) \quad \text{for all } k, \tag{3d}$$

where $\widehat{\lambda}_k$ is the estimated population share of type k, n is the sample size and round_1 is a round-to-integer operator. Adding (3d) to the program adds considerable computational complexity.[5] An exhaustive search would take excessively long time, but it is well known that often a *good* solution can be found in a much shorter time. Running for 24 hours, our algorithm stabilized at a classification that gave a value of the objective (3a) of -43.03, a value that can be compared to the value of -33.04 from the simple Bayesian classification without imposing (3d), and to the value of -78.08 that we obtain by using the self-reported classification.[6]

[Table 4 about here.]

Table 4 reports the estimates of the random utility model based on self-reported classification and Bayesian classifications with and without imposing (3d). We observe that the Bayesian approach clearly outperforms the model that relies on self-reported fairness ideals.[7] There is a substantial difference in the

[5]Binary integer programming in general is NP-hard. An exhaustive search of all ways of finding 19 strict egalitarians, 40 liberal egalitarians and 33 libertarians in our sample of 92 individuals would have to evaluate $\binom{92}{19}\binom{92-19}{40} \approx 1.4 \cdot 10^{40}$ combinations.

[6]We solve (3) using the `lp_solve` library (Berkelaar, Eikland, and Notebaert, 2007).

[7]Both the self-reported and the Bayesian classifications perform better than the model in the previous section in terms of likelihood. This should not be surprising, as estimating conditional

log likelihood values.[8] The Bayesian models estimate a much higher γ, which translates into less room for idiosyncratic "noise" in behavior. This suggests that even though self-reported data contains substantial informational value, there is even more to learn from studying behavioral data.

We also note that the estimates for the two Bayesian models are very close, which reflects that the two classification methods mainly differ in terms of how they treat people who behave in a non-informative way, e.g., those who take everything for themselves. These estimates are also very close to the estimates reported in the previous section for the E-treatment, which may serve as a consistency check of our specification.

4.2 Who misreports?

[Figure 1 about here.]

To study the match between self-reported fairness ideals and actual behavior at the individual level, we present Figure 1, in which we plot the distribution of a posteriori probabilities for each individual who has self-reported this fairness type. If an individual has chosen in such a manner that we can perfectly identify his fairness ideal, he is located at the corner representing this fairness ideal. On the other hand, if the person has not revealed any information about his fairness ideal through his choices, he is located at the point that represents the estimated population shares in this ternary diagram.

As we can see from Figure 1, for all three fairness ideals, there are individuals who are identified perfectly as having the fairness ideal they have self-reported. This is the case for 5 out of 9 strict egalitarians, 13 out of 43 liberal egalitarians and 14 out of 40 libertarians. For the rest, we cannot rule out that they have acted on a different ideal than the one they self-reported. We consider a person to have misreported his fairness ideal if, through his distributive choices, he has decreased the likelihood of having his self-reported fairness ideal. In the diagram, such an individual would be located further away from his fairness ideal than the point

on a classification is similar to introducing person fixed effects for the fairness ideal. If there is informational value in a classification, such a conditional model should allow a much closer fit to the data.

[8]The models are not nested, so likelihood ratio tests do not make sense. However, all models are discrete choice models of the same data on a fixed grid, so the log likelihood values are informative as measures of model fit.

of the estimated population shares. Table 5 summarizes this information for all three fairness ideals. We observe that there are 2 self-reported strict egalitarians, 6 self-reported liberal egalitarians and 8 self-reported libertarians misreporting. On average, 17.4% of the participants misreported, and interestingly this seems to be about equally prevalent for the different self-reported fairness ideals.

[Table 5 about here.]

5 Self-serving bias

A potential explanation for the misreporting is that the participants have a self-serving bias with respect to their perception of fairness in the distributive situations. A self-serving bias of this kind would be that not only do people trade fairness against income, but their fairness ideal is also influenced in some conscious or unconscious way by what is most beneficial for them in a particular situation (Messick and Sentis, 1983; Dana, Weber, and Kuang, 2007).

5.1 Self-serving bias among the misreporting individuals

It is interesting to study the extent to which people misreporting have acted on the fairness ideal that benefited them most. We do so by considering the correlation between the relative gain of acting on one of the two fairness ideals they did not self-report and the relative change in a posteriori probabilities.

