Responsibility for what? Fairness and individual responsibility

Alexander W. Cappelen, Erik Ø. Sørensen, Bertil Tungodden

Department of Economics, Norwegian School of Economics and Business Administration, N-5045 Bergen, Norway
ESOP, University of Oslo, N-0317 Oslo, Norway
Chr. Michelsen Institute, N-5892 Bergen, Norway

Abstract

What should individuals be held responsible for? This is a fundamental question in much of the contemporary debate on distributive justice. Different fairness ideals, such as strict egalitarianism, and different versions of equal opportunity ethics and libertarianism give different answers to this question. In order to study the prevalence of these fairness ideals in society, we present the results from a dictator game where the distribution phase is preceded by a production phase. Each participant’s contribution is a result of working time, productivity, and price. We estimate what factors the participants hold each other responsible for and the weight they attach to fairness. In addition, we discuss how fairness preferences relate to business education and labour market experiences by comparing the estimates for business students at different stages of their education, and by comparing the estimates for final-year business students with the estimates for former business students with some years of work experience.

1. Introduction

What should individuals be held responsible for? This is a core question both in the modern political debate on distributive justice and in normative reasoning. In particular, there is substantial disagreement about the extent to which people should be held responsible for various factors affecting their pre-tax income. In the normative literature, there has been an increasing focus on equal opportunity theories of distributive justice that combine an egalitarian commitment with a concern for individual responsibility (Dworkin, 1981; Roemer, 1998; Rawls, 1971). The role of individual responsibility is also a controversial issue in political debates over redistribution. An important difference between left-wingers and right-wingers is the role they assign to individual responsibility in the design of redistributive policies. Typically, right-wingers argue that people should be held responsible for a large fraction of the factors that determine their incomes, while left-wingers defend a more limited role for individual responsibility.

Despite the importance of the normative question of what people should be held responsible for in distributitional situations, the positive question of what people actually do hold each other responsible for has received little attention in the experimental literature. People’s behaviour in games such as the ultimatum game and the dictator game has been widely interpreted as indicative of people being motivated by fairness considerations and being willing to sacrifice monetary gains in order to avoid large deviations from what they consider to be a fair outcome (Camerer, 2003). These games, however, are not well suited for studying attitudes towards individual responsibility, because the participants do
not contribute to the production of the value that is distributed. To address the question of what people hold each other responsible for, which is the first aim of this paper, it is necessary to study distributional preferences in more complex situations involving production.

We study a dictator game where the distribution phase is preceded by a production phase.\(^1\) In the production phase, the participants were randomly assigned one of two documents and asked to type the text on a computer. The value of their production depended on the number of minutes they decided to work, the number of correct words they typed per minute and a price for each correctly typed word that depended on the assigned text. The participants could therefore differ with respect to three factors: working time, productivity and price. For which of these factors, then, do people hold others responsible?

Different normative theories have different implications for what factors we should hold people responsible for in such situations. Two opposing views are strict egalitarianism and libertarianism. Strict egalitarianism does not hold people responsible for any of the factors determining production, while libertarianism holds people responsible for all factors. Equal opportunity ethics can be seen as an intermediate view, which holds people responsible for some, but not all, factors. In our paper, we study two interesting interpretations of this view, which we refer to as choice egalitarianism and meritocratism. Choice egalitarianism holds people responsible for their choices, while meritocratism in addition holds them responsible for all factors that can be considered personal traits.

People’s views on responsibility may be formed by the institutions they have been exposed to, which may partly explain the huge differences in fairness perceptions that are found both within and between countries (Alesina et al., 2001; Alesina and Angeletos, 2005; Alesina and La Ferrara, 2005; Piketty, 1995). The second aim of this paper is to discuss the role of institutions in shaping fairness preferences. There is a substantial literature arguing that both education and exposure to market institutions could shape people’s fairness preferences, and in this paper we try to shed some light on this issue. We do so by studying the behaviour of four different subject groups; first, second and final-year business students and business graduates with labour market experience.

We consider an economic environment where efficiency considerations play no role in the distributional phase, and where people make a trade-off between monetary gains and fairness considerations when proposing a distribution of the total production. Based on a random utility model, we estimate for each subject group the share that finds it fair to hold people responsible for a given set of factors and the importance these subject groups attach to fairness considerations. We then compare the estimates for first-year business students with the estimates for second and final-year business students, and the estimates for final-year business students with the estimates for business graduates. Given these estimates, we discuss the extent to which the observed differences may be due to business education and labour market experience.

Section 2 describes the basic model in more detail, including a further discussion of the different fairness views on responsibility. Section 3 provides a discussion of the experimental design, while the results and robustness tests are reported in Section 4. Section 5 contains some concluding comments.

2. The model

Consider an economy in which individuals differ in working time \((q)\), productivity \((a)\) and the price \((p)\) they receive per unit they produce, and where individual \(i\)’s production value \((x_i)\) is the product of these factors, such that \(x_i(a_i, q_i, p_i) = a_i q_i p_i\). Given this economic structure, we are concerned with distributional situations involving pairs of individuals, who are referred to as person 1 and person 2. The total production value to be distributed in any given distributional situation is given by \(X(a, q, p) = x_1(a_1, q_1, p_1) + x_2(a_2, q_2, p_2)\), where \(a = (a_1, a_2)\), \(q = (q_1, q_2)\), and \(p = (p_1, p_2)\).

