The pluralism of fairness ideals: An experimental approach

Alexander W. Cappelen Astri D. Hole
Erik Ø. Sørensen Bertil Tungodden*

2nd September 2005

Discussion Paper 12/2005

Abstract

A core question in the contemporary debate on distributive justice is how the fair distribution of income is affected by differences in talent and effort. Important theories of distributive justice, such as strict egalitarianism, liberal egalitarianism and libertarianism, all give different answers to this question. This paper presents the results from a version of the dictator game where the distribution phase is preceded by a production phase. Each player's contribution is a result of an exogenously given talent and a chosen effort. We estimate simultaneously the prevalence of three main principles of distributive justice among the players as well as the distribution of weights they attach to fairness considerations.

*Cappelen: The University of Oslo and Norwegian School of Economics and Business Administration, Bergen, Norway. e-mail: alexander.cappelen@nhh.no. Hole: Bergen University College, Norway. e-mail: astri.drange.hole@hib.no. Sørensen: Norwegian School of Economics and Business Administration, Bergen, Norway. e-mail: erik.sorensen@nhh.no. Tungodden: Norwegian School of Economics and Business Administration and Chr. Michelsen Institute, Bergen, Norway. e-mail: bertil.tungodden@nhh.no. We would like to thank Chris Ferrall and Steinar Holden for helpful comments.

1 Introduction

People are motivated by fairness considerations and are willing to sacrifice pecuniary gains in order to avoid large deviations from what they consider a fair solution. This type of behaviour has been extensively documented in laboratory experiments with games such as the ultimatum game and the dictator game (Camerer, 2003). However, while these games show us that a substantial fraction of the players are motivated by fairness considerations, they do not provide much information on the pluralism of fairness ideals present in society. In the standard versions of the ultimatum game and the dictator game, the money to be distributed by the players is essentially "manna from heaven", and it seems rather uncontroversial to assume that people in general view the fair solution to be to distribute money equally in these cases.

The core question in both the modern political debate on distributive justice and in normative theoretical reasoning, however, is how to understand fairness in more complex situations involving production. In particular, there is substantial disagreement about the extent to which people should be held responsible for various factors affecting their pre-tax income. The controversy between the left wing and the right wing of the political spectrum can be interpreted to a large extent as a disagreement about how differences in effort and talent should be allowed to affect the income distribution. The prevalence of the view that luck determines income in a society seems to play an important role in explaining cross-country variation in choices of re-distributive policies (Alesina and Angeletos, 2004).

Three fairness ideals are prominent in this debate. Proponents of the *strict egalitarian* doctrine argue that people should not at all be held responsible for their effort and talent, thus considering equal sharing as the fair solution even in cases involving production. *Libertarians*, on the other hand, claim that people should be held responsible for both their talent and their effort, which implies that the fair solution is to give each person what she produces. As an intermediate position, *liberal egalitarians* view effort as within and talent beyond individual control, and thus believe that redistributive policies should aim at equalising differences due to differences in talent but should allow for inequalities due to differences in effort.

Which of these fairness ideals is more prevalent in society? This question is not easily answered, because in actual behaviour, fairness considerations are usually balanced against self-interest considerations. Differences in observed behaviour therefore may be due to two different sources. People may differ both in the importance they assign to fairness considerations and with respect to what they consider to be a fair distribution. As a result, the most common ways to elicit data on the prevalence of different fairness ideals have been to use surveys or experiments where the proposer is not a stake-holder, thereby avoiding any self-serving bias (see Konow (2003) for an overview of this literature). However, these approaches have the weakness that the participants do not have to demonstrate any willingness to act on the endorsed fairness ideals, and consequently they can be very sensitive to framing effects.

The aim of this paper is to show how one may estimate simultaneously the prevalence of different fairness ideals and the degree of importance people attach to fairness considerations in an experiment where participants have a stake in the outcome. We study a dictator game in which the distribution phase is preceded by a production phase. The players differ with respect to both effort and talent, and thus different fairness ideals provide different answers to the question of what is a fair distribution of the total production. Given a simple random utility model where people make a trade-off between pecuniary gains and fairness considerations when proposing a distribution of the production, we estimate the share of the population motivated by each of the three fairness ideals (strict egalitarianism, liberal egalitarianism and libertarianism) and the mean value and variance in the parameter measuring the importance people attach to fairness considerations. We also provide a simple test of whether there is "moral wriggling" among the participants, that is, whether they decide opportunistically on a fairness ideal after the distributional situation is known (Dana, Weber and Kuang, 2004).

Section 2 describes the basic model in more detail, including the fairness ideals. Section 3 provides a discussion of the experimental design, and the results are reported in Section 4. Section 5 contains a discussion of related literature and some concluding comments.

2 The model

We study a situation in which individuals differ in both effort and talent, and in which effort is clearly within individual control, whereas talent is clearly beyond individual control. Effort, q_i, is the amount of money an individual i chooses to invest in the production phase. Talent, a_i, is the rate of return on the investment. The income generated by individual i in the production phase is then $x_i = a_i q_i$. The experiment is designed such that there is no need to model the choice of effort in the production phase.

