Principal's Distributive Preferences and the Incentivization of Agents *

Authors: Sophie Cetre,† Max Lobeck,‡

November 23, 2020

Abstract

Do principals' distributive preferences affect the allocation of incentives within firms? We document a robust relationship between French employers' fairness preferences and the incentive contracts they choose for their workers. To establish causality, we run a Principal-Agent lab experiment in a firm-like setting. Principals must choose piece rate wage contracts for two workers that differ in terms of ability. Workers have to choose an effort level that is non-contractible. Principals are either paid in proportion to the output produced (Stakeholder treatment) or paid a fixed wage (Spectator treatment). We study how principals make trade-offs between incentive concerns (motivating workers to maximize output) and their own normative distributive preferences. We find that, despite the firm-like setting and the moral hazard situation, principals do hold egalitarian concerns, but they are sensitive to both extensive and intensive margin incentives. We characterize the heterogeneity in distributive preferences by positing a utility function that incorporates the principal's other-regarding preferences and we estimate it using a finite mixture model. We use these estimates to make counterfactual analyses and show that principals' egalitarian preferences gain in efficiency when workers hold relative income concerns.
JEL classification: D63, C49, C91

Keywords: fairness, distributive preferences, principal-agent, social preferences

* The research leading to the results of this study has received funding from the CEPREMAP and EUR grant ANR-17-EURE-0001. We are immensely grateful to Nicolas Jacquemet, Shachar Kariv and Claudia Senik for particularly helpful comments. We also thank Ghazala Azmat, Béatrice Boulu-Reshef, Andrew Clark, Stefano DellaVigna, Jana Friedrichsen, Jeanne Hagenbach, Dorothea Kubler, Fabrice LeLec and Alvin Roth for the insightful discussions. We also thank seminar participants at PSE, ESSEC and Stanford, and conference participants at FAIR-UCSD Spring School in Behavioral Economics, at the European and North American meetings of the Economic Science Association 2019, at the Economic Science Association world meeting 2020 and at ASFEE 2019 for their feedback and comments.

†Sciences Po Paris, sophie.cetre@sciencespo.fr

‡Paris School of Economics, max.lobeck@psemail.eu

1 Introduction

Employers and managers are, first and foremost, citizens with views about what is fair or not. Are these personal preferences interfering with their managerial choices? Several studies suggest that managers' social preferences play an important role in the organization of firms and more specifically in the way incentives are allocated among workers (Bertrand and Schoar, 2003; Bastos and Monteiro, 2011; Cronqvist and Yu, 2017). However, the extent to which these preferences affect firms' performance, and the context in which they are revealed and used to take managerial decisions, remain unclear. Understanding the relationship between managers' fairness preferences and their managerial decisions is important because there are still substantial variations in management practices that are insufficiently understood. These variations cause persistent gaps in total factor productivity across firms, within and between countries (Bloom et al., 2014).

We provide evidence of a robust correlation between the distributive preferences of executive managers and the incentive structures of their firms. We use a French survey of 4,000 employers and executive managers that includes an extensive set of questions related to workers' wage compensations. We show that when managers think that a policy of individualized wages may be unfair, they are less likely to implement performance pay. Of course, reverse causality can explain the result and the correlation could also be driven by strategic concerns, instead of purely normative preferences. Workers may exert less effort in excessively competitive environments, and this can be anticipated by managers. We show that the relationship declines in strength but remains sizeable and statistically significant when we include strategic motives for using or avoiding performance pay, such as the prevalence of unions, whether they think performance pay motivates workers, whether it is likely to create tensions, etc. This correlation is also robust to a wide array of manager and firm specific controls.

Establishing causality in such a context is complicated. Ideally, we would need a random allocation of managers to firms – to ensure that their normative preferences vary exogenously – and to then measure the type of incentive schemes they subsequently implement. A more realistic approach is to consider exogenous shocks on managers' preferences or their disclosure. For instance, some managers may face stronger efficiency incentives than others because their pay is indexed on the company's performance. This implies that an inequality-averse manager would face a stronger conflict between her normative preferences and incentivization concerns, thereby reducing the influence of her preferences. However, incentive schemes for managers vary non-exogenously across firms and self-sorting of managers into firms leads to a reverse causality problem.

To work around these issues, we run a principal-agent lab experiment, randomizing subjects into manager (principal) or worker (agent) positions. Each principal is matched with two workers of differing ability levels. Both workers choose a costly effort level to produce output, and effort is non-contractible. Principals choose between a series of binary piece rate wage contracts for both workers. These piece rates generate a variable pay-for-performance share of labor income. We randomly allocate principals to either a Stakeholder group (principals' income is proportional to the output produced by the workers), or a Spectator group (fixed income). The Spectator group makes the moral hazard situation irrelevant since the principal no longer has an incentive to maximize output.