2.1. Fairness and choice

We assume that individuals, when proposing an income distribution, are motivated by own income, \(y_i\), and a desire to act according to a fairness ideal. Individual \(i\)’s fairness ideal is denoted by \(k(i)\) and specifies a unique distribution of income, \(m^{k(i)}\) to individual \(i\) and \(X - m^{k(i)}\) to the other individual, in any given distributional situation. We also assume that the marginal disutility of deviating from the fairness ideal is strictly increasing in the size of the deviation from the fair distribution. Specifically, we assume that person \(i\) is maximizing the following utility function when proposing a

\[^1\] For other studies of real-effort dictator games, see Cappelen et al. (2007), Cherry et al. (2002), Frohlich et al. (2004), Konow (2000), and Konow et al. (2009). The importance of differences in individual endowments has also been studied in the context of ultimatum games, public-good games and redistributive voting games; see among others Cherry et al. (2005), Clark (1998), Durante and Putterman (2007), and Hoffman and Spitzer (1982).

\[^2\] There is an independent question of whether people with different preferences self-select into different types of educations and occupations, for example if market-friendly people to a greater extent choose to study subjects such as economics and business (see, e.g., Frey and Meier, 2005).

\[^*\] For an interesting survey on this issue, see Faravelli (2007).
distribution:
\[
V_i(y; \mathbf{a}, \mathbf{q}, \mathbf{p}) = y_i - \frac{\beta_i}{2} (y_i - m^{ik}(\mathbf{a}, \mathbf{q}, \mathbf{p}))^2 / X(\mathbf{a}, \mathbf{q}, \mathbf{p}),
\]
where \( \beta_i \geq 0 \) is the weight individual \( i \) assigns to fairness considerations. For an interior solution, the optimal proposal, \( y_i^* \), is
\[
y_i^* = m^{ik}(\mathbf{a}, \mathbf{q}, \mathbf{p}) + X(\mathbf{a}, \mathbf{q}, \mathbf{p}) / \beta_i.
\]

The optimal proposal thus depends on the fairness ideal endorsed by the individual and the importance assigned to fairness considerations. Note that this specification of the utility function implies that the proposed share \( y_i^*/X \) is given by the fair share \( m^{ik}/X \) plus a constant determined by the weight attached to fairness considerations. Alternative formulations of the utility function are considered in Section 4.4.

2.2. Fairness and the responsibility cut

We assume that the individuals endorse one of the following fairness principles: strict egalitarianism, one of two versions of equal opportunity ethics (choice egalitarianism or meritocratism) or libertarianism. The difference between these four fairness principles lies in the implications they have for what is often referred to as the responsibility cut, that is, for what factors individuals should be held responsible and for what factors individuals should not be held responsible. In formalizing the responsibility cut, it is useful to introduce the responsibility set \( R^k \subset [a, q, p] \). The responsibility set represents the factors for which people are held responsible under fairness ideal \( k \).

According to strict egalitarianism, total income should be distributed equally between the two individuals, independently of how it came about. Hence, the individuals are not held responsible for any of the factors affecting production, which implies that the responsibility set is simply the empty set, \( R^{SE} = \emptyset \). The strict egalitarian view is closely related to the motivation captured in the inequality-aversion models, which assume that people dislike unequal outcomes, independently of the source of inequality (see Fehr and Schmidt, 1999; Frohlich et al., 2004).

Equal opportunity ethics objects to strict egalitarianism because it believes that people should be held responsible for some of the factors affecting the outcome of their actions (Roemer, 1998). There are different versions of this position, which reflect different views on how to justify the responsibility cut. In this paper, we consider two versions. First, we consider choice egalitarianism, which follows from the view that people should not be held responsible for factors beyond their control. If only working time is considered to be within individual control, as we will argue shortly is the most likely scenario in our experiment, the responsibility set is given by \( R^{CE} = \{q\} \), and the fair distribution is to give each person a share of the total income equal to his or her share of the total working time. If additional factors are considered to be within individual control, the choice egalitarian responsibility set would expand correspondingly.

Second, we consider meritocratism, which argues that an individual should be held responsible for all factors that can be considered personal traits, but not for factors that are unrelated to individual merits (Arrow et al., 2000). In our setting, where prices are randomly assigned, it seems natural to assume that the meritocratic responsibility set is given by \( R^{M} = \{a, q\} \). A fair distribution according to meritocratism would be to give each person a share of the total income equal to his or her share of the total production.

The libertarian fairness ideal is at the opposite extreme of strict egalitarianism and implies that people are held responsible for all factors affecting their income. It has been defended as the only way of respecting people’s self-ownership (Nozick, 1974). The responsibility set is given by \( R^{L} = \{a, q, p\} \), and the fair distribution is simply to give each person the value of what they produce. The libertarian solution may thus involve an unequal distribution of income due to differences in prices as well as in individual productivity and working time.

There are other logically possible ways of drawing the responsibility cut, \( \{a\} \), \( \{a, p\} \), \( \{p\} \), and \( \{q, p\} \). However, there does not seem to be any reasonable justification for these approaches, and we therefore ignore them in the following.

We can capture the four relevant fairness ideals by the following general fairness function:
\[
m^{ik}(\mathbf{a}, \mathbf{q}, \mathbf{p}) = \frac{r(R^k, a_1, q_1, p_1)}{r(R^k, a_1, q_1, p_1) + r(R^k, a_2, q_2, p_2)} X(\mathbf{a}, \mathbf{q}, \mathbf{p}),
\]
where
\[
r(R^k, a_i, q_i, p_i) = \begin{cases} 1/2 & \text{if } k = SE, \\ q_i & \text{if } k = CE, \\ a_i q_i & \text{if } k = M, \\ a_i q_i p_i & \text{if } k = L. \end{cases}
\]

The function \( r(R^k, a_i, q_i, p_i) \) may be interpreted as representing the fair claim of each individual, and thus the general fairness function may be seen as distributing the total production proportionally to the individuals’ fair claims. To illustrate, consider a situation where the working time of person 1 is three times the working time of person 2. A choice egalitarian would argue that the fair claim of each individual is determined by working time, and consequently that the fair claim of
person 1 is three times the fair claim of person 2. As a result, the choice egalitarian would argue that it is fair that person 1 should receive three-fourths of the total production.

