Responsibility for what? Fairness and individual responsibility

Alexander W. Cappelen Erik Ø. Sørensen
Bertil Tungodden*

December 29, 2006

*Cappelen: Norwegian School of Economics and Business Administration, Bergen, Norway. e-mail: alexander.cappelen@nhh.no. Sørensen: University of Mannheim, Germany. e-mail: esorense@rumms.uni-mannheim.de. Tungodden: Norwegian School of Economics and Business Administration and Chr. Michelsen Institute, Bergen, Norway. e-mail: bertil.tungodden@nhh.no. We are grateful for excellent research assistance by Ingvild Almås, Sigbjørn Birkeland, Lise Gro Bjørnsen, Marit Dolve, Hanne Falkanger, Preben Falkenberg, Andreas Friis, Daniel Frøyland, Nils Anders Gjerland, Jan Vidar Håtuft, Anne Marie Johansen, Magne Ludvigsen, Knut Nygaard, Kira Pronin, Alexander Sandvik and Jørgen Krog Sæbø. 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 Wien and the Social Choice and Welfare Meeting 2006 in Istanbul.

Abstract

For what should individuals be held responsible? 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 can be interpreted as giving 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 player'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 examine to what extent institutions such as education and the labor market shape fairness preferences. We do this by comparing the estimates for business students at different stages in 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.

For what should individuals be held responsible? This is a core question both in the modern political debate on distributive justice and in normative theoretical 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 (Rawls, 1971; Dworkin, 1981; Roemer, 1998). In the political debate, the controversy between left-wingers and right-wingers may, to a large extent, be interpreted as disagreement over the role individual responsibility should play 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, the positive question of what people actually hold each other responsible for has received little attention in the experimental literature. Experimental work has established, with games such as the ultimatum game and the dictator game, that people are motivated by fairness considerations and that they are willing to sacrifice pecuniary gains in order to avoid large deviations from what they consider to be a fair outcome (Camerer, 2003). These games, how-

ever, are not well suited for studying attitudes towards individual responsibility, because none of the players contribute to the production of the money that is distributed. To address the question of what people hold each other responsible for more broadly, 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 players 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 randomly assigned price for each correctly typed word. The participants may therefore differ with respect to three factors; time, productivity and price. In the experiment, work time was clearly a factor within individual control and price was clearly a factor beyond individual control. It is less clear, however, how to classify individual productivity. Most likely, the participants considered it a factor shaped in a complex way by other factors both within and beyond individual control. In this study, we are particularly interested in how people deal with differences in productivity in distributive situations. Much of the political debate on redistribution centers around the question of whether individual productivity is within or beyond individual control, and thus we believe it important to shed some light on the extent to which people find it fair to hold others responsible for this factor.

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 position, which holds people responsible for some, but not all, factors. This position has been given different interpretations. Choice egalitarianism holds people responsible for factors within individual control but not for factors beyond individual control.[2] Meritocratism, on the other hand, holds in-

[1]The dictator game with a production phase is also studied in Cappelen, Hole, Sørensen, and Tungodden (Forthcoming), Cherry, Frykblom, and Shogren (2002), Frohlich, Oppenheimer, and Kurki (2004) and Konow (2000).

[2]In Cappelen et al. (Forthcoming), we show that in an experiment where all factors clearly

dividuals responsible for a broad set of personal traits, independent of whether these traits are a result of individual choices or not. Therefore, meritocrats hold people responsible for individual productivity, while choice egalitarians only do so if they consider individual productivity a factor within individual control. In the experiment, we were not in a position to measure to what extent people believe that individual productivity is within or beyond individual control. Consequently, we cannot sharply distinguish between choice egalitarianism and meritocratism. However, we can study the prevalence of people who hold others responsible only for working time and people who hold others responsible for working time and individual productivity, and for simplicity of exposition, we will refer to the first group as choice egalitarians and the second as meritocrats. The share of individuals who are either meritocrats or libertarians is thus an estimate of the share of individuals who hold others responsible for their productivity.

People's views on responsibility may partly be formed by the institutions they have been exposed to, which may also explain the huge differences in fairness perceptions that we find both within and between countries (Piketty, 1995; Alesina, Glaeser, and Sacerdote, 2001; Alesina and La Ferrara, 2005; Alesina and Angeletos, 2005). The second main aim of this paper is to study the role of institutions in shaping fairness preferences by looking at whether a particular education, in our case business studies, and participation in the labor market make people more or less ready to hold others responsible for the factors affecting their production. There is a substantial literature arguing that both education and market integration have an effect on the importance people assign to self-interest considerations (Henrich, Boyd, Bowles, Camerer, Fehr, and Gintis, 2004; Frank, Gilovich, and Regan, 1996), but interestingly, these studies have not considered the possibility that such institutions may also have an effect on what people consider fair. Intuitively, it seems plausible to claim that both business studies and participation in the labor market make people more inclined to consider the market outcome as fair, independently of any effect that such institutions may have on the importance people attach to self-interest considerations. We propose a framework that makes it possible to study simultaneously both the effect institutions may have

could be identified as either within or beyond individual control, the most prevalent fairness ideal was to hold individuals responsible only for the factors within individual control.

on the role of self-interest and the effect institutions may have on people's notions of fairness. Based on a simple random utility model where people make a trade-off between pecuniary gains and fairness considerations when proposing a distribution of the total production, we estimate for each subject group the share that find it fair to hold people responsible for a given set of factors and the importance people in this group attach to fairness considerations. We then look at the possible effects of business education by comparing the estimates for first-year business students with the estimates for second and final-year business students, and the effect of participating in the labor market by comparing the estimates for final-year business students with the estimates for business graduates with work experience.