For each individual, let $1(i)$ and $2(i)$ be an (arbitrary) order of the two fairness ideals not self-reported by individual i, while $SR(i)$ is the self-reported ideal. Furthermore, let $M^{1(i)}$, $M^{2(i)}$, and $M^{SR(i)}$ be what individual i could justify taking according to each of these fairness ideals, summed over both distributive situations. We can now define the relative gain as $(M^{1(i)} - M^{SR(i)}) - (M^{2(i)} - M^{SR(i)}) = M^{1(i)} - M^{2(i)}$. Let p_{i1} and p_{i2} be the a posteriori probabilities that individual i has acted on fairness ideal 1 and 2, and let $\widehat{p}_1$ and $\widehat{p}_2$ be the corresponding population estimates. We can now define the relative change in a posteriori probability as $(p_{i1} - \widehat{p}_{i1}) - (p_{i2} - \widehat{p}_{i2})$.

A self-serving bias would imply a positive correlation between the gain and the relative change in a posteriori probabilities. Figure 2 plots these for each individual that misreported. If there were a self-serving bias in the sense that people acted on the fairness ideal that benefited them most, we should observe a positive correlation between the relative gain and the relative change in a posteriori probability. However, we observe that there is no such systematic relationship in Figure 2. We therefore conclude that there is no strong evidence of a self-serving bias among the misreporting individuals.

[Figure 2 about here.]

5.2 Self-serving bias and the rate of return

Another interesting test of self-serving bias is to consider whether individuals' perception of fairness is affected by the rate of return they have been assigned. Is it the case that the individuals who were assigned a high rate of return tend to act on the libertarian fairness ideal (which would provide a moral justification for keeping the benefits from a high price), and that the individuals who were assigned a low rate of return tend to act on the strict egalitarian fairness ideal?

[Table 6 about here.]

Table 6 compares the a posteriori probabilities for the high rate of return group and the low rate of return group. For both treatments, the average probability of being a libertarian is higher for the high return group than for the low return group. Similarly, there is a movement in the opposite direction for the probability of being a strict egalitarian. Considering the sample sizes, however, this effect is not particularly large. Hence, neither of our tests provide strong evidence of self-serving bias, which confirms simpler tests performed in Cappelen et al. (2007).

We also note that the difference between the high rate of return and the low rate of return group is smaller in the E-treatment than in the B-treatment. Although the evidence is weak, this may indicate that the introduction of moral reflection makes people even less susceptible to self-serving bias.[9]

6 Concluding remarks

Our findings suggest that moral reflection has a significant impact on individual behavior. It increases the weight attached to fairness in distributive choices,

[9]Haisley and Weber (2005) reach a similar conclusion, although in a different setting.

and it changes people's perception of what is a fair distribution. While we find that strict egalitarianism is the most prevalent fairness ideal in the base treatment, it is the least prevalent fairness ideal in the elicitation treatment where the participants had to reflect in advance on their view of a fair distribution. One interpretation of this finding is that the strict egalitarian fairness ideal corresponds to a widely adopted heuristic decision rule that is acceptable in many social situations, but reflecting on the fairness issue makes some people consider such a rule not justifiable in the context of this experiment.

Studies of social preferences have employed both questionnaires and experiments. Our design allows us to compare the relative informational value of the data from these two approaches. We find that a majority of the participants do not misreport when asked a hypothetical question about fairness, and hence, self-reported data have substantial informational value. However, we show that a specification based purely on behavioral data performs better than a specification that incorporates self-reported data.

Finally, we find no strong evidence of self-serving bias among the participants. In particular, it does not seem to be the main explanation for misreporting. An alternative explanation might simply be that some people do not put much thought into their answers when there is no monetary incentive. This provides support for the experimental approach of introducing real stakes in the study of social preferences.

References

Amiel, Yoram and Frank A. Cowell (1992). "Measurement of income inequality: Experimental test by questionnaire", *Journal of Public Economics*, 47(1): 3–26.

Berkelaar, Michel, Kjell Eikland, and Peter Notebaert (2007). "lp_solve: Open source (mixed-integer) linear programming system", `http://lpsolve.sourceforge.net/5.5/`.

Berndt, Ernst R., Bronwyn H. Hall, Robert. E. Hall, and Jerry A. Hausman (1974). "Estimation and inference in nonlinear structural models", *Annals of Economic and Social Measurement*, 3(4): 653–665.