The distribution phase will always be in a two-person setting, where we refer to the individuals as person 1 and person 2. The total income to be distributed is given by $X(\mathbf{a}, \mathbf{q}) = x_1(a_1, q_1) + x_2(a_2, q_2)$, where $\mathbf{a} = (a_1, a_2)$ and $\mathbf{q} = (q_1, q_2)$. Each individual is to propose an amount of income y for herself and $X - y$ for her opponent.

2.1 Individual motivation: income and fairness

We assume that the individuals are motivated by both a desire for income and a fairness ideal, where individual i's fairness ideal is denoted $m^{k(i)}$ and specifies a unique distribution in any given situation. We also assume that the marginal disutility of deviating from the fairness ideal is increasing in the size of the deviation from the fair distribution. More formally, we assume that person i is maximising the following utility function when proposing a distribution (this is a generalisation of the utility function studied by Bolton and Ockenfels (2000)),

$$V_i(y; \mathbf{a}, \mathbf{q}) = \gamma y - \frac{\beta_i}{2} \left((y - m^{k(i)}(\mathbf{a}, \mathbf{q})\right)^2, \tag{1}$$

where the parameters $\gamma > 0$ and $\beta_i \geq 0$ determine the weight individual i gives to income and to fairness considerations. The optimal proposal, y^*, is (given an interior solution)

$$y^* = m^{k(i)}(\mathbf{a}, \mathbf{q}) + \gamma/\beta_i. \tag{2}$$

It follows immediately that the optimal proposal depends on both the fairness ideal endorsed by the individual and the importance assigned to fairness considerations.

A player with $\beta_i = 0$ would always keep all the money for herself.

2.2 The fairness ideals

We assume that an individual endorses some version of strict egalitarianism, libertarianism or liberal egalitarianism. Each of the fairness ideals satisfies the no-waste condition, and thus we can index the fair distribution such that m^k and $X - m^k$ is what fairness ideal k assigns to person 1 and person 2 respectively.

Strict egalitarians do not hold people responsible for their effort and talent, and therefore they view equal sharing as the fair distribution. This fairness ideal may be interpreted in two different ways in the present context. First, one may defend the simplest and strongest notion of equality (see, for example, Nielsen (1985)), where fairness is to distribute gross total income equally. We call this the strong version of strict egalitarianism (SES).

$$m^{SES}(\mathbf{a}, \mathbf{q}) = X(\mathbf{a}, \mathbf{q})/2. \tag{3}$$

Alternatively, one may interpret strict egalitarianism as equally distributing the net total income, which implies that the two persons receive the same overall income from the game (see also Iversen, Jackson, Kebede, Munro and Verschoor (2005)). We call this the weak version of strict egalitarianism (SEW).

$$m^{SEW}(\mathbf{a}, \mathbf{q}) = q_1 + (X(\mathbf{a}, \mathbf{q}) - q_1 - q_2)/2. \tag{4}$$

The strict egalitarian view is closely related to the inequality-aversion models in the experimental literature, which assume that people dislike inequitable outcomes (see Fehr and Schmidt (1999) and Frohlich, Oppenheimer and Kurki (2004). The weaker version may also be given a welfarist interpretation. If we assume that the fair distribution is what maximises a quasi-concave social welfare function, that individuals derive the same welfare from income and that marginal welfare decreases with income, then it follows that the fair solution is to distribute the net total income equally.

The libertarian fairness ideal is at the opposite extreme of strict egalitarianism. The fair distribution is simply to give each person exactly what she produces,

$$m^L(\mathbf{a}, \mathbf{q}) = a_1 q_1. \tag{5}$$

This view may be defended by arguing that people should be held responsible for both their effort and their talent, and hence that a low talent does not justify any redistribution among individuals (Nozick, 1974). The fair solution may thus involve an unequal distribution of income due to differences in both effort and talent.

Liberal egalitarianism, on the other hand, defends the view that people should only be held responsible for their choices (Roemer, 1998). A reasonable interpretation of this fairness ideal in the present context is to view the fair distribution as giving each person a share of the total income equal to her share of the total effort.

$$m^{LE}(\mathbf{a}, \mathbf{q}) = \frac{q_1}{q_1 + q_2} X(\mathbf{a}, \mathbf{q}). \tag{6}$$

This principle is equivalent to what has been described as the accountability principle (Konow, 1996, 2000). It implies that if two persons make the same choice, then the fair solution is to give them the same income. If they make different choices, the liberal egalitarian fairness ideal justifies an unequal distribution of income between them.

Even though these fairness ideals provide different solutions to the distributional problem, it is important to note that on average they instruct individuals to offer the same amount to the other person. In any particular game and for any fairness ideal k, the fair solution would be for person 1 to offer $X - m^k$ to person 2 and for person 2 to offer m^k to person 1, which implies that the average fair offer in the game is $X/2$. Hence, it is not possible to extract any information about the prevalence of the various fairness ideals from the size of the average offer. In order to establish such information, we need to study how each individual's offer depends on the distribution of effort and talent in the situation.