Section 1 describes the basic model in more detail, including a further discussion of the different fairness views on responsibility. Section 2 provides a discussion of the experimental design, while the results and robustness tests are reported in Section 3. Section 4 contains some concluding comments.

1 The model

Consider an economy in which the individuals differ in working time (q), production per time unit (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 three factors, such that $x_i(a_i, q_i, p_i) = q_i a_i p_i$. Working time is assumed to be completely within individual control, while the price is assumed to be completely beyond individual control. Individual productivity, on the other hand, is partly a result of factors within individual control and partly a result of personal characteristics beyond individual control.

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 $X(\boldsymbol{a}, \boldsymbol{q}, \boldsymbol{p}) = x_1(a_1, q_1, p_1) + x_2(a_2, q_2, p_2)$, where $\boldsymbol{a} = (a_1, a_2)$, $\boldsymbol{q} = (q_1, q_2)$, and $\boldsymbol{p} = (p_1, p_2)$. Each individual proposes an amount of income y to herself and $X - y$ to her opponent.

1.1 Fairness and choice

We assume that individuals, when proposing an income distribution, are motivated by income and a desire to act according to a fairness ideal, where individual i's *fairness ideal* is denoted by $m^{k(i)}$ and specifies a unique distribution of income $(m^{k(i)}, X - m^{k(i)})$ 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 distribution:

$$V_i(y_i; \boldsymbol{a}, \boldsymbol{q}, \boldsymbol{p}) = y_i - \frac{\beta_i}{2} \frac{(y_i - m^{k(i)}(\boldsymbol{a}, \boldsymbol{q}, \boldsymbol{p}))^2}{X(\boldsymbol{a}, \boldsymbol{q}, \boldsymbol{p})}, \tag{1}$$

where $\beta_i \geq 0$ is the weight an individual i assigns to fairness considerations. For an interior solution, the optimal proposal, y_i^*, is:

$$y_i^* = m^{k(i)}(\boldsymbol{a}, \boldsymbol{q}, \boldsymbol{p}) + X(\boldsymbol{a}, \boldsymbol{q}, \boldsymbol{p})/\beta_i. \tag{2}$$

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 $m^{k(i)}/X$ plus a constant determined by the weight attached to fairness considerations. Alternative formulations of the utility function are considered in Section 3.4.

1.2 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 $\mathcal{R}^k \subset \{a, p, q\}$. 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, $\mathcal{R}^{SE} = \emptyset$. 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) and 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. Choice egalitarians argue that individuals should be held responsible for factors within their control but not for factors beyond their control and thus view an inequality as fair if it reflects differences in choices. In our experiment, the only factor that is clearly within individual control is the choice of working time. We therefore refer to choice egalitarianism as the view that the responsibility set is given by $\mathcal{R}^{CE} = \{q\}$ and that 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.

An alternative version of equal opportunity ethics, often called *meritocratism*, is to argue that an individual should be held responsible for a broad set of personal traits, including individual productivity, independently of whether these characteristics are a result of individual choices or not (Arrow, Bowles, and Durlauf, 2000). This implies that meritocrats only consider it fair to eliminate inequalities due to factors that are unrelated to the individual merits, in other words, the responsibility set is given by $\mathcal{R}^{M} = \{q, a\}$. 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 $\mathcal{R}^{L} = \{q, a, p\}$, and the fair distribu-

tion is simply to give each person the value of what she produces. 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 ignore them in the following.

We can capture the four relevant fairness ideals in the following general fairness function:

$$m^k(\boldsymbol{a}, \boldsymbol{q}, \boldsymbol{p}) = \frac{r_1(\mathcal{R}^k, \boldsymbol{a}, \boldsymbol{q}, \boldsymbol{p})}{r_1(\mathcal{R}^k, \boldsymbol{a}, \boldsymbol{q}, \boldsymbol{p}) + r_2(\mathcal{R}^k, \boldsymbol{a}, \boldsymbol{q}, \boldsymbol{p})} X(\boldsymbol{a}, \boldsymbol{q}, \boldsymbol{p}), \tag{3}$$

where:

$$r_i(\mathcal{R}^k, \boldsymbol{a}, \boldsymbol{q}, \boldsymbol{p}) = \begin{cases} 1 & \text{if } k = SE, \\ q_i & \text{if } k = CE, \\ a_i q_i & \text{if } k = M, \\ a_i q_i p_i & \text{if } k = L \end{cases}.$$

As the general fairness function makes clear, the fairness ideals, except for strict egalitarianism, distribute the total production proportionally to the product of the values of the responsibility factors for each person. Given our specific interest in how people deal with differences in productivity in distributive situations, it is important to note that we have two fairness ideals that do not hold people responsible for productivity (strict egalitarianism and choice egalitarianism) and two fairness ideals that do hold people responsible for this factor (meritocratism and libertarianism).

2 Experimental design

The experiment was a dictator game in which the distribution phase was preceded by a production phase. The participants in the experiment were recruited from first-, second- and fourth-year students at the Norwegian School of Economics and Business Administration. In addition, we invited alumni with some years of work experience (2–10 years). 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. 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.