Camerer, Colin F. (2003). *Behavioral Game Theory: Experiments in Strategic Interaction*, Princeton, NJ: Princeton University Press.

Cappelen, Alexander W., Astri Drange Hole, Erik Ø. Sørensen, and Bertil Tungodden (2007). "The pluralism of fairness ideals: An experimental approach", *American Economic Review*, 97(3): 818–827.

Cherry, Todd L., Peter Frykblom, and Jason F. Shogren (2002). "Hardnose the dictator", *American Economic Review*, 92(4): 1218–1221.

Dana, Jason, Roberto A. Weber, and Jason Xi Kuang (2007). "Exploiting moral wriggle room: Experiments demonstrating an illusory preference for fairness", *Economic Theory*, 33(1): 67–80.

Ferrall, Christopher (2005). "Solving finite mixture models: Efficient computation in economics under serial and parallel execution", *Computational Economics*, 25(4): 343–379.

Frohlich, Norman, Joe Oppenheimer, and Anja Kurki (2004). "Modeling other-regarding preferences and an experimental test", *Public Choice*, 119(1-2): 91–117.

Gaertner, Wulf (1994). "Distributive justice: Theoretical foundations and empirical findings", *European Economic Review*, 38(3-4): 711–720.

Gaertner, Wulf and Lars Schwettmann (2007). "Equity, responsibility and the cultural dimension", *Economica*, 74(296): 627–649.

Haisley, Emily C. and Roberto A. Weber (2005). "Favorable interpretations of ambiguity and unstable preferences for fairness", mimeo, Carnegie Mellon University.

Konow, James (2000). "Fair shares: Accountability and cognitive dissonance in allocation decisions", *American Economic Review*, 90(4): 1072–1091.

McLachlan, Geoffrey and David Peel (2000). *Finite Mixture Models*, Wiley.

Messick, David M. and Keith Sentis (1983). "Fairness, preference and fairness biases", in "Equity Theory: Psychological and Sociological Perspectives", (edited by David M. Messick and Karen S. Cook), New York: Praeger, (pp. 61–94).

Schokkaert, Erik and Kurt Devooght (2003). "Responsibility-sensitive fair compensation in different cultures", *Social Choice and Welfare*, 21(2): 207–242.

Schokkaert, Erik and Leo Lagrou (1983). "An empirical approach to distributive justice", *Journal of Public Economics*, 21(1): 33–52.