3 Experimental design

Our experiment is a version of the dictator game with production, where production is dependent on both factors within and factors beyond individual control. At the beginning of the experiment, each participant was given money credits equal to 300 Norwegian Krone (NOK), approximately 50 USD, and informed about the rules of the game.[1] Each participant was then 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, while those who were assigned a high rate of return would quadruple their investment.

In the production phase the participants were asked to determine how much they wanted to invest in two different games. Their choice alternatives

[1]The complete instructions are available on request from the authors.

were limited to 0, 100 and 200 NOK, and the total amount invested in the two games could not exceed the initial money credit they received. Any money they chose not to invest they could keep after the experiment ended, and thus they faced a genuine choice of investment.

In the distribution phase, the participants were paired with a player who had the same rate of return in one game and with a player who had a different rate of return in another game. In each game, they were given information about the other participant's rate of return, investment level and total contribution and were then asked to propose a distribution of the total income. The participants were not informed about the outcome of the first game before the second game was completed. For each participant, one of the two games and one of the two proposals in that game (the participant's own that of the opponent) were randomly selected to determine the final outcome. The total earnings from the experiment for a participant were then the sum of the final outcome and the amount of money not invested.

At the end of the experiment, the participants were assigned a code and instructed to mail the code and the bank account numbers to the accounting division of the Institute for Research in Economics and Business Administration (SNF). Independently, the research team mailed a list with the codes and total payment to the accounting division, who then disbursed the earnings directly to the participants' bank accounts. This procedure ensured that neither the participants nor the research team was in a position to identify how much each participant earned in the experiment.

The participants in the experiment were all recruited among the first-year students at the Norwegian School of Economics and Business Administration. They were not informed about the purpose of the experiment but were only invited to take part in a research project. In the invitation, they were told that they would initially receive 300 NOK for use in an experiment that would last for about 40 minutes and that their total earnings from the experiment would depend on their choices. The hourly opportunity cost for most of these students would be about 100 NOK, while the average payout was 447 NOK. Each student was only permitted to participate once. We had one session with 20 participants, one session with 12, and four sessions with 16, comprising a to total of 96 participants. The participants were in the same computer lab during a session, but all communication was anonymous and was conducted through a web-based interface.

In Table 1, we see the distribution of investments in the first and the second game. No one kept the full endowment, one participant (with a low rate of return) invested only 100 NOK and 10 participants (four with a high rate of return and six with a low rate of return) invested 200 NOK. The remaining 85 participants invested the full endowment of 300 NOK, evenly distributed between investing $(200, 100)$ and $(100, 200)$. The fact that some participants did not invest the full endowment indicates that they perceived the choice of investment as a genuine choice. However, since most did invest the full amount, we doubt that the variation in choices in the production phase introduces any important bias in our analysis of the distribution phase.

[Table 1 about here.]

In the distribution phase, the paired players could differ with respect both to their rate of return and their investment, which implies that there were four different classes of distributional situations in the experiment. First, there were situations where the players were identical with respect to both their rate of return and their investment. All the four fairness ideals imply the same fair distribution in this case, namely that both players get an equal share of the total income. Second, there were situations where the players had the same rate of return but differed in their investment level. This would make the liberal egalitarian and the libertarian fairness ideal coincide, whereas the two versions of strict egalitarianism would imply different views of the fair distribution. Third, there were situations where the players had made the same investment but differed in their rate of return. All the fairness ideals except for libertarianism consider an equal distribution fair in such a case. Finally, there were situations where the players differed along both dimensions. In these situations, the strong version of strict egalitarianism and libertarianism imply the same fair offer if the player with the high talent is the player with the low effort. Otherwise, all the fairness ideals differ in this case. Table 2 reports the empirical distribution of the four classes of distributional situations in the experiment.

[Table 2 about here.]

As we can see from Table 2, there was almost a balanced design with respect to the four distributional situations.[2] We have 44 situations where the prevalence of different fairness ideals cannot influence the distribution of offers made. In the remaining observations, the differences in observed behaviour may be due to the fact that people endorse different fairness ideals. In order to get a clearer view of the potential variation caused by the prevalence of different fairness ideals, we present in Figure 1 pair-wise scatter plots of how the various fairness ideals correlate to each other in all the distributional situations in the experiment. If two fairness ideals coincide for all the distributional situations, then all the points should be at the diagonal in the respective comparison. Figure 2 shows that the fairness ideals imply considerable variation for the distributional situations in the experiment, possibly with the exception of the two versions of strict egalitarianism. These overlap to a great extent, and thus we should not expect to gain much by including both in the empirical analysis of this experiment.

[Figure 1 about here.]

[2] There are 190, not 192, distributional situations in total, since a single incidence of a software problem caused a pair of participants to enter invalid data in one distributional situation. This pair was dropped from all further analysis.

4 Results

We begin by presenting some descriptive statistics before we formulating and estimating a random utility model. Finally, we consider the possibility of "moral wriggling" by the participants.

4.1 Descriptive statistics

[Table 3 about here.]