So Spectators can implement their preferred income distribution at no cost, which gives us a measure of the distribution of income principals believe is fair. In the Stakeholder group, principals must take into account workers' incentives if they want to increase joint output and maximize their own income. This gives us a measure of principals' willingness to pay for implementing their preferred distribution. The difference in behavior between these two groups isolates normative distributive preferences at the extensive margin. The comparison across treatment groups also characterizes the possible effects of institutional factors such as competitive pressure through market forces on the importance of distributional concerns in incentivization decisions.

Moreover, our framework allows us to pin down the relative importance of various fairness ideals (egalitarian, efficiency-minded, and equal-procedure) among principals. Piece rate wage contracts are an innovation compared to the existing literature because comparing the piece rates chosen for each worker, depending on their ability level, leads to direct classification into three distributive preferences types. Choosing to reward the high ability worker with a higher piece rate is evidence of efficiency concerns, since in our setting this approach is output-maximizing if workers best respond to wage contracts. Rewarding both workers with the same piece rate implies to paying them in proportion to the output they have produced. This leads to procedural fairness since both workers are treated equally with the same piece rate. Finally, giving the low ability worker a higher piece rate shows an egalitarian concern, since differences in productivity will be offset. We calibrate these egalitarian contracts in such a manner that if both workers exert the same level of effort, then they are paid the same final total wage. This corresponds to a common situation in real firms, in which both workers are paid the same final wage, despite their different production levels.

The analysis crucially depends on (i) whether or not agents optimally respond to piece rates and (ii) whether principals anticipate such behavior. Before asking principals to choose their preferred wage contracts, we elicit their beliefs concerning workers' responses to piece rates. This provides control over the efficiency-equality trade-off that principals believe they face before workers start working.

We find that despite the firm-like setting and the moral hazard situation, principals do hold egalitarian concerns. On average, they are willing to accept a trade-off between higher output and reduced within-firm inequality. This willingness is significantly lower if principals are Stakeholders (extensive margin incentives) and it is also the case within treatments when there is a large trade-off between efficiency and equality. Stakeholders are also more sensitive to these intensive margin incentives than Spectators. When the alternative to the output-maximizing (high-inequality) contract is the equal piece rate contract (rather than the egalitarian contract), principals are not more likely to choose it on average. This indicates that equality in procedure as such is not seen as a particularly attractive contract characteristic and principals are more interested in distributive outcomes.

We then calibrate a simple utility function that takes principals' other-regarding concerns into account. The estimates for the representative principal suggest that (i) intrinsic motives are 30% as strong as extrinsic motives in maximizing output and (ii) that principals are averse to extreme inequalities. The structural estimates allow us to make counterfactual statements on how the implications of these preferences on firm performance change once we assume that agents hold horizontal social preferences.

Furthermore, we are interested in examining different profiles of principals and identifying which types actually generate inefficiencies in the allocation of incentives. We use a finite mixture model to characterize heterogeneity in preferences. We quantify the importance principals attach to the payoff of high- and low-ability agents, allowing for the variation in these importances according to whether one agent is paid a higher or lower piece rate than the other agent. The Normalized Entropy Criterion (see McLachlan and Peel, 2004, p.214) recommends assigning principals to one of three types: (1) Output maximizers who always favor the contract that maximizes joint output. These principals do not attach any importance to agents' well-being. (2) Strong redistributors who always attach considerable importance to the low-ability agent's income, and (3) an intermediate group that attaches positive importance to the low-ability agent's income if the difference in piece rates becomes too great. We show that all principals in the *Spectator* treatment care to some extent about the distributive consequences of their decisions. On the contrary, 40% of *Stakeholder* principals are classified as output maximizers and are never willing to relinquish income to compress wages. This implies a sizable crowding out of inequality concerns through the provision of extensive margin incentives. Nevertheless, 60% of stakeholders are allocated to either type (2) or (3), suggesting that moral concerns persist on average, even if principals hold a stake in the workers' output. Counterfactual simulations that vary *workers'* other-regarding preferences show that egalitarian concerns are not always associated with a loss in profit for the firm. Sophisticated output-maximizing principals will mimic the behavior of egalitarian principals because they ultimately make the most efficient choices if agents are egalitarian. But when principals are naive and do not update their effort beliefs, then the egalitarian principals perform better for moderate agent inequality aversion levels.

We contribute to the large and growing body of literature that explores the role of social preferences and inequality in the workplace. Managers' preferences have rarely been the main focus in the theoretical, empirical and experimental literature, despite the important consequences of managerial decisions on wage inequality and firm performance.