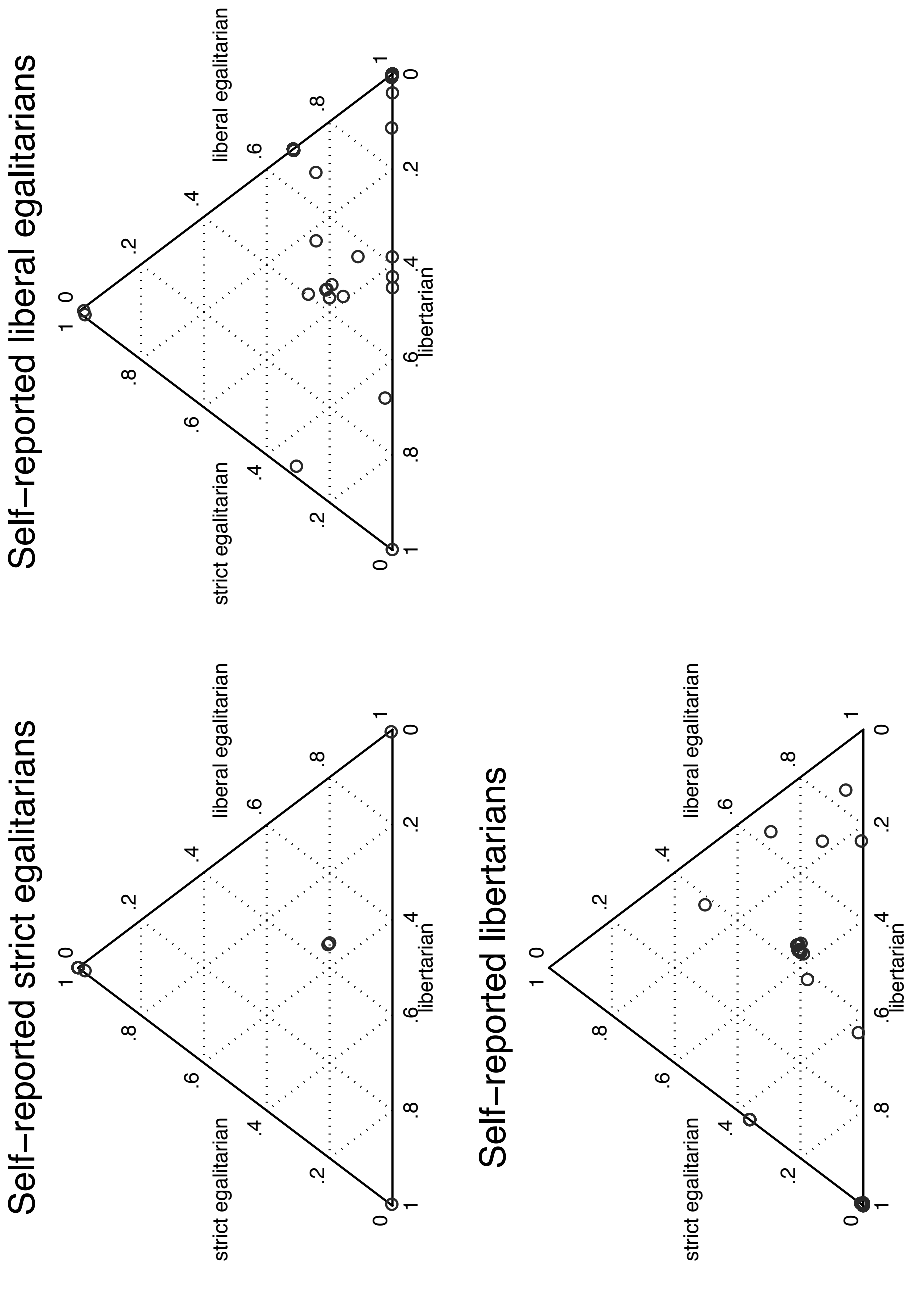

Figure 1: The distribution of a posteriori probabilities by self-reported fairness ideals

Note: The a posteriori probabilities are based on equation (2) and the estimates in Table 2.

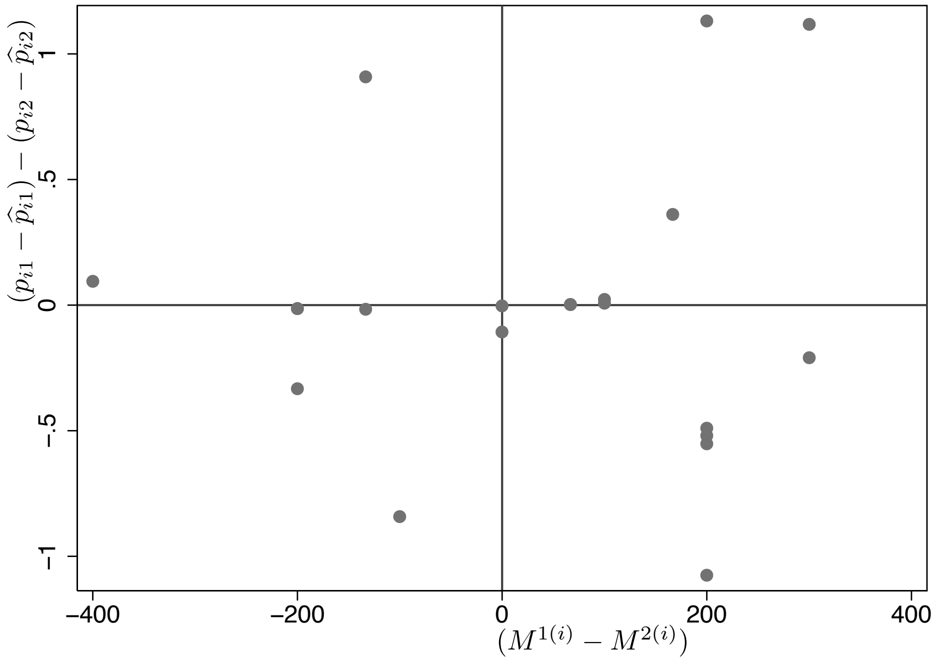

Figure 2: The extent of self-serving bias among the misreporting individuals
Note: For those individuals in the E-treatment with a posteriori probabilities of their self-reported fairness ideal more than one percentage point less than the estimated population share of this fairness ideal.