Table 3 summarises some main features of the offers made. The average offer to the opponent is 27.1% (which amounts to 229 NOK), while the median is 29.2%. This is slightly higher than what is commonly observed in the standard dictator games without production (Andreoni and Miller, 2002; Camerer, 2003), and may indicate that the presence of a production phase causes people to care more about fairness considerations. The maximum offer is of 75% of the total income.

[Table 4 about here.]

Table 4 contains the full distribution of offers made. We see that there are marked steps in the distribution. In fact, out of 190 proposed distributions, 184 are of even 100 NOK amounts. The remaining six proposals are of even 50 NOK amounts. While 31% of the offers leave the opponent with nothing, some offer substantial amounts; 20 out of 190 offers are NOK 600 (about USD 100) or above. 27% of the offers are exactly fifty-fifty (not reported in table).

[Table 5 about here.]

In Table 5 we present some descriptive regressions. We see from the first regression that the participants demand almost all of their own production

(a_1q_1) but only two-thirds of the opponent's production (a_2q_2). This difference is statistically significant. Hence, in the distribution phase, it seems to matter who contributes to the production of the total income. The second and third regressions show that it also matters how the contribution came about. The participants seem to take more of the opponent's production if this is due to a high rate of return than if it is due to a high investment. This is consistent with the hypothesis that there are individuals who care about the distinction between effort and talent, but in itself these regressions are not very informative about individual preferences.

The model outlined in Section 2 implies that there might be identification from the fact that no one should ever offer *more* than what is implied by their fairness ideal. If one observes an offer of more than what is implied by a fairness ideal, then the model rules out the possibility that this person is motivated by this particular fairness ideal. 75 out of the 96 participants, however, demand more than what is implied by *all* the fairness ideals, and hence our experimental data are not well suited to a revealed-preference approach to the identification of the prevalence of different fairness ideals. Moreover, it turns out that two out of the 96 participants demand less than what is implied by all the fairness ideals. Estimation of any model that does not allow for some smoothing of choices will therefore fail.

4.2 Empirical model

We adapt the model to bring it into line with two features in the experimental data. First, given that the participants have a very strong tendency to choose round numbers, we restrict the choice of y to the set $\mathbf{Y}(\mathbf{a}, \mathbf{q}) = \{0, 50, 100, \ldots, X(\mathbf{a}, \mathbf{q})\}$. Second, we introduce random variation that is idiosyncratic to each choice. Given the utility function V defined in (1), we introduce the random utility model

$$U_i(y; \cdot) = V_i(y; \cdot) + \varepsilon_y. \tag{7}$$

We assume that the ε_y's are i.i.d. extreme value distributed, and that individuals choose a y, call it y^*, such that $U_i(y^*; \cdot) \geq U_i(y; \cdot)$ for all y in $\mathbf{Y}$.[3]

The model we propose has a mixed logit structure where each person is characterised by their fairness ideal, $k(i)$, as well as the parameter β_i determining the importance a person assigns to fairness considerations. We cannot classify individuals by $(k(i), \beta_i)$, but we estimate the distribution of these characteristics. The distribution of moral types is discrete in nature, and we approximate the distribution of β by a log-normal distribution, such that $\log \beta \sim N(\zeta, \sigma^2)$. Since the fairness ideal and the importance a person assigns to fairness considerations are unobserved by us, these must be integrated out for the unconditional choice probabilities as functions of the observed variables. We provide the likelihood function in an appendix.

Formal proofs of identification are difficult to provide in this situation where there is a large (but discrete) set of outcomes. However, consider what can be learned from the situations where $a_1 = a_2$ and $q_1 = q_2$. In these situations all fairness ideals coincide at $X/2$. The mean offer in these situations reflect the mean weight given to fairness considerations. The variance of offers in these situations reflect both the distribution of β and the smoothing introduced by the extreme value distributed ε's. There is, however, also a discontinuity in the design, in that all offers above $X/2$ must result from the smoothing alone. With the parametric assumption of log-normality of $f(\beta)$, these situations provide information about (γ, ζ, σ). Repeated observations, and the fact that we expose individuals to very different distributional situations, provide information about the distribution of moral ideals and further precision about the distribution of β.

[3] The random utility structure of discrete offers made in our empirical model is similar to that of Andreoni, Castillo and Petrie (2004), but our model is estimated on the full population, and we do not estimate individual-specific utility functions.

4.3 Structural estimates

In Table 6, we present the estimates of the structural model. Column 1 presents the structural estimates with all the fairness ideals, including both the weak and the strong version of strict egalitarianism. Columns 2–5 drop one of the fairness ideals in turn. In all columns, the estimate for each of the fairness ideals is the share of the participants who are motivated by this particular fairness ideal.

[Table 6 about here.]

From the different specifications 1–5, and as we could expect from Figure 1, we see that the strong and the weak version of strict egalitarianism are not well separated in our data. Neither the log-likelihood nor the other parameters are much affected in specification 3 where the weak version of strict egalitarianism is excluded. There are, however, large effects of dropping any of the other fairness ideals. Specification 3, in which we have 39.7% strict egalitarians, 43.4% liberal egalitarians and 16.8% libertarians, is therefore our preferred specification.