3. Experimental design

Before presenting the details of the actual experiment, we provide a discussion of how our subject pool was recruited. Finally, we discuss how the variation in the actual distributional situations make it possible to identify heterogeneity in the fairness perceptions of the participants.

3.1. The subject pool

The participants in the experiment were recruited from first-, second- and fourth-year students at the Norwegian School of Economics and Business Administration (NHH); in addition, we also invited former students from the same school to constitute the alumni group. The participants were not informed about the purpose of the experiment but only invited to take part in a research project. They were told that they would receive a participation fee of 100 NOK (approximately 15 USD) and that they might earn additional money in the experiment. In the invitation, they were also asked whether they wanted to participate in a long or a short version of the experiment, and they were told that their monetary gain might depend on the length of the experiment.

We had nine student sessions, where, in total, 238 students participated: 82 first-year students, 84 second-year students and 72 fourth-year students. In addition, we had one alumni session with 57 participants. The share of female participants in the four subject groups was 28%, 43%, 39%, and 32%, respectively. Each session consisted only of participants from the same subject group and each contained a long and a short version of the experiment.

The students were recruited with the same set of procedures, and at each level we recruited approximately 20% of the student population. The drop out rate from the first to the fourth year at NHH is small, approximately 10%. Hence, the main differences between the student subject groups are that the students at the higher levels have received more business education and are, on average, older. In our analysis, we cannot distinguish these two effects, so we cannot rule out that what we as a shorthand description sometimes refer to as the effect of “business education” may partly be an age effect.

The recruitment of alumni participants was based on a list of all former students at the school with 2–10 years of potential work experience and a postal address in Bergen. Bergen is the second largest city in Norway, and the population is in general nationally representative. We recruited approximately 30% of the former students on this list. The alumni group differ from the student groups in many respects, including the fact that they are older and more established in life. Hence, we cannot rule out that what we as a shorthand description sometimes refer to as the effect of “labor market experiences” may partly be due to these other factors.

3.2. The experiment

At the beginning of the experiment, all participants are informed about the rules of the game, and given a complete description of how the game will proceed. They are also told that all communication will be anonymous and conducted through a web-based interface. Each participant is then randomly, with equal probability, assigned document A or document B. Both documents are reports from public commissions in Norway. The average length of the words in each document is approximately the same. In the production phase, the participants are asked to type the text in their assigned document into a word processing file on their computer, where correctly typed words are given the value of 1 NOK (approx. 0.16 USD) for document A and 0.5 NOK (approx. 0.08 USD) for document B. The price for a correctly typed word is thus completely outside the control of the participants. In order to make it easier for the participants to make calculations in the distribution phase, we round off their reported production to the nearest 50. The value of each participant’s production is therefore equal to the reported production multiplied by the value of each correctly typed word. Our choice of prices and production time ensures that participants exercise high effort in the production phase and that high stakes are involved in the distribution phase.

---

4 The complete instructions are available upon request from the authors.

5 We introduced different documents in order to avoid individuals being paid differently for the exact same product. A natural interpretation of the price differential could then be that it reflected differences in the demand for the two products. To make the production process as realistic as possible, the participants were allowed to use the Word spell check to make sure that they were spelling the words correctly.
In the distribution phase, each participant is matched with a sequence of other participants, maximally six persons, where no information is given about the outcome of any distributional situation before all proposals are made. The participants are matched both with people who have the same working time and with people who have a different working time than themselves. For each match the participants are given information about the working time, production, and production value of both participants. They are then asked to propose a distribution of the total production value for the two participants. The proposals have to be rounded off to the nearest 25 NOK. When they have made proposals for all the matches, they are given an overview of all their proposals and asked to give a final confirmation. If they regret some of their proposals, they are allowed to start over again and make new proposals. In total, the 295 participants made 1230 distributional choices.

When everyone has confirmed their proposals, actual payment for each person is determined by randomly drawing a proposal from one of the distributional situations that this person has been involved in. There is an equal probability that this is his or her own proposal or the proposal of one of the other participants. Hence, if a person makes four distributional choices, there is a one in eight chance that a given choice is implemented. A participant’s total earnings are then the sum of the amount assigned to this person in the selected proposal and the participation fee. At the end of the experiment, the participants are assigned a code and asked to write this code together with their bank account number on a claim form. This form is placed in an envelope addressed to the accounting division of NHH. The procedure ensures that neither the participants nor the research team are in a position to identify how much each participant earns in the experiment.

3.3. The distributional situations

In the distribution phase, the paired participants might differ with respect to the working time, their production per minute and the price they receive per correctly typed word. Differences in working time would be due to different individual choices, whereas differences in the price would be beyond individual control. Differences in production per minute, reported in Fig. 1, might be due both to differences in the chosen intensity of work effort in the experiment and to differences in the innate ability to type.\footnote{Most participants were paired with four other participants, but we allowed for some variation in this number in order to deal with sessions of different sizes.\footnote{We observe that there is larger dispersion in the productivity among the alumni; a standard deviation of 9.7 versus 7.2 in the student groups taken together. However, this difference is not significant; a Levene’s test of equality of variance gives a \( p \)-value of 0.12.}}
Given the high piece-rate payment and the short period for production, however, we find it quite unlikely that the participants would choose not to work at maximum effort in the experiment, or that they would suspect others not to do so.\(^8\) We therefore interpret differences in production per minute as reflecting different innate abilities beyond individual control, and we assume that the participants also hold this view when making distributional choices. Hence our preferred interpretation of choice egalitarianism in this experiment is that people are only held responsible for working time.