Our contribution to this literature is threefold. First, we show that principals' normative distributive preferences affect workers' incentive schemes, which in turn affect the firm's overall performance. Second, we establish that these preferences play a variable role in decision-making according to whether the principal has a stake in the firm's outcome. Finally, we take an ex-ante perspective, in the sense that principals choose incentive schemes *before* agents have exerted any effort. We thus take account of the fact that managers typically make decisions in an uncertain environment.

More precisely, we contribute to the experimental literature on social preferences and distributive fairness. This literature studies distributional preferences using relatively abstract dictator games to infer whether subjects' allocation decisions are guided by concerns about selfishness, efficiency, inequality, or maximin preferences (e.g. Engelmann and Strobel, 2004; Fisman et al., 2007). Similarly, allocation games have been used to infer whether subjects are primarily concerned about inequality, or rather inequity (Konow, 2000; Cappelen et al., 2007; Almås et al., 2018). These studies do involve the (re)allocation of income *after* a production stage. Therefore, they do not consider the role played by distributional preferences in contract creation that is decided *before* production occurs. Furthermore, Balafoutas et al. (2013) study the conflict between equality, equity,

and incentives using a public goods game.

The theoretical literature on social preferences in the workplace has incorporated social preferences into principal-agent models with a focus on team production. Bartling and Von Siemens (2010), Englmaier and Wambach (2010), and von Siemens (2010) incorporated workers' envy and social comparisons into the derivation of optimal contracts, and found that this affects the optimal incentive structure. However, principals are modelled as output-maximizers.

Field and lab experiments have shown that agents compare their income horizontally (e.g. Clark et al., 2010; Bandiera et al., 2005; Breza et al., 2017; Cohn et al., 2014; Gross et al., 2015; Eisenkopf and Teyssier, 2013; Abeler et al., 2010), and that they care about being treated equally (Gagnon et al., 2020). Similarly, workers may have social preferences towards principals and reciprocate high unconditional wages with high effort as shown in the gift-exchange literature (Bellemare and Shearer, 2009; DellaVigna et al., 2016; Fehr et al., 1993).

Few papers study how other-regarding concerns may affect the allocation of incentives within a firm. Existing work shows that principals' incentives affect how they allocate their supervision (Bandiera et al., 2007). Principals take into account fairness concerns in a context in which they are matched with a single agent (Fehr and Schmidt, 2004; Fehr et al., 2007). Brandts et al. (2019) study principals' distributive concerns in a gift-exchange setting, where principals' strategic motives are muted. Kocher et al. (2013) show that social preferences correlate with preferences concerning managerial leadership styles. Cabrales et al. (2010) also document a correlation between social preferences and choices concerning contracts but in a setting in which principals have to compete for workers.

The remainder of the paper is structured as follows: section 2 presents survey evidence, section 3 introduces the design, section 4 presents our main results, section 5 presents the structural model, and section 6 concludes.

2 Managers' preferences: survey evidence

We use the two waves (2011 and 2017) of a French survey on Professional Relationships and Firm Negociations (REPONSE), that has been conducted every six years since 1993.[1] The survey was administered in 4,023 firms in 2011 and 4,364 in 2017[2] and three types of questionnaires are distributed, one for a representative of the executive managerial positions[3], one for a personnel representative[4], and one for employees of the firm.

We use the questionnaire dedicated to managers in which they are asked whether workers benefited from individualized pay rises and also whether they received bonuses related to individual performance in 2010 and 2016, for the 2011 and 2017 waves, respectively. The survey asks both questions to white- and blue-collar workers.[5] We use these outcome variables as indicators of whether

[1] We only had access to the last two waves via the *Réseau Quetelet*, as the earliest ones are no longer available.

[2] To be more precise, the survey is conducted at the plant (*établissement*) level, unless the firm has an independent status. We use "firm" in the text for simplicity. These plants are representative of 196,434 plants with 11 employees or more, and approximately 10 million employees in France, according to the 2017 wave documentation.

[3] Either CEO, Secretary General, Plant Director, Head of Human Resources, or another top managerial position.

[4] Either a union representative or a staff representative.

[5] *Cadres* and *Non Cadres* in French.

the firm engaged in pay-for-performance and thus wage (or bonus) differentiation based on effort or ability, for both types of workers. In our experiment, we proxy such kinds of choices by the series of decisions between two binary piece rate wage contracts.

Regarding our main explanatory variable, we use a question to proxy principals' distributive preferences: the questionnaire asks whether the manager believes that individualized wage raises are unfair. Managers who agree with this statement can be categorized as averse to inequalities among their employees.