Based on these estimates, we make three observations. First, there is considerable pluralism in the fairness ideals that motivate people, even in rather simple distributional situations involving a homogeneous group of students. Second, the majority of the participants (the liberal egalitarians and the libertarians) care about the investments made by the opponent when they decide how much to offer. This implies that fairness considerations cannot be reduced to income inequality aversion in these distributional situations. Third, the estimated share of strict egalitarians is larger than the share of offers that are fifty-fifty. This is due to the fact that the fairness ideals overlap in some distributional situations and that people make active trade-offs.

The distribution of the parameter β, which determines the importance that people attach to fairness considerations, is assumed to be log-normal and characterised by parameters (ζ, σ), while the parameter γ determines the weight given to deterministic utility relative to the smoothing implied from the extreme value distributed ε's.[4] To get a handle on the effect of our estimated parameters, we provide Figure 2. This figure takes as the point of departure a situation where the total production is 1000 and the fairness ideal endorsed by a hypothetical individual specifies an equal split. We then provide, for every inner decile of the distribution of β, the deterministic utility and, plotted as solid bars, the implied choice probabilities for all even 50 NOK amounts for this hypothetical individual. By way of illustration, consider the case where $CDF(\beta) = 0.5$. The deterministic part of the utility function reaches its maximum when the individual offers 350 NOK, and thus the individual makes an active trade-off between fairness and self-interest considerations. The smoothing, however, implies that there is a positive but small probability of observing such a person offering more than what is considered just by the fairness ideal she endorses (as seen by the small mass to the left of the fairness ideal).

[Figure 2 about here.]

Our general impression from Figure 2 is that the population can be divided into three main groups. About 30% of the participants assign so little importance to fairness considerations that they have no inner maximum in their choice problem. Thus the most common choice among them is to offer the opponent nothing. 40% of the participants make active trade-offs between fairness and self-interest considerations, whereas 30% of the participants care mainly about fairness considerations.

To see how well our estimates predict the actual distribution of offers, we simulate a distribution of offers for the distributional situations in the experiment. As we can see from Figure 3, there is a close fit. In particular, we note that we fit the large mass at the two most distinct points in the distribution (offers of 0% and of 50%). At the ends of the support, the

[4] The model is normalised by the constant variance of ε_i, which is $\pi^2/6$.

smoothing can only operate one way, and hence we slightly underpredict the number of proposals that offer nothing, and we slightly overpredict the number of very high offers. This is to be expected given the random utility structure of the model.

[Figure 3 about here.]

4.4 True pluralism or moral wriggling?

We have assumed that individuals have a fairness ideal that is independent of the distributional situation in which they find themselves. Alternative approaches emphasise self-deception (Konow, 2000) or “moral wriggling” (Dana et al., 2004), where the idea is that individuals may use ambiguity in the distributional situation to further their own pecuniary self-interest at the expense of fairness. In a setting such as the one we are examining in this paper, a natural application of this train of thought is to allow for the possibility that people have no firm view about the fairness ideal to which they should adhere, and that they choose opportunistically the fairness ideal that benefits them most in any particular distributional situation.

In distributional situations where the rate of return and the investment level is the same for the two participants, all the fairness ideals defend an equal sharing of the outcome. Hence, moral wriggling is only applicable in situations where there is some inequality in either the rate of return or the investment. Even in these situations, every fairness ideal has an average offer of 50%. In Table 7, however, we see in the second column that in the ambiguous situations, choosing a fairness ideal self-interestedly would on average justify increasing one’s own share of the total income with 9.3 percentage points. A simple test of the idea of moral wriggling is therefore to see whether the participants consistently ask for a larger share in distributional situations where there is scope for such moral wriggling. In the third column, we see that there is indeed a difference of 2.2 percentage points between the actual amount demanded in the non-ambiguous and ambiguous situations, but this difference is small and not statistically significant. We conclude from this that while we cannot rule out that some individuals exploit such scope for moral wriggling, there is little reason to suspect that this is pervasive to a degree that would invalidate our analysis.

[Table 7 about here.]

5 Concluding remarks

Our analysis relates to the interesting studies of Konow (2000) and of Frohlich et al. (2004), which also apply versions of the dictator game with production in order to analyse the role of fairness considerations in individual choices. In line with our findings, both studies find that the distinction between effort and talent matters for many people. At the same time, there are important differences between these studies and ours.

The focus of Konow (2000) is to examine the extent to which fairness considerations can be explained by a single fairness ideal, namely the liberal egalitarian principle. In contrast, our aim has been to examine the prevalence of different fairness ideals among the participants, including liberal egalitarianism as one possibility. Moreover, even though liberal egalitarianism turns out to be the most prevalent fairness ideal among our participants, the majority of them hold other fairness ideals.

Frohlich et al. (2004) share our focus on the pluralism of fairness ideals, and they also find that there is substantial heterogeneity in their group of participants. They study this issue in an environment where it is not possible to distinguish libertarians from liberal egalitarians. More importantly, their choice model does not allow for any active trade-off between a fairness ideal and pecuniary self-interest, and thus they are unable to distinguish clearly between a fairness ideal and the weight people attach to fairness considerations. This implies that they are unable to study possible heterogeneity in the weight people attach to fairness considerations (while such heterogeneity would bias their classification procedure).