At the beginning of the experiment, all participants were informed about the rules of the game, and they were given a complete description of how the game would proceed.[3] Each participant was then randomly, with equal probability, assigned document A or document B. Both documents were reports from public commissions in Norway. The average length of the words in each document was approximately the same. In the production phase, the participants were asked to type the text in their assigned document into a word processing file on their computer, where correctly typed words were 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 was thus completely outside the control of the participants. The participants were also informed that the two texts were equally difficult to type and thus that the price differential did not reflect any difference in the time it took to copy the two texts.

The length of the production phase depended on which version of the experiment they chose to participate in. Those who participated in the short version worked for 10 minutes, while those who participated in long versions of the experiment worked for 30 minutes. The short and the long version of the experiment took place in two different computer labs. Those who participated in the long version were asked to meet up 20 minutes before those who participated in the short version. Hence, the production phase ended at the same time for the two groups. When the production phase ended, they submitted the text, and this was automatically compared with the original document. In order to make it easier for the participants to make calculations in the distribution phase, we rounded off the reported production to the nearest 50. The value of each participant's production was therefore equal to the reported production multiplied by the value of each correctly typed word.

[3]The complete instructions are available upon request.

In the distribution phase, each participant was matched with a sequence of not more than six other players, where no information was given about the outcome of any distributional situations before all proposals were made. The participants were matched with people who had the same working time and with people who had a different working time. All communication was anonymous and conducted through a web-based interface. For each match the participants were given information about the working time, production, and production value of both players. They were then asked to propose a distribution of the total production value for the two players. The proposals had to be rounded off to the closest 25 NOK. When they had made proposals for all the matches, they were given an overview of all their proposals and asked to give a final confirmation. If they regretted some of their proposals, they were allowed to start over again and make new proposals.

When everyone had confirmed their proposals, one proposal was randomly selected to determine the actual outcome for this person. A participant's total earnings were then the sum assigned to this person in the selected proposal plus the participation fee. At the end of the experiment, the participants were assigned a code and asked to write this code together with their bank account number on a claim form. This form was placed in an envelope addressed to the accounting division of the Norwegian School of Economics and Business Administration. The procedure ensured that neither the participants nor the research team were in a position to identify how much each participant earned in the experiment.

We had nine student sessions, each containing a short and a long version of the experiment. Each student was only permitted to participate once. In total, 238 students participated: 82 first-year students, 84-second year students and 72 fourth-year students. We had one alumni session with 57 participants.

In the production phase of the experiment, there were four possible combinations of working time and price. As shown in Table 1, the average productivity of the individuals for each of these combinations was approximately the same. There is thus no evidence that prices had an important effect on effort in the experiment.

[Table 1 about here.]

In the distribution phase, the paired players might differ with respect to the working time, their production per minute and the price they received per correctly

typed word. In general, the fairness principles justify different fair distributions, but they coincide in certain distributional situations. Figure 1 shows that we had only a small number of distributional situations where the players were identical with respect to production per minute.

[Figure 1 about here.]

Given unequal productivity, we can distinguish between four different categories of distributional situations; the two players could differ with respect to the time they worked and the price they received. The various fairness principles would still coincide in some of the distributional situations. First, if the players were identical with respect to both the working time and the price they received, then the strict egalitarian and the choice egalitarian ideals would imply the same fair distribution, as would the meritocratic and the libertarian positions. Second, if the players differed in working time but were given the same price for their production, the meritocratic and the libertarian positions would stipulate the same fair distribution. Third, if the players had the same working time but were assigned different prices, then the choice egalitarian and strict egalitarian fairness ideal would coincide. If the players differed along all dimensions, then none of the fairness ideals would coincide perfectly.

To obtain a better picture of the potential variation in offers implied by the different fairness ideals, Figure 2 shows how the various fairness ideals correlate in all the distributional situations in the experiment. The size of the dots in the scatter plot indicates the number of distributional situations covered by this point. 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. In general, there is substantial variation between the fairness ideals, and thus potentially, they could all contribute in the empirical analysis.

[Figure 2 about here.]

3 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.

3.1 Descriptive statistics

Table 2 presents some key information about individual offers. 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 also observe that the second-year students offered a significantly lower share than the other subject groups and that the alumni offered a significantly higher share than the other subject groups. 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 fair average offer is the same for all the 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 distributional situation where we observed a very high offer to the opponent of 95% of the total income. In this case, player 1 had worked for 10 minutes, produced 100 words and received a price of 0.5 NOK. Player 2, on the other hand, had worked for 30 minutes, produced 850 words and received a price of 1 NOK. Hence, they had 900 NOK to split. Given this information, player 1 acted in accordance with the libertarian fairness ideal and offered player 2 almost all of the total income, 850 NOK.

Interestingly, in 29.5% of the offers, the player proposed more to the opponent than to him- or herself. In 9.3% of the offers, the player proposed everything to him- or herself. As we can see from Table 2, there were substantial differences among the subject groups in the number of offers where the player proposed everything to him- and herself. Such offers were quite rare among first-year students and alumni, while about 15% of the offers from second and fourth-year students left nothing to the opponent.

[Table 2 about here.]