Columns 1 and 4 of Table 1 use a logit specification to show that there is a strong negative correlation between both variables. Managers who think that individualized wage rises are unfair are 20 (20.9) percentage points less likely to run a company that implements individualized wage rises among white-collar (blue-collar) workers. Obviously, this correlation is likely to suffer from reverse causality or self-selection since managers are not randomly allocated across firms. An omitted variable bias is also likely: this correlation may capture other motives. Principals may answer the fairness question by considering what workers think is fair instead of their own personal distributive preferences. For instance, principals may believe that individualized wages generate tensions among their employees, and might therefore avoid implementing them in order to maintain levels of production and profit. They may declare that individualized wages are unfair, by considering their employees' opinions rather than their own preferences.

Table 1: Individualized wage raises and managers' distributive preferences

	White-collar workers			Blue-collar workers		
	(1)	(2)	(3)	(4)	(5)	(6)
Dep var = Did white/blue-collar workers benefit from individualized wage raises						
Individualized wage raises are unfair	-0.200***	-0.103***	-0.0830***	-0.209***	-0.138***	-0.119***
	(0.0127)	(0.0140)	(0.0151)	(0.0118)	(0.0130)	(0.0145)
Individualized wage raises create tension		-0.146***	-0.0993***		-0.116***	-0.0881***
		(0.0102)	(0.0110)		(0.0101)	(0.0113)
Individualized wage raises motivate		0.0678***	0.0579***		0.0425***	0.0317*
		(0.0165)	(0.0178)		(0.0156)	(0.0174)
Individualized wage raises are subjective		-0.111***	-0.0418***		-0.0926***	-0.0650***
		(0.0120)	(0.0131)		(0.0114)	(0.0128)
Wave dummy	Yes	Yes	Yes	Yes	Yes	Yes
Individual controls	No	No	Yes	No	No	Yes
Firm controls	No	No	Yes	No	No	Yes
Observations	7666	7566	5771	8139	8033	6156
Pseudo R^2	0.026	0.069	0.185	0.028	0.055	0.104

Robust standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The Table displays marginal effects from logit specifications. We regress a binary variable for whether white-collar workers benefited from individualized wage raises in columns (1) to (3) (blue-collar workers in columns (4) to (6)) on binary variables for whether the manager answering the survey thinks that individualized wage raises are unfair, whether they create tensions, motivate, or are subjective. A 2017 wave dummy is added in all regressions. We additionally control for individual and firms' controls in columns (3) and (6). See Appendix Table A1 for a description of all the variables.

Fortunately, the survey is extensive enough to control for such beliefs. To isolate normative preferences as much as possible, in the rest of the columns we control for other strategic concerns that may lead the firm to avoid or adopt performance pay, such as the belief that it motivates

workers, can create tensions, or is difficult to base on objective criteria.[6] In columns (3) and (6), we also control for the socio-demographic characteristics of the manager answering the survey and a large set of firm-related controls, in order to minimize the self-selection issue.[7]

We see that the normative distributive preference variable (individualized wage raises are unfair) still leads to a negative and significant coefficient, even after the inclusion of all these controls. It is similar in magnitude to the coefficient associated with the belief that individualized wages generate tensions within the firm.[8] Appendix Table A2 shows the same regressions but focuses on the implementation of performance-based bonuses, rather than pay raises, as the dependent variable. Results are similar in sign and significance.

This representative survey of French managers reveals a robust correlation between the implementation of individualized wage raises or bonuses and normative distributive preferences. This survey indicates that managers in real situations are sometimes willing to abandon a (theoretically) efficient tool to increase production (performance pay) in order to avoid a conflict with their own normative distributive preferences. We use this evidence as a starting point to investigate the causal impact of principals' distributive preferences in a more controlled lab environment.

3 Experimental design

3.1 Lab setting

Each session of our laboratory experiment consists of 18 to 24 subjects that are randomly assigned as either an agent or a principal at the beginning of the session. Furthermore, each principal is randomly matched with two agents and the groups and roles are fixed throughout the experiment. The experiment is framed as an interaction in a firm, which is the most natural setting in which principal-agent interactions and wage distribution take place (see Alekseev et al., 2017, for a discussion on contextual instructions). Agents are called "workers" and principals are called "Managers".[9] We inform all participants that the currency used during the experiment is the ECU with the following conversion rate: 1€ = 10 ECU. The detailed instructions (translated from French to English) are included in Appendix D.

We ran the experiment at the Laboratoire d'Economie Experimentale de Paris between December 2018 and January 2019. All sessions were in French with French-speaking subjects who were recruited using ORSEE (Greiner, 2004). Sessions were computerized using zTree (Fischbacher, 2007), average payments were 15 € and sessions lasted 90 minutes, on average. Overall, 339 sub-

[6] Appendix Table A1 describes all these variables, including an English translation of the original questions.

[7] Of course, we do not claim that such an empirical strategy is sufficient to establish a causal relationship.

[8] Whether managers believe that individualized wages motivate workers is smaller in magnitude and only significant at the 10% level for blue-collar workers. This can be explained by the fact that 90% of the sample agrees with this