The main aim of our study has been to show how we can estimat simultaneously the degree of heterogeneity in fairness ideals and in the weight people attach to fairness considerations. It turns out that both of these kinds of heterogeneity matter in explaining individual behaviour in our experiment, but we believe that this is also true more generally. Value pluralism is a characteristic feature of modern societies, and thus it could also potentially constitute an important ingredient in the explanation of economic phenomena.

Appendix: The likelihood function

In order to take into account the fact that individuals make repeated choices, it is neccessary to introduce the notation J_i for the number of choices individual i makes. The likelihood of an individual i of type k making a proposal y_{ij} from the set of feasible proposals $\mathbf{Y}_{ij}$ given a parameter vector $\theta = (\lambda_1, \lambda_2, \lambda_3, \lambda_4, \gamma, \zeta, \sigma)$ is

$$L_{ik}(\theta) = \int_0^\infty \left(\prod_{j=1}^{J_i} \frac{e^{V^k(y_{ij};\mathbf{a}_{ij},\mathbf{q}_{ij},\beta,\gamma)}}{\sum_{s \in \mathbf{Y}_{ij}} e^{V^k(s;\mathbf{a}_{ij},\mathbf{q}_{ij},\beta,\gamma)}} \right) f(\beta;\zeta,\sigma) d\beta. \tag{8}$$

Revelt and Train (1998) calls this a “mixed logit with repeated choices”. We assume that $f(\beta;\cdot)$ is log-normal, parameterised such that $\log(\beta) \sim N(\zeta, \sigma^2)$. The total likelihood, integrating over the distribution of unobserved moral type, is a finite mixture over the type distribution determined by the discrete distribution induced by λ,

$$L_i(\theta) = \sum_{k=1}^{4} \lambda_k L_{ik}(\theta). \tag{9}$$

The estimation is with simulated maximum likelihood, with 250 random draws with antithetics for the numerical integration over the $f(\beta)$ distribution. The estimation is performed with FmOpt, Christopher Ferrall's efficient routines for finite mixture models (Ferrall, 2005).

References

Alesina, A. and G.-M. Angeletos (2004). "Fairness and redistribution." forthcoming, *American Economic Review*.

Andreoni, J., M. Castillo and R. Petrie (2004). "Revealing preferences for fairness in ultimatum bargaining." mimeo.

Andreoni, J. and J. Miller (2002). "Giving according to GARP: An experimental test of the consistency of preferences for altruism." *Econometrica*, **70**(2), 737–753.

Berndt, E. R., B. E. Hall, R. E. Hall and J. A. Hausman (1974). "Estimation and inference in nonlinear structural models." *Annals of Economic and Social Measurement*, **3**(4), 653–665.

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

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

Dana, J., R. A. Weber and J. X. Kuang (2004). "Exploiting moral wriggle room: Behavior inconsistent with a preference for fair outcomes." mimeo, Carnegie Mellon University.

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

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

Frohlich, N., J. Oppenheimer and A. Kurki (2004). "Modeling other-regarding preferences and an experimental test." *Public Choice*, **119**(1-2), 91–117.

Iversen, V., C. Jackson, B. Kebede, A. Munro and A. Verschoor (2005). "Contributions, efficiency and inequality: An experimental approach to testing economic and other theories of household behaviour in rural Uganda." mimeo, University of East Anglia.

Konow, J. (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*, **41**(4), 1188–1239.

Nielsen, K. (1985). *Equality and Liberty.* Rowman and Allanheld.

Nozick, R. (1974). *Anarchy, State and Utopia.* Basic Books.

Revelt, D. and K. Train (1998). "Mixed logit with repeated choices: Households' choices of appliance efficiency level." *Review of Economics and Statistics*, **80**(4), 647–657.

Roemer, J. E. (1998). *Equality of Opportunity.* Harvard University Press.