Table 1: Descriptive statistics of offers made to opponent

	B-treatment	E-treatment
mean	0.271	0.346
median	0.292	0.500
standard deviation	0.219	0.235
minimum	0	0
maximum	0.75	1.0
share demanding everything	0.305	0.255

Note: Offers are reported as the share of total production given to the opponent.

Table 2: Estimates of the random utility model

	B-treatment	E-treatment	$B+E$-treatments
Share strict egalitarian	0.435	0.202	0.301
	(0.090)	(0.066)	(0.055)
Share liberal egalitarian	0.381	0.440	0.417
	(0.088)	(0.074)	(0.057)
Share libertarian	0.184	0.358	0.283
	(0.066)	(0.068)	(0.049)
γ, marginal utility of money	28.359	31.373	28.503
	(3.589)	(4.915)	(2.913)
ζ, mean of $\log(\beta)$	5.385	7.665	5.997
	(0.349)	(0.749)	(0.389)
σ, standard deviation of $\log(\beta)$	3.371	5.887	4.374
	(0.530)	(1.467)	(0.601)
Log likelihood	-337.58	-261.68	-610.26

Note: Standard errors (in parentheses) are calculated using the BHHH method (Berndt, Hall, Hall, and Hausman, 1974). Money is scaled in units of 1000 NOK. One population share and its standard error is calculated residually. The likelihood is maximized using the FmOpt library (Ferrall, 2005).

Table 3: Estimates and self-reported data in E-treatment

	estimates	self-reported
Share strict egalitarian	0.202	0.098
	(0.066)	(0.031)
Share liberal egalitarian	0.440	0.467
	(0.074)	(0.052)
Share libertarian	0.358	0.435
	(0.068)	(0.051)

Note: The first column is taken from the second column of Table 2. For the self-reported data, standard errors (in parentheses) are calculated based on binomial outcomes.

Table 4: Estimates using individual classifications, E-treatment

	self-reported	Bayes rule	Bayes rule (restricted)
γ, marginal utility of money	17.857	34.354	37.048
	(2.301)	(4.189)	(5.555)
ζ, mean of $\log(\beta)$	7.049	7.861	7.420
	(0.808)	(0.741)	(0.656)
σ, standard deviation of $\log(\beta)$	6.320	5.800	5.684
	(1.601)	(1.261)	(1.296)
Log likelihood	-240.41	-207.59	-206.98

Note: The numbers are calculated using the restriction that $k(i)$ is equal to the self-reported ideal or the Bayesian classifications proposed in (3). Standard errors (in parentheses) are calculated using the BHHH method (Berndt et al., 1974) and are not corrected for the first-step estimation to do the classification. Money is scaled in units of 1000 NOK.

Table 5: Behavior versus self-reported fairness ideal

classification status	Self-reported ideal SE	LE	L	total
Decreased	22.2%	14.0%	20.0%	17.4%
No change	22.2%	11.6%	20.0%	16.3%
Increased	55.6%	74.4%	60.0%	66.3%
total	100%	100%	100%	100%
n	9	43	40	92

Note: Decreased: a posteriori probability of the self-reported ideal is less than the population estimate minus one percentage point. No change: a posteriori probability of the self-reported ideal is within one percentage point of the population estimate. Increased: a posteriori probability of the self-reported ideal is more than one percentage point larger than the population estimate.

Table 6: Averages of a posteriori type probabilities, by treatment and rate of return

	B-treatment		E-treatment	
average of	$a_1 = 2$	$a_1 = 4$	$a_1 = 2$	$a_1 = 4$
$P(SE\|Y_i, Z_i)$	0.458	0.410	0.215	0.190
$P(LE\|Y_i, Z_i)$	0.404	0.360	0.447	0.433
$P(L\|Y_i, Z_i)$	0.138	0.231	0.339	0.377

Note: The a posteriori probabilities are based on the estimates in Table 2.

Norges
Handelshøyskole

Norwegian School of Economics
and Business Administration

NHH
Helleveien 30
NO-5045 Bergen
Norway

Tlf/Tel: +47 55 95 90 00
Faks/Fax: +47 55 95 91 00
nhh.postmottak@nhh.no
www.nhh.no