3.2 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 $\mathcal{Y}(\boldsymbol{a}, \boldsymbol{q}, \boldsymbol{p}) = \{0, 25, 50, \ldots, X(\boldsymbol{a}, \boldsymbol{q}, \boldsymbol{p})\}$, and we introduce random variation that is idiosyncratic to each choice:

$$U_i(y;\cdot) = V_i(y;\cdot) + \varepsilon_{iy}/\gamma. \tag{4}$$

The constant γ determines the weight given to the random elements ε_{iy}, which we assume are 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 $\mathcal{Y}(\boldsymbol{a}, \boldsymbol{q}, \boldsymbol{p})$.

This model has a mixed logit structure where each person is characterized by his or her fairness ideal, k, as well as the parameter β_i. We cannot classify individuals by (k, β), but we estimate the distribution of these characteristics. The distribution of fairness ideals is discrete, and we approximate the distribution of β by a log-normal distribution, such that $\log \beta \sim N(\zeta, \sigma^2)$. 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 λ_k denote the share of the subject group that hold the fairness ideal k, the likelihood of an individual i making proposals $(y_{i1}, \ldots, y_{iJ_i})$ from the sets of feasible proposals $(\mathcal{Y}_{i1}, \ldots, \mathcal{Y}_{iJ_i})$ given a parameter vector $\boldsymbol{\theta} = (\lambda_{SE}, \lambda_{CE}, \lambda_M, \lambda_L, \gamma, \zeta, \sigma)$, is:

$$L_i(\boldsymbol{\theta}) = \sum_k \lambda_k \int_0^\infty \left(\prod_{j=1}^{J_i} \frac{e^{\gamma V(y_{ij};\boldsymbol{a}_{ij},\boldsymbol{q}_{ij},\boldsymbol{p}_{ij},k,\beta)}}{\sum_{s\in\mathcal{Y}_{ij}} e^{\gamma V(s;\boldsymbol{a}_{ij},\boldsymbol{q}_{ij},\boldsymbol{p}_{ij},k,\beta)}} \right) dF(\beta;\zeta,\sigma). \tag{5}$$

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 identifies the weight people attach to fairness on average. 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 β and the population shares.

3.3 Structural estimates

Table 3 reports the estimates of the structural model for each of the four subject groups, where the estimate for each of the fairness ideals is the proportion of the participants motivated by this particular fairness ideal.[4]

[Table 3 about here.]

Consider first our estimates for the first-year students. 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 is, in other words, considerable pluralism in where 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 their working time, that is, for a factor that was fully within individual control. Second, more than 70% (strict egalitarians, choice egalitarians and meritocrats) do not find it fair to hold people responsible for the price, which was completely beyond individual control. Finally, 70% of the first-year students (meritocrats and libertarians) consider it fair to hold people responsible for their productivity.

We find it particularly interesting to observe that most participants hold others responsible for their productivity, but not for the price assigned to them. This observation challenges the common view that the responsibility cut only relies on a distinction between factors within and factors beyond individual control. If both productivity and price are seen as factors beyond individual control, then we need a further explanation of why people hold others responsible for one and not the other. We suggest that the explanation might be that people assign importance to

[4]Note that the restriction that a common model applies to all groups can be rejected with $p < 0.001$ using a standard likelihood ratio test.

the distinction between personal and impersonal factors, and are willing to hold others responsible for personal factors even if they are beyond individual control. Alternatively, one may argue that productivity is viewed by the participants as fully within individual control and thus that our estimates show a strong support for choice egalitarianism.

How is the prevalence of different fairness ideals affected by business studies? If we compare columns (1) and (2) in Table 3, there is a striking similarity in the population share estimates. The estimated shares for the second-year students are almost identical to those of the first-year students. Furthermore, there is hardly any difference in the share of strict egalitarians and the share of choice egalitarians in a comparison of the different student groups. However, in a comparison of first-year and fourth-year students, we see that the share of libertarians falls from 28.4% to 15.0%, and the share of meritocrats increases from 41.5% to 57.8%. This result is surprising because it suggests that business studies do not have the effect that one might expect, namely that more students during their studies come to view the libertarian solution as fair.

In order to study the effect of participation in the labor market, we compare the fourth-year students with the alumni group. As we observe from comparing (3) and (4), there are large and systematic differences in the prevalence of the four fairness ideals among these two groups. In particular, the prevalence of libertarians triples from 15.0% to 44.8%. This suggests that deeper involvement in the labor market makes this group more likely to accept inequalities due to differences in prices.

It is also interesting to study how business studies and labor market experience affect the weight people attach to fairness considerations. To understand the effect of the estimated values of γ, ζ and σ, we provide Figure 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.

[Figure 3 about here.]

The general impression from Figure 3 is that business studies do not have much impact on the weight people attach to fairness considerations except for at the 10th percentile. We observe that the expected offer of a hypothetical individual at the 10th percentile of the estimated β distribution is much higher for the first-year students than for the second- and fourth-year students, which is in line with the finding in Table 2 that the share of individuals offering nothing increases significantly after the first year. At the other percentiles, however, the differences are rather minor.

Interestingly, labor market experience appears to have a positive impact on the weight attached to fairness considerations. At all percentiles, the expected offer is higher for the alumni than for the fourth-year students, and roughly in line with what we observe among the first-year students. We should be careful, however, about how to interpret this result because the average income among the alumni is much higher than for the fourth-year students.

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 groups in the experiment. As we can see from Figure 4, the fit is good.

[Figure 4 about here.]

3.4 Robustness tests

Our estimates rely on assumptions about the distribution of β and functional form of the utility function. In order to assess how sensitive our estimates are to our particular specifications, we provide specification tests in Table 4.[5]

[Table 4 about here.]