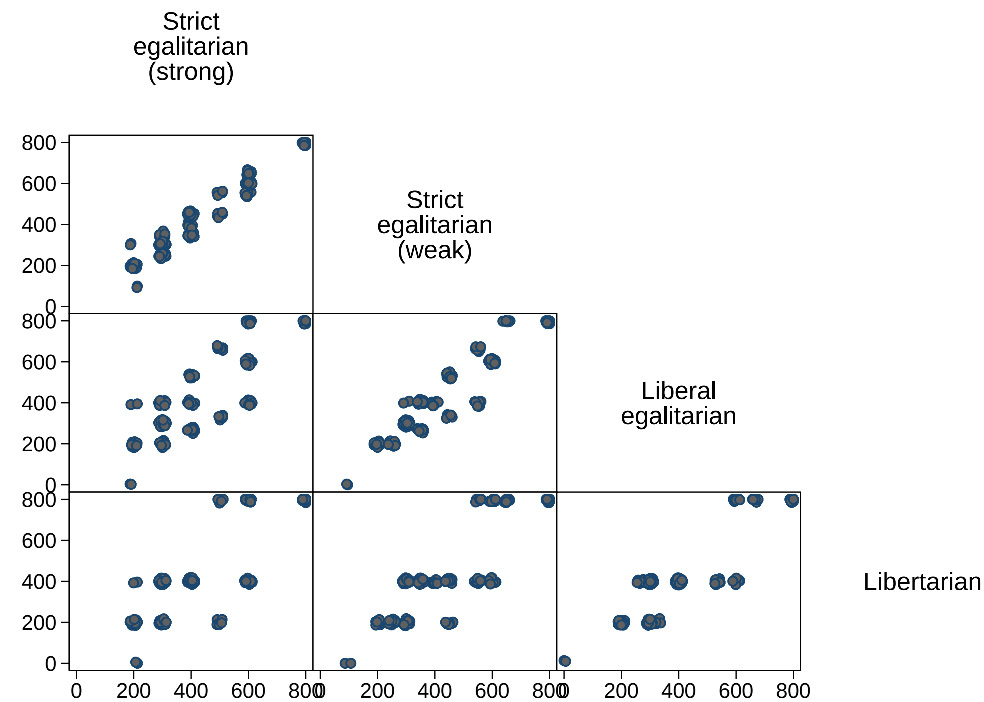

Figure 1: Scatter plots of fairness ideals. Pair-wise plots of $m^k(\mathbf{a}, \mathbf{q})$ against $m^j(\mathbf{a}, \mathbf{q})$ for all the distributional situations in our data. The weight of dots indicates the number of observations at that point.

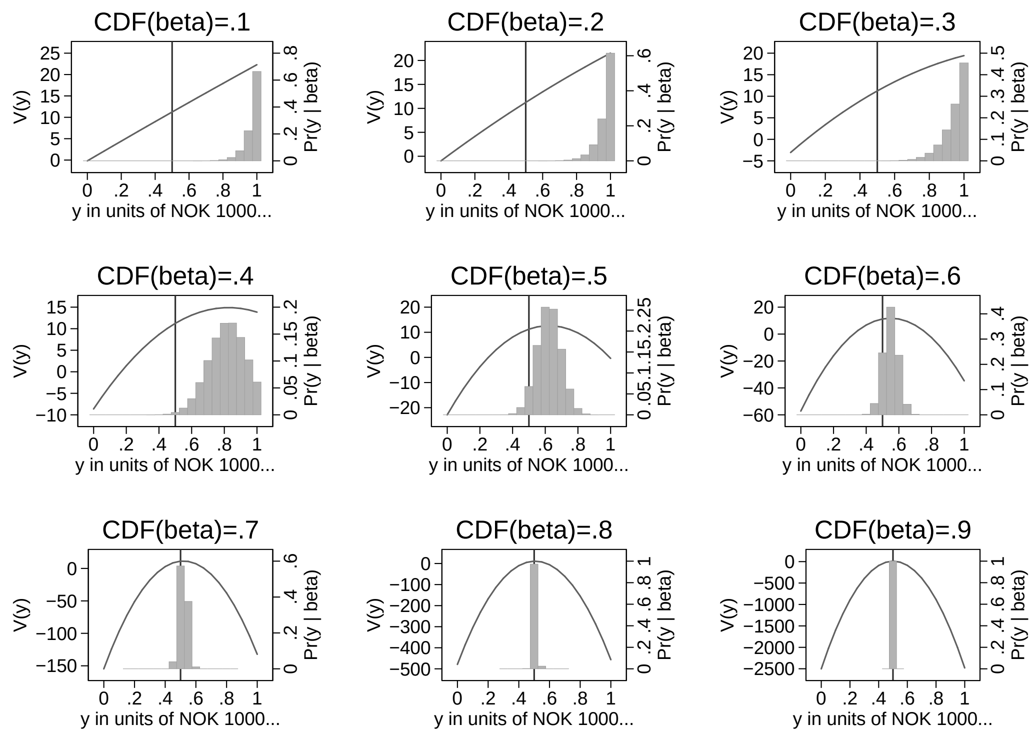

Figure 2: Deterministic utility, $V(y) = \gamma y - \beta(y - m)^2/2$, for an individual with $m = 0.5$ (marked by a vertical line). Calculated at the deciles of the estimated β distribution using the estimates in the preferred specification (3) in Table 6. Money, y, is measured in units of thousands of NOK.