In general, the fairness principles justify different fair distributions, but they coincide in certain situations. First, if the participants are identical with respect to both working time and price, then the strict egalitarian and the choice egalitarian fairness ideals imply the same fair distribution, as would the meritocratic and the libertarian fairness ideals. And all four fairness ideals coincide if the participants also have the same productivity. Second, if the participants differ in working time but are given the same price, the meritocratic and the libertarian fairness ideals justify the same fair distribution, which again would be equal to the choice egalitarian distribution if the participants also have the same productivity. Third, if the participants have the same working time but are assigned different prices, then the choice egalitarian and strict egalitarian fairness ideal coincide, and they justify a distribution equal to the meritocratic position if the participants also have the same productivity. If the participants differ along all dimensions, then none of the fairness ideals would coincide perfectly.

To obtain a better picture of the potential variation in proposals implied by the different fairness ideals, Fig. 2 shows how the various fairness ideals correlate in all the distributional situations faced by the individuals in the experiment. If two fairness ideals were to coincide in all the distributional situations, then all the points should be at the diagonal in the respective comparison. To illustrate, consider the comparison of the strict egalitarian and the choice egalitarian fairness ideals. The two fairness ideals coincide when the individuals have the same working time, and these distributional situations, a total of 550, are therefore located at the diagonal. In 340 distributional situations, person 1 has worked for 30 min and person 2 has worked for 10 min. The strict egalitarian fairness ideal still justifies equal sharing, whereas the choice egalitarian fairness ideal justifies that person 1 takes three-fourths of the total production value. Consequently, these distributional situations are located at the line above the diagonal. By design, given that both individuals in a pair make a proposal, each distributional situation has a symmetric counterpart, and hence there are also 340 distributional situations where person 1 has worked for 10 min and person 2 has worked for 30 min. These distributional situations are located at the line below the diagonal.

More generally, we observe from Fig. 2 that there are substantial differences in how much the fairness ideals justify that person 1 can take in a given distributional situation. Thus, in the analysis, we should be able to identify whether there is heterogeneity in the fairness perceptions of the individuals.

### Results

As a background for the analysis, we present some descriptive statistics. We then formulate a random utility model and estimate how our subjects actually draw the responsibility cut and the weight they attach to fairness considerations.

---

\(^8\) See Dickinson (1999) for an experimental study of the distinction between extensive effort, for example working time, and the intensity of work effort.

\(^9\) Indeed, the participants assigned a low price were slightly more productive than those with a high price (a standard t-test of the hypothesis of no mean difference gives a p-value of 0.052 in the 10-min group and 0.0005 in the 30-min group). While this, in principle, could be because a lower price induced more effort, we believe that the reason for this difference is a slight difference in the difficulty of the two texts that we did not uncover in our pre-experimental testing. Note that the difference was too small for the participants to learn about this from the few distributional situations they observed, and thus there is no reason to believe that they did not accept the statement in the procedures saying that the two texts were equally difficult to type.
4.1. Descriptive statistics

Table 2 presents some key information about individual offers. The average total production value to be distributed was 735 NOK (approximately 124 USD), and the average offer to the opponent was 38.5% (which amounted to 286 NOK; approximately 48 USD). This is substantially higher than is commonly observed in standard dictator games without production (Andreoni and Miller, 2002; Camerer, 2003) and may indicate that the presence of a production phase caused people to care more about fairness considerations. We observe that the fourth-year students offered almost the same as the first-year students, the alumni offered more than the fourth-year students, and females offered more than males. Note, however, that the size of the average offer does not provide us with any information about the prevalence of the different fairness principles. The average fair offer is the same for all four fairness ideals, namely $X/2$. This is easily seen by observing that, for any particular distributional situation and for any fairness ideal $k$, the fair solution would be that person

![Graph showing scatter plots of fairness ideals.](image-url)

**Fig. 2.** Scatter plots of fairness ideals. Pair-wise plots of $m^k(a, q, p)$ against $m^j(a, q, p)$ for all the distributional situations in our data set (i.e., each plot represents what the fairness ideals $m^k$ and $m^j$, respectively, justify that person 1 takes in a given situation).

<table>
<thead>
<tr>
<th>Share, $(X - y)/X$</th>
<th>1st year</th>
<th>2nd year</th>
<th>4th year</th>
<th>Alumni</th>
<th>All students</th>
<th>All males</th>
<th>All females</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.396</td>
<td>0.338</td>
<td>0.383</td>
<td>0.430</td>
<td>0.373</td>
<td>0.360</td>
<td>0.385</td>
<td></td>
</tr>
<tr>
<td>Standard deviation</td>
<td>0.188</td>
<td>0.233</td>
<td>0.223</td>
<td>0.172</td>
<td>0.216</td>
<td>0.215</td>
<td>0.185</td>
<td>0.208</td>
</tr>
<tr>
<td>Minimum</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Maximum</td>
<td>0.789</td>
<td>0.944</td>
<td>0.862</td>
<td>0.973</td>
<td>0.944</td>
<td>0.920</td>
<td>0.973</td>
<td>0.973</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Amount, $X - y$, in NOK</th>
<th>1st year</th>
<th>2nd year</th>
<th>4th year</th>
<th>Alumni</th>
<th>All students</th>
<th>All males</th>
<th>All females</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>274</td>
<td>253</td>
<td>266</td>
<td>361</td>
<td>264</td>
<td>268</td>
<td>320</td>
<td>286</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>189</td>
<td>223</td>
<td>199</td>
<td>203</td>
<td>204</td>
<td>205</td>
<td>208</td>
<td>207</td>
</tr>
<tr>
<td>Maximum</td>
<td>1000</td>
<td>1000</td>
<td>850</td>
<td>900</td>
<td>1000</td>
<td>1000</td>
<td>1000</td>
<td>1000</td>
</tr>
<tr>
<td>Share offering nothing</td>
<td>0.043</td>
<td>0.135</td>
<td>0.156</td>
<td>0.040</td>
<td>0.108</td>
<td>0.129</td>
<td>0.026</td>
<td>0.093</td>
</tr>
<tr>
<td>Share offering more than half</td>
<td>0.296</td>
<td>0.270</td>
<td>0.299</td>
<td>0.319</td>
<td>0.288</td>
<td>0.261</td>
<td>0.358</td>
<td>0.295</td>
</tr>
</tbody>
</table>

Some key features of offers made to opponent.