In specification R1 (in Table 4), we examine whether our assumption of a log-normal distribution of β is crucial to our results. We do so by adding a mass ϕ_0 at the point $\beta = 0$, such that β is now distributed as a mixture of 0 with probability

[5] We do not report new estimates for γ, ζ, σ, since they are in line with what is reported in Table 3.

ϕ_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 choice of y depend on the absolute value of the deviation from the fairness ideal, regardless of the total income produced. We see that this specification does uniformly worse than our preferred estimates in terms of likelihood values. In R3, we let the utility loss depend on the size of the deviation relative to the size of the total income produced. We see that this specification does slightly worse in terms of likelihood in the specification where parameters vary by subject group, -2719.3 versus -2716.0 in our preferred specification. However, for two subject groups (the first- and fourth-year students), the log-likelihood is slightly higher in R3 than in our preferred specification. Because the quantitative changes in estimated population shares are not very large, we do not consider the choice between R3 and our preferred specification crucial to our results.

4 Concluding remarks

The notion of responsibility plays an important role in 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. The aim of the present study has been to conduct an experiment that allows us to examine this question in more detail.

Our analysis shows that there is a large majority of students that do not hold people responsible for the price they were assigned, an impersonal factor beyond individual control. The support for this view was significantly lower among former students with some years of work experience. However, even within this group, a majority of the participants did not hold people responsible for the price.

There seems also to be broad support for the view that individuals should be held responsible for factors under their control. More than 75% of the participants

in all the four subject groups reject a strict egalitarian position. This conclusion is also illustrated by the fact that in 29.5% of the offers in the experiment, the player proposed more to the opponent than to him- or herself.

A main finding in this paper is that the majority of both students and alumni hold others responsible for their productivity. We believe that the more plausible interpretation of this result is that the majority of our participants are ready to hold individuals responsible for personal traits even if they are outside individual control. At the same time, a large majority of this group are not willing to hold people responsible for another factor outside individual control, namely price. This suggests that for many of our participants, the core distinction when drawing the responsibility cut is not between choices and circumstances but between impersonal and personal factors. This finding is particularly interesting given the prominence of the choice egalitarian view in normative theoretical reasoning, and questions whether this view really plays a prominent role in the political debate. Note, however, that this is not to say that choices are unimportant in the responsibility debate. The meritocratic position also justifies holding people responsible for their choices, but the justification is not that choices are under individual control, rather that they are personal characteristics that merit reward.

The second main finding in this paper is that institutions such as education and the labor market appear to shape people's fairness preferences in a way that is not much discussed in the experimental literature. The focus has been on whether institutions make people more or less self-interested, but the equally interesting question is whether institutions affect people's notion of fairness. Our results show that experience with the labor market seems to increase the likelihood that people view the market outcome as fair. Equally interesting, we do not find this pattern in our study of business education. One might think that the strong focus on the efficiency of markets in the teaching of business and economics should make people more inclined to think that people should receive the value of their production, but this is not the case in our group of business students. On the contrary, we see a move away from libertarianism towards meritocratism.

Our study supports earlier studies that find little evidence that business education makes people more self-interested (Frey and Meier, 2005). On average, there is only a slight decrease in the estimated weight that people assign to fairness.

The main effect seems to be that there is an increase in the number of participants that offer the opponent nothing among the second-year and fourth-year students. Interestingly, however, the group of people that offer the opponent nothing almost disappears among the alumni, who also on average seem to assign more weight to fairness considerations than students.

Our results show that 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 further 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 labor market appear to shape individual fairness preferences. Interestingly, the main effect of these institutions seems to be that they change people's notion of fairness, while there are only moderate effects on the weight people attach to fairness considerations.

References

Alesina, Alberto and George-Marios Angeletos, 2005. "Fairness and Redistribution." *American Economic Review*, 95(4): 960–980.

Alesina, Alberto, Edward L. Glaeser, and Bruce Sacerdote, 2001. "Why Doesn't the United States Have a European-Style Welfare State?" *Brookings Papers on Economic Activity*, 2001(1): 187–254.

Alesina, Alberto and Eliana La Ferrara, 2005. "Preferences for redistribution in the land of opportunities." *Journal of Public Economics*, 89(5-6): 897–931.

Andreoni, James and John Miller, 2002. "Giving According to GARP: An Experimental Test of the Consistency of Preferences for Altruism." *Econometrica*, 70(2): 737–753.

Arrow, Kenneth, Samuel Bowles, and Steven Durlauf, eds., 2000. *Meritocracy and Economic Inequality*. Princeton, NJ: Princeton University Press.

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

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

Cappelen, Alexander W., Astri D. Hole, Erik Ø. Sørensen, and Bertil Tungodden, Forthcoming. "The pluralism of fairness ideals: An experimental approach." *American Economic Review*.

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

Dworkin, Ronald, 1981. "What is Equality? Part 2: Equality of Resources." *Philosophy and Public Affairs*, 10(4): 283–345.

Fehr, Ernst and Klaus M. Schmidt, 1999. "A Theory of Fairness, Competition and Cooperation." *Quarterly Journal of Economics*, 114(3): 817–868.

Ferrall, Christopher, 2005. "Solving Finite Mixture Models: Efficient Computation in Economics Under Serial and Parallel Execution." *Computational Economics*, 25(4): 343–379.

Frank, Robert H., Thomas D. Gilovich, and Dennis T. Regan, 1996. "Do Economists Make Bad Citizens?" *Journal of Economic Perspectives*, 10(1): 187–192.

Frey, Bruno S. and Stephan Meier, 2005. "Selfish and Indoctrinated Economists?" *European Journal of Law and Economics*, 19(2): 165–171.

Frohlich, Norman, Joe Oppenheimer, and Anja Kurki, 2004. "Modeling Other-Regarding Preferences and an Experimental Test." *Public Choice*, 119(1-2): 91–117.

Henrich, Joseph, Robert Boyd, Samuel Bowles, Colin Camerer, Ernst Fehr, and Herbert Gintis, eds., 2004. *Foundations of Human Sociality: Economic Experiments and Ethnographic Evidence from Fifteen Small-Scale Societies*. Oxford, UK: Oxford University Press.

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

Nozick, Robert, 1974. *Anarchy, State, and Utopia*. New York: Basic Books.

Piketty, Thomas, 1995. "Social Mobility and Redistributive Politics." *Quarterly Journal of Economics*, 110(3): 551–584.

Rawls, John, 1971. *A Theory of Justice*. Cambridge, MA: Harvard University Press.

Roemer, John E., 1998. *Equality of Opportunity*. Cambridge, MA: Harvard University Press.

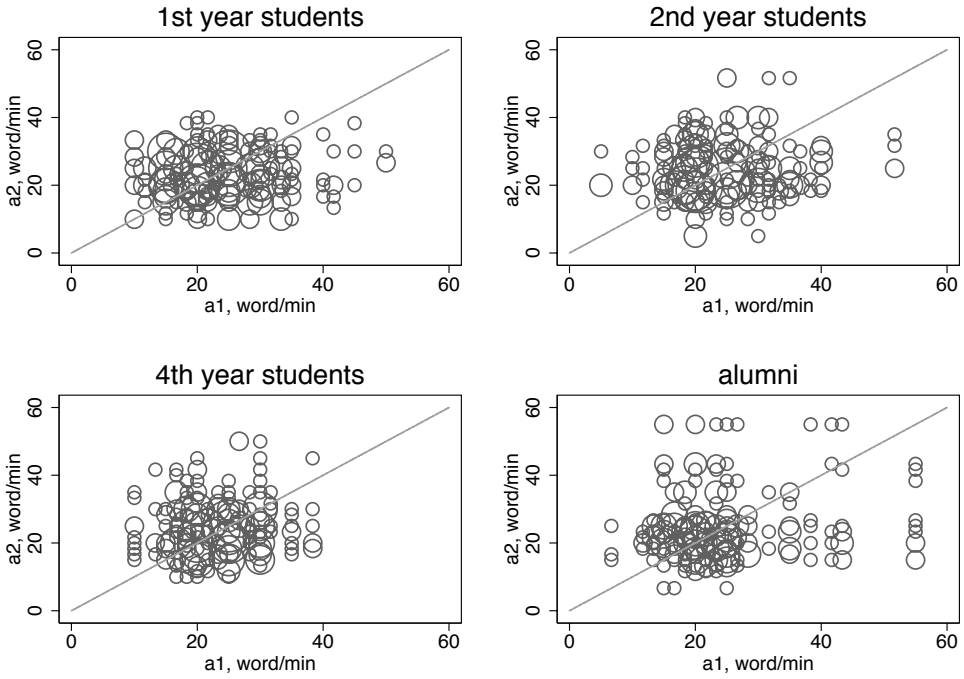

Figure 1: **Distribution of productivity.** The figure plots the productivity of player 1 and player 2 against each other for all distributional situations. The size of the markers indicates the number of situations at that point.