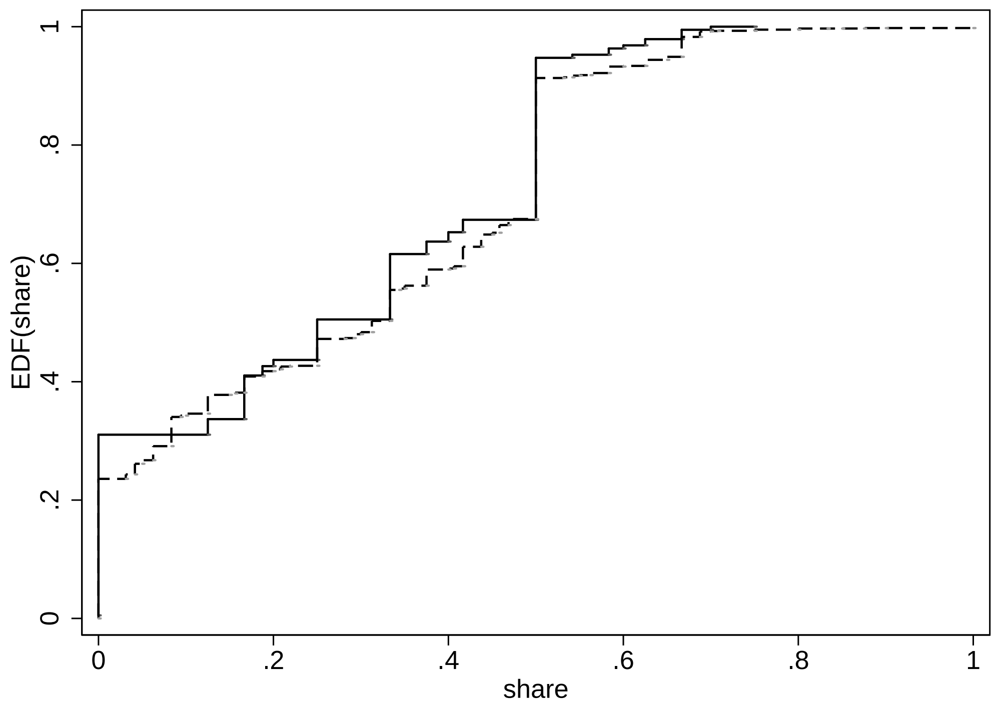

Figure 3: Empirical distribution function of offers made (as share of total production) and predictions from the estimated model. The solid line is our experimental data while the dashed line is predictions made from the estimates in specification (3). Predictions are made at the distributional situations in our dataset.

	second game			
first game	0	100	200	total
0	0	1	0	1
100	0	9	39	48
200	1	46	·	47
total	1	56	39	96

Table 1: Investments in the first and the second game.

	investment		
talent	same	different	total
same	44	50	94
different	54	42	96
total	98	92	190

Table 2: Number of observations in each of the four classes of distributional situations.

	offer	
	share	amount (in NOK)
mean	0.271	229
median	0.292	200
standard deviation	0.219	219
minimum	0	0
maximum	0.75	800

Table 3: Descriptive statistics of offers made to opponent.

offer (in NOK)	frequency	share	cumulative share
0	58	30.53	30.53
100	15	7.89	38.42
150	3	1.58	40.00
200	39	20.53	60.53
250	1	0.53	61.05
300	25	13.16	74.21
400	23	12.11	86.32
500	6	3.16	89.47
600	8	4.21	93.68
650	1	0.53	94.21
700	3	1.58	95.79
750	1	0.53	96.32
800	7	3.68	100.00

Table 4: Full distribution of offers made to opponent.

	specification			
y_1 on	1	2	3	4
constant	-56.48	-260	-289	-798
	(39.60)	(90.3)	(78.3)	(81.7)
$a_1 q_1$	0.936			
	(0.069)			
$a_2 q_2$	0.667			
	(0.069)			
a_1		157		125
		(21)		(16)
a_2		143		101
		(21)		(16)
q_1			3.62	3.07
			(0.37)	(0.31)
q_2			2.84	2.19
			(0.37)	(0.31)
R^2	0.66	0.36	0.45	0.64

Table 5: Descriptive regressions.

parameter	specification 1	2	3	4	5
λ_1, strict egalitarian, strong	0.3342 (0.0969)		0.3971 (0.0900)	0.3249 (0.1362)	0.3719 (0.1130)
λ_2, strict egalitarian, weak	0.0877 (0.0759)	0.4032 (0.1086)		0.3958 (0.1449)	0.0963 (0.0965)
λ_3, liberal egalitarian	0.4078 (0.0924)	0.3967 (0.1054)	0.4338 (0.0923)		0.5318 (0.1017)
λ_4, libertarian	0.1703 (0.0641)	0.2001 (0.0707)	0.1681 (0.0641)	0.2792 (0.0813)	
ζ	5.235 (0.487)	4.794 (0.447)	5.167 (0.475)	4.474 (0.549)	4.611 (0.455)
σ	3.710 (0.728)	3.017 (0.524)	3.703 (0.690)	3.468 (0.702)	3.129 (0.576)
γ	23.68 (4.10)	21.56 (3.35)	22.48 (3.54)	16.48 (2.44)	19.86 (3.09)
Log likelihood	-346.07	-357.22	-347.25	-369.92	-358.33

Table 6: Estimates of structural model. Standard errors, calculated using the outer product of the gradient (Berndt et al., 1974) in parentheses.

	n	means: $\max_k m^k/X$	y/X
non-ambiguous situations	44	0.500	0.711
ambiguous situations	146	0.593	0.733
difference		0.093	0.022
p-value, t-test of no difference			0.559

Table 7: Scope for moral wriggling. Non-ambiguous situations are distributional situations where all the principles that we consider agree on what is fair. Ambiguous situations are situations where the principles disagree. On average, every fairness ideal is to offer 50%.