4.1. Descriptive statistics

Table 2 presents some key information about individual offers. The average total production value to be distributed was 735 NOK (approximately 124 USD), and the average offer to the opponent was 38.5% (which amounted to 286 NOK; approximately 48 USD). This is substantially higher than is commonly observed in standard dictator games without production (Andreoni and Miller, 2002; Camerer, 2003) and may indicate that the presence of a production phase caused people to care more about fairness considerations. We observe that the fourth-year students offered almost the same as the first-year students, the alumni offered more than the fourth-year students, and females offered more than males.

Note, however, that the size of the average offer does not provide us with any information about the prevalence of the different fairness principles. The average fair offer is the same for all four fairness ideals, namely $X/2$. This is easily seen by observing that, for any particular distributional situation and for any fairness ideal $k$, the fair solution would be that person
1 offers $X - m^k$ to person 2 and that person 2 offers $m^k$ to person 1. Hence, it is not the case that a high average offer indicates, for example, a large share of strict egalitarians and a small share of libertarians.

This point can be illustrated further by looking at a particular distributional situation where we observed a very high offer of 95% of the total income. In this case, person 1 had worked for 10 min, produced 100 words and received a price of 0.5 NOK. Person 2, on the other hand, had worked for 30 min, produced 850 words and received a price of 1 NOK. Hence, they had 900 NOK to split between them. Given this information, person 1 acted in accordance with the libertarian fairness ideal and offered person 2 almost all of the total income, 850 NOK.

Interestingly, in 29.5% of the offers, the participant proposed more to the opponent than to him or herself, whereas in 9.3% of the offers, the participant proposed everything to him or herself. Such offers were quite rare among first-year students and alumni, while about 15% of the offers from second-year and fourth-year students left nothing to the opponent.

### 4.2. Empirical model

We adapt the choice model to bring it into line with the fact that we asked the participants to round off their proposals to the nearest 25 NOK. The choice of $y$ is therefore restricted to the set $\{0, 25, 50, \ldots, X(a, q, p)\}$, and we introduce a random element in the utility model:

$$U_i(y; \cdot) = V_i(y; \cdot) + \epsilon_{iy}/\gamma.$$  

(5)

The constant $\gamma$ determines the weight given to the random element $\epsilon_{iy}$, which we assume specific to each element in $y$ and i.i.d. extreme value distributed. The individuals choose $y^*$ such that $U_i(y^*; \cdot) \geq U_i(y; \cdot)$ for all $y$ in $\{a, q, p\}$.

This model has a mixed logit structure where each person is characterized by his or her fairness ideal, $(k, i)$, as well as by the parameter $\beta_k$. We cannot classify individuals by $(k, \beta)$, but we estimate the distribution of these characteristics. The distribution of fairness ideals is discrete, and we approximate the distribution of $\beta$ by a log-normal distribution, such that $\log \beta \sim \mathcal{N}(\zeta, \sigma^2)$. Because the fairness ideal and the importance a person assigns to fairness considerations are unobserved, these must be integrated out for the unconditional choice probabilities as functions of the observed variables. If we introduce a subscript $j = 1, \ldots, J_i$ indexing the situations an individual $i$ faces, and let $\lambda_k$ denote the share of the subject group that holds the fairness ideal $k$, the likelihood of an individual $i$ making proposals $(y_1, \ldots, y_J)$ from the sets of feasible proposals $(Y_1, \ldots, Y_J)$ given a parameter vector $\theta = (\lambda_{SE}, \lambda_{CE}, \lambda_M, \lambda_L, \gamma, \zeta, \sigma)$, is

$$L_i(\theta) = \sum_k \lambda_k \int_0^\infty \left( \prod_{j=1}^J \frac{e^{\gamma/(\epsilon_{ij} a_k \epsilon_{ij} q_k \epsilon_{ij} p_k k^\beta)}}{\sum_k \sum_{j=1}^J e^{\gamma/(\epsilon_{ij} a_k \epsilon_{ij} q_k \epsilon_{ij} p_k k^\beta)}} \right) dF(\beta; \zeta, \sigma).$$

(6)

Repeated observations and the fact that we expose individuals to different distributional situations provide information about the distribution of fairness ideals and the weight people attach to them. Given that the distributional situations are symmetrical and all fairness ideals specify that an equal split is fair on average, the average offer provides identification of the weight people on average attach to fairness. On the basis of how intra-individual and inter-individual variance in offers relate to observable characteristics of the distributional situation, we estimate the distribution of $\beta$ and the population shares $\lambda_k$. Note that for the inference, each individual’s sequence of choices is treated as a single observation.

### 4.3. Estimates of the model

Table 3 reports the estimates of the model for each of the four subject groups, where the population share for each of the fairness ideals is estimated with the proportion of participants acting in accordance with this particular fairness ideal. It also reports the estimates for all students taken together and all participants broken down by gender.\(^{10}\)

Consider first our estimates for the first-year students in Table 3. The estimated share of strict egalitarians is 23.5%, the share of choice egalitarians is 6.6%, the share of meritocrats is 41.5% and the share of libertarians is 28.4%. There seems, in other words, to be considerable pluralism in how these students draw the responsibility cut. However, at the same time, the analysis reveals overlapping consensus on a number of issues. First, almost 80% of these first-year students (choice egalitarians, meritocrats and libertarians) find it fair to hold people responsible for the randomly assigned price. Second, more than 70% (strict egalitarians, choice egalitarians and meritocrats) do not find it fair to hold people responsible for their chosen working time. Second, while only a relatively small minority hold others responsible for the price. If differences in productivity are considered to be beyond individual control, which we find likely given the design of the experiment, this implies that many people justify the responsibility cut on the basis of individual merits, and not on the distinction between factors that are beyond and factors that are within individual control.