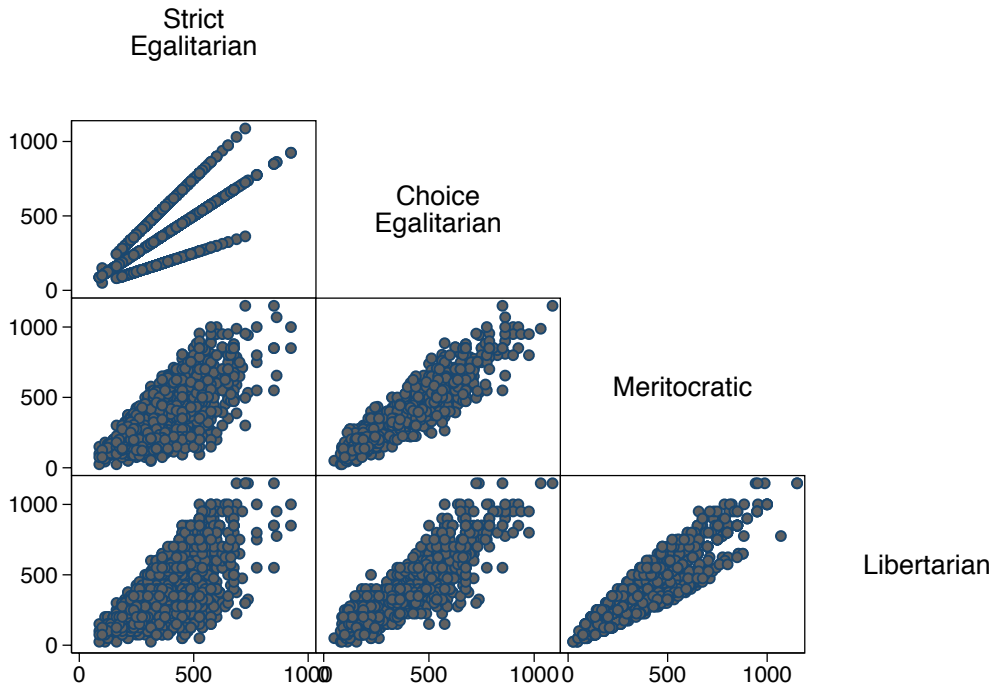

Figure 2: **Scatter plots of fairness ideals.** Pair-wise plots of $m^k(\boldsymbol{a}, \boldsymbol{q}, \boldsymbol{p})$ against $m^j(\boldsymbol{a}, \boldsymbol{q}, \boldsymbol{p})$ for all the distributional situations in our data. The weight of dots indicates the number of observations at that point.

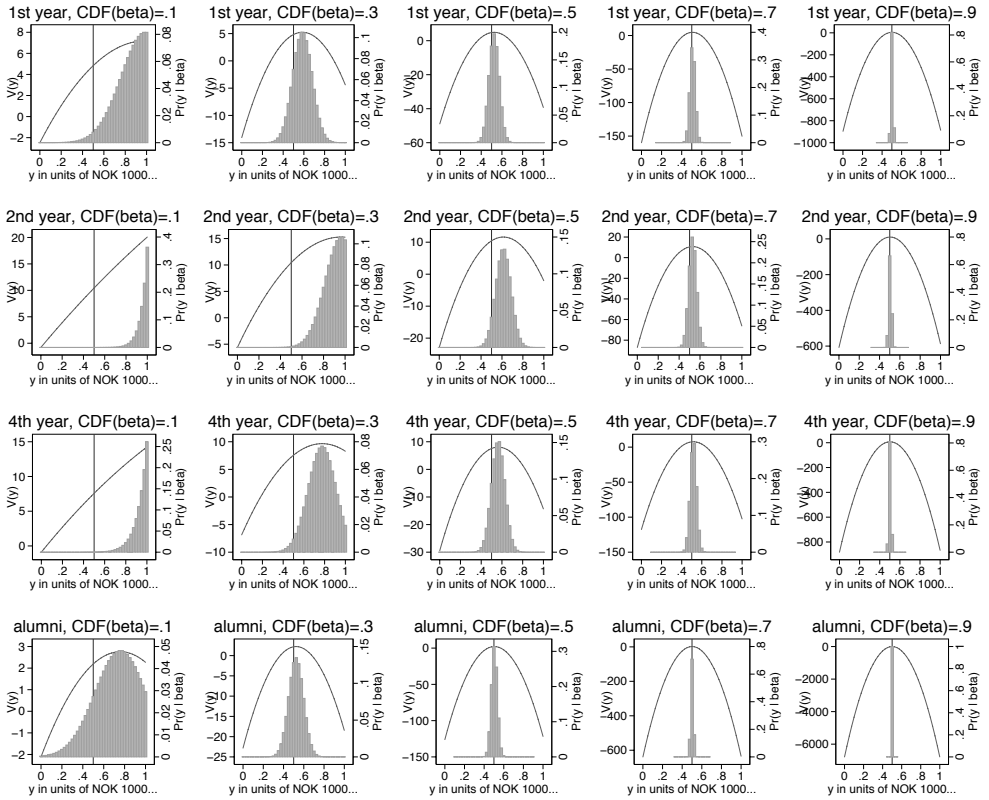

Figure 3: **Illustration of implied choice probabilities.** The graphs plot $\gamma V(y;\cdot) = \gamma(y - \beta(y-m)^2/2X)$, for a hypothetical individual with $m = 0.5$ (marked by a vertical line). Calculated at the 10th, 30th, 50th, 70th and 90th percentiles of the estimated β distribution using the estimates in Table 3. Money, y, is measured in thousands of NOK.

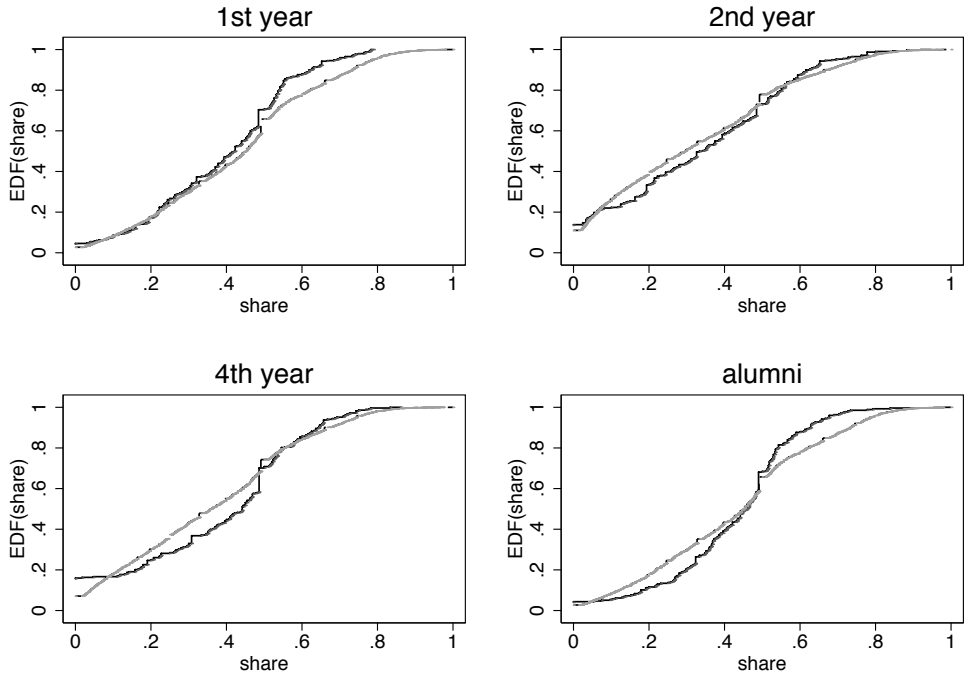

Figure 4: **Empirical distribution function of offers.** Distributions of offers made (as share of total production) and predictions from the estimated model for each experimental group. The black line is our experimental data while the gray line is predictions made from the estimates in Table 3. Predictions are made at the distributional situations in our dataset.

	duration in minutes		
price/word (NOK)	10	30	total
0.5	25.72	25.22	25.43
	$n = 62$	$n = 85$	$n = 147$
1.0	23.28	21.33	22.17
	$n = 64$	$n = 84$	$n = 148$
total	24.48	23.28	23.80
	$n = 126$	$n = 169$	$n = 295$

Table 1: **Statistics on productivity.** The cells show average productivity and the number of observations in each of the four categories.

	Share, $(X-y)/X$				
	1st year	2nd year	4th year	alumni	total
mean	0.396	0.338	0.383	0.430	0.385
standard deviation	0.188	0.233	0.223	0.172	0.208
minimum	0	0	0	0	0
maximum	0.789	0.944	0.862	0.973	0.973
	Amount, $X-y$, in NOK				
	1st year	2nd year	4th year	alumni	total
mean	274	253	266	361	286
standard deviation	189	223	199	203	207
maximum	1000	1000	850	900	1000
share offering nothing	0.043	0.135	0.156	0.040	0.093
share offering more than half	0.296	0.270	0.299	0.319	0.295

Table 2: **Descriptive statistics.** Some key features of offers made to opponent.

	Subject group				
parameter	1st year	2nd year	4th year	alumni	total
λ_{SE}, share strict egalitarian	0.235	0.211	0.213	0.121	0.180
	(0.065)	(0.061)	(0.062)	(0.060)	(0.030)
λ_{CE}, share choice egalitarian	0.066	0.078	0.059	0.023	0.046
	(0.047)	(0.046)	(0.050)	(0.023)	(0.019)
λ_{M}, share meritocratic	0.415	0.421	0.578	0.408	0.470
	(0.078)	(0.080)	(0.088)	(0.084)	(0.042)
λ_{L}, share libertarian	0.284	0.289	0.150	0.448	0.305
	(0.076)	(0.073)	(0.066)	(0.088)	(0.039)
ζ, mean of $\log \beta$	3.675	2.351	2.819	4.912	3.153
	(0.319)	(0.261)	(0.332)	(0.459)	(0.162)
σ, standard deviation of $\log \beta$	2.293	2.502	2.618	3.073	2.672
	(0.298)	(0.297)	(0.306)	(0.445)	(0.168)
γ, inverse weight on ε	9.669	21.358	15.058	4.512	12.432
	(1.547)	(1.987)	(1.253)	(0.785)	(0.618)
Log likelihood of subject group	−777.12	−728.65	−641.61	−568.65	
Total log likelihood			−2716.03		−2754.90

Table 3: **Estimates of structural model.** Standard errors, calculated using the outer product of the gradient (Berndt et al., 1974), in parentheses. The model is estimated with Simulated Maximum Likelihood (with 250 antithetic random draws) using the FmOpt library (Ferrall, 2005).

	Subject group				
	1st year	2nd year	4th year	alumni	total
R1: probability mass, ϕ_0, at $\beta = 0$.					
λ_{SE}	0.248	0.210	0.228	0.128	0.186
λ_{CE}	0.071	0.077	0.059	0.023	0.045
λ_M	0.411	0.424	0.576	0.404	0.471
λ_L	0.270	0.289	0.137	0.445	0.287
ϕ_0	0.111	0.130	0.170	0.049	0.120
$\log L$/subject group	−774.56	−726.50	−634.86	−567.66	
$\log L$/total		−2703.58			−2744.55
R2: utility loss in absolute terms: $V(y;\cdot) = y - \beta(y-m)^2/2$.					
λ_{SE}	0.297	0.219	0.256	0.140	0.217
λ_{CE}	0.078	0.083	0.078	0.025	0.054
λ_M	0.396	0.475	0.562	0.402	0.470
λ_L	0.229	0.223	0.104	0.434	0.260
$\log L$/subject group	−786.41	−739.82	−651.51	−567.30	
$\log L$/total		−2745.05			−2784.42
R3: utility loss in relative terms: $V(y;\cdot) = y - \beta((y-m)/X)^2/2$.					
λ_{SE}	0.180	0.196	0.183	0.110	0.156
λ_{CE}	0.052	0.053	0.041	0.022	0.038
λ_M	0.403	0.498	0.538	0.410	0.451
λ_L	0.366	0.342	0.238	0.458	0.355
$\log L$/subject group	−774.04	−736.07	−637.97	−571.23	
$\log L$/total		−2719.30			−2751.98

Table 4: **Tests of robustness.** The population shares under three alternative specifications of the empirical model.