\(^{10}\) The restrictions that a common model applies to all students, or to both students and alumni, or to both genders, can all be rejected with $p<0.01$ using likelihood ratio tests.
How is the prevalence of different fairness ideals related to years of business education? If we compare the estimates for the first-year students and the fourth-year students, we see a sharp increase in the share of meritocrats from 41.5% to 57.8%. At the same time, surprisingly, there is a substantial drop in the share of libertarians from 28.4% to 15.0%. Given the strong focus on the efficiency of markets in business education, we expected final year students to be more inclined to find it fair that people receive the value of their production. Instead we observe that a majority of students endorse the meritocratic fairness ideal. A possible explanation for this finding is that the students are shaped by the meritocratic features of the educational system, which aims to reward effort and innate ability, but not pure luck. For the other two positions, strict egalitarianism and choice egalitarianism, we hardly see any difference at all between the subject groups.

In order to study how fairness perceptions relate to participation in the labour market, we compare the fourth-year students with the alumni group. As we observe from Table 3, there are large and systematic differences in the prevalence of the four fairness ideals between the two groups. We see a sharp decrease in the prevalence of all fairness ideals except for libertarianism. In fact, the share of libertarians triples from 15.0% to 44.8%. This may reflect the same kind of mechanism as that people receive the value of their production. Instead we observe that a majority of students endorse the meritocratic fairness ideal. A possible explanation for this finding is that the students are shaped by the meritocratic features of the educational system, which aims to reward effort and innate ability, but not pure luck. For the other two positions, strict egalitarianism and choice egalitarianism, we hardly see any difference at all between the subject groups.

It is also interesting to study how business education and labour market experience relate to the weight people attach to fairness considerations. We observe from Table 3 that there are only minor differences in the estimated means of logβ for first-year students and fourth-year students, whereas there is a substantial increase when comparing fourth-year students and alumni. To understand the overall effect of the estimated values of γ, ζ and σ, we provide Fig. 3. This figure takes as the point of departure a situation where the total produced value is 1000 and the fairness ideal endorsed by a hypothetical individual specifies an equal split. We then provide for each of the four subject groups, and for the 10th, 30th, 50th, 70th and 90th percentiles of the distribution of β, the deterministic utility and, plotted as solid bars, the implied choice probabilities for all multiples of 25 NOK for this hypothetical individual.

The general impression from Fig. 3 is that business education does not have much impact on the weight people attach to fairness. Except for a dip in the expected offer at the lower percentiles for second-year students, the distribution of choice probabilities is very close for all student groups. In contrast, labour market experience appears to relate positively to the weight attached to fairness. At all percentiles, the expected offer is higher for the alumni than for the fourth-year students. We should be careful, however, about how we interpret this result, because there are many possible explanations. For example, there are more parents among the alumni, and it is already established in Peters et al. (2004) that parents seem to act more in line with fairness considerations than what is typically observed among students. Moreover, alumni are also wealthier than students, and this may also contribute to make them more inclined to act in line with fairness considerations.

To see how well our estimates predict the actual distribution of offers, we simulate distributions of offers for the distributional situations of each of the four subject groups in the experiment. As we can see from Fig. 4, the fit is good.

There are rather small differences in the gender composition between the subject groups, and, as seen from Tables 2 and 3, also small differences in the choices of males and females. Females assign more weight to fairness than males, but still this difference is too small to have an impact on our main findings on how the weight attached to fairness relates to...
business education and labour market experiences. In a comparison of the first-year and the fourth-year students, the predicted effect of the change in gender composition amounts to an increase in the average offer of less than 1%; in a comparison of the fourth-year students and the alumni, taking the gender composition into account merely strengthens

Fig. 3. Illustration of implied choice probabilities. The graphs plot $g(y/C_1) = g(y/C_0) b(y/C_0 m)$, for a hypothetical individual with $m = 0.5$ (marked by a vertical line). Calculated at the 10th, 30th, 50th, 70th and 90th percentiles of the estimated $b$ distribution using the estimates in Table 3. Money, $y$, is measured in thousands of NOK.

Fig. 4. Empirical distribution function of offers. Distributions of offers made (as share of total production) and predictions from the estimated model for each experimental group. The black line is our experimental data while the grey line is our prediction, for the distributional situations in the data set, made from the estimates in Table 3.
Table 4
Robustness tests.

<table>
<thead>
<tr>
<th>Subject group</th>
<th>1st year</th>
<th>2nd year</th>
<th>4th year</th>
<th>Alumni</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>R1: probability mass, $\phi_0$, at $\beta = 0$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\lambda_{SE}$</td>
<td>0.248</td>
<td>0.216</td>
<td>0.228</td>
<td>0.131</td>
<td>0.186</td>
</tr>
<tr>
<td>$\lambda_{CE}$</td>
<td>0.071</td>
<td>0.071</td>
<td>0.058</td>
<td>0.024</td>
<td>0.045</td>
</tr>
<tr>
<td>$\lambda_{M}$</td>
<td>0.411</td>
<td>0.428</td>
<td>0.574</td>
<td>0.401</td>
<td>0.471</td>
</tr>
<tr>
<td>$\lambda_{L}$</td>
<td>0.270</td>
<td>0.285</td>
<td>0.140</td>
<td>0.445</td>
<td>0.297</td>
</tr>
<tr>
<td>$\phi_0$</td>
<td>0.111</td>
<td>0.191</td>
<td>0.168</td>
<td>0.045</td>
<td>0.120</td>
</tr>
<tr>
<td>Log likelihood</td>
<td>$-774.6$</td>
<td>$-726.2$</td>
<td>$-635.0$</td>
<td>$-567.9$</td>
<td>$-2744.5$</td>
</tr>
<tr>
<td>R2: utility loss in absolute terms: $V(y_{i, \gamma}) = y - \beta(y - m)^2/2$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\lambda_{SE}$</td>
<td>0.297</td>
<td>0.219</td>
<td>0.258</td>
<td>0.143</td>
<td>0.217</td>
</tr>
<tr>
<td>$\lambda_{CE}$</td>
<td>0.078</td>
<td>0.082</td>
<td>0.078</td>
<td>0.025</td>
<td>0.054</td>
</tr>
<tr>
<td>$\lambda_{M}$</td>
<td>0.396</td>
<td>0.476</td>
<td>0.568</td>
<td>0.399</td>
<td>0.470</td>
</tr>
<tr>
<td>$\lambda_{L}$</td>
<td>0.229</td>
<td>0.223</td>
<td>0.097</td>
<td>0.433</td>
<td>0.260</td>
</tr>
<tr>
<td>Log likelihood</td>
<td>$-786.4$</td>
<td>$-739.5$</td>
<td>$-651.3$</td>
<td>$-567.1$</td>
<td>$-2784.4$</td>
</tr>
<tr>
<td>R3: utility loss in relative terms: $V(y_{i, \gamma}) = y - \beta(y - m)/X_{iy_{i, \gamma}}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\lambda_{SE}$</td>
<td>0.180</td>
<td>0.198</td>
<td>0.185</td>
<td>0.115</td>
<td>0.156</td>
</tr>
<tr>
<td>$\lambda_{CE}$</td>
<td>0.052</td>
<td>0.052</td>
<td>0.042</td>
<td>0.022</td>
<td>0.038</td>
</tr>
<tr>
<td>$\lambda_{M}$</td>
<td>0.403</td>
<td>0.409</td>
<td>0.535</td>
<td>0.410</td>
<td>0.451</td>
</tr>
<tr>
<td>$\lambda_{L}$</td>
<td>0.366</td>
<td>0.340</td>
<td>0.238</td>
<td>0.453</td>
<td>0.355</td>
</tr>
<tr>
<td>Log likelihood</td>
<td>$-774.0$</td>
<td>$-737.2$</td>
<td>$-638.0$</td>
<td>$-570.9$</td>
<td>$-2752.0$</td>
</tr>
<tr>
<td>R4: heteroskedastic random utility shocks: $U_i(y_{i, \gamma}) = V(y_{i, \gamma}) + X_{iy_{i, \gamma}}/\gamma$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\lambda_{SE}$</td>
<td>0.227</td>
<td>0.272</td>
<td>0.214</td>
<td>0.123</td>
<td>0.197</td>
</tr>
<tr>
<td>$\lambda_{CE}$</td>
<td>0.067</td>
<td>0.053</td>
<td>0.056</td>
<td>0.023</td>
<td>0.045</td>
</tr>
<tr>
<td>$\lambda_{M}$</td>
<td>0.412</td>
<td>0.433</td>
<td>0.522</td>
<td>0.405</td>
<td>0.449</td>
</tr>
<tr>
<td>$\lambda_{L}$</td>
<td>0.294</td>
<td>0.252</td>
<td>0.208</td>
<td>0.449</td>
<td>0.309</td>
</tr>
<tr>
<td>Log likelihood</td>
<td>$-773.75$</td>
<td>$-733.10$</td>
<td>$-647.24$</td>
<td>$-569.18$</td>
<td>$-2753.9$</td>
</tr>
<tr>
<td>R5: heteroskedastic random utility shocks: $U_i(y_{i, \gamma}) = V(y_{i, \gamma}) + (\gamma - V)e_{iy_{i, \gamma}}/\gamma$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\lambda_{SE}$</td>
<td>0.240</td>
<td>0.231</td>
<td>0.250</td>
<td>0.122</td>
<td>0.211</td>
</tr>
<tr>
<td>$\lambda_{CE}$</td>
<td>0.076</td>
<td>0.0+</td>
<td>0.047</td>
<td>0.0+</td>
<td>0.023</td>
</tr>
<tr>
<td>$\lambda_{M}$</td>
<td>0.421</td>
<td>0.481</td>
<td>0.581</td>
<td>0.367</td>
<td>0.457</td>
</tr>
<tr>
<td>$\lambda_{L}$</td>
<td>0.263</td>
<td>0.288</td>
<td>0.123</td>
<td>0.511</td>
<td>0.308</td>
</tr>
<tr>
<td>Log likelihood</td>
<td>$-832.7$</td>
<td>$-754.9$</td>
<td>$-666.4$</td>
<td>$-697.8$</td>
<td>$-2973.4$</td>
</tr>
</tbody>
</table>

Estimated population shares under five alternative specifications of the empirical model. An asterisk indicates that this number was driven to 0 in estimation, and fixed at this value to achieve convergence.

our finding (since there is a larger share of females among the fourth-year students). Turning to the analysis of the prevalence of fairness ideals, we observe from Table 3 that the estimates are remarkably similar for males and females. There is a slight difference in the composition of the egalitarians, but overall the shares of egalitarians, meritocrats, and libertarians are almost identical. Thus nor for our discussion of how fairness perceptions relate to business education and labour market experiences, should the small differences in gender composition matter.

4.4. Robustness tests

Our estimates rely on assumptions about the distribution of $\beta$ and the functional form of the utility function. In order to assess how sensitive our estimates are to our particular specifications, we provide specification tests of our estimates for the student groups and the alumni in Table 4.

In specification R1 (in Table 4), we examine whether our assumption of a log-normal distribution of $\beta$ is crucial to our results. We do so by adding a mass $\phi_0$, at the point $\beta = 0$, such that $\beta$ is now distributed as a mixture of 0 with probability $\phi_0$ and a log-normal variate with probability $1 - \phi_0$. We see that this probability, estimated to be in the range 0.049–0.170, is well above zero. However, the increase in the log-likelihood is not dramatic, and the estimated population shares hardly change at all. From this we conclude that the assumption of log-normality is not crucial to our results.

In specifications R2 and R3, we consider alternative functional forms of the utility function. In R2, we let the utility loss from a deviation from the fairness ideal depend on the magnitude of the deviation, regardless of the total income produced. We see that this specification does uniformly worse than our preferred estimates in terms of likelihood values.

---

11 The predicted effect of the change in gender composition from the first-year to the fourth-year is calculated from Table 2, $(0.39 - 0.28) \cdot (0.433 - 0.360) = 0.008$.
12 We do not report new estimates for $\gamma$, $\zeta$, $\sigma$, since they are in line with what is reported in Table 3.
5. Concluding remarks

The notion of responsibility plays an important role in the political debate and has been the focus of much work in modern normative theory. Nonetheless, the positive question of what people actually hold each other responsible for has received little attention. One aim of the present study has been to conduct an experiment that allows us to examine this question in more detail. More specifically, we have studied the extent to which people hold each other responsible for price, working time, and productivity in distributional choices.

Our analysis shows that a large majority of the students who took part in the experiment did not hold people responsible for the randomly assigned price, an impersonal factor beyond individual control. The support for this position was significantly lower among business graduates with some years of work experience, but even within this group, a majority of the participants did not hold people responsible for the price.

There was also broad support for the view that individuals should be held responsible for their choice of working time. More than 75% of the participants in all the four subject groups rejected a strict egalitarian position. This conclusion is also illustrated by the fact that in 29.5% of the offers in the experiment, the participants proposed more to their opponent than to themselves.

Finally, the majority of both students and alumni held others responsible for their productivity. This is a factor which we consider beyond the control of the participants in the experiment, and thus it is quite interesting to observe that they viewed productivity differently from the randomly assigned price. Our preferred interpretation of this finding is that for many of our participants, the core distinction when drawing the responsibility cut was not between choices and circumstances, but between impersonal and personal factors. This finding is particularly interesting given the prominence of the choice egalitarian position in the normative literature. Choices are, however, not unimportant to most people in distributional situations. The meritocratic position also justifies holding people responsible for their choices, though not with the justification that choices are under individual control, but because they are personal characteristics that merit reward. An alternative interpretation of the observation that the participants made a distinction between productivity and price, is that most students are choice egalitarians and consider differences in productivity to be due to differences in the chosen intensity of work effort. Given the design of the experiment, with the high piece-rate payment and a short period of production, we find this interpretation unlikely, but we believe that it would be interesting to establish further experimental data on these two competing hypotheses.

A second aim of the paper has been to study how institutions, such as education and the labour market, shape people’s fairness preferences and the weight they attach to fairness. First, we establish that the alumni group to a greater extent than students is motivated by the libertarian fairness ideal, which, we argue, relates to their labour market experience. However, when interpreting this result, one should keep in mind that this alumni group, overall, are having successful careers. Thus an important extension of our research would be to study subject groups that have been less successful in the labour market. Second, and somewhat surprisingly, we establish that final-year students of business education do not follow the libertarian fairness ideal to a greater extent than first-year students. Rather, final-year students seem to be more attracted to the meritocratic fairness ideal, an ideal which is in line with the reward structure of the educational system.

Third, our study supports previous papers reporting little evidence of business education making people more self-interested (Frey and Meier, 2005). On average, we find only a small difference in the weight attached to fairness between first-year students and final-year students. Finally, however, we establish that the alumni participants attach more weight to fairness considerations than final-year students. This is in line with one of the main findings in Henrich et al. (2004), namely that exposure to market institutions seems to trigger fair-mindedness.\footnote{Note, however, that our empirical strategy is much narrower than in Henrich et al. (2004), who provide an extensive cross sectional analysis of generosity by comparing societies with different levels of market integration.}
of this finding, since there are many important differences between students and alumni, including differences in parenthood and wealth. Still we consider it interesting to observe that fairness considerations continue to be important as people move from being students to becoming more active participants in the economy.

There is substantial heterogeneity in people’s notion of fairness and the weight they attach to fairness considerations, and we believe that an important task for future research should be to seek deeper understanding of the explanations for these differences. We have made one attempt in this direction, by studying the extent to which important institutions such as education and the labour market appear to shape individual fairness preferences. Interestingly, the main effect of these institutions seems to be that they change people’s fairness ideal, while there are only moderate effects on the weight people attach to fairness.

Acknowledgements

We would like to thank the associate editor and two referees for extremely valuable comments and suggestions. We are grateful for excellent research assistance by Ingvið Almás, Sigbjørn Birkeland, Lise Gro Bjørnsen, Marit Dolve, Hanne Falkanger, Preben Falkenberg, Andreas Friis, Daniel Freyland, Nils Anders Gjerland, Jan Vidar Håftu, Anne Marie Johansen, Magne Ludvigsen, Knut Nygaard, Kira Pronin, Alexander Sandvik and Jørgen Krog Sæbe. We have also benefited from helpful comments by Geir Asheim, Dirk Engelmann, Gorm Grønnevet, Peter Hammond, Karl Ove Moene, Kjell G. Salvanes, Agnar Sandmo, seminar participants at the Bank of Norway, Katholieke Universiteit Leuven, the Norwegian University of Science and Technology, Stanford University, the University of Bergen, the University of Copenhagen, the University of East Anglia and the University of Mannheim, and conference attendees at the European Economic Association Congress 2006 in Vienna and the Social Choice and Welfare Meeting 2006 in Istanbul. This research was supported by The Research Council of Norway, Grant 151955/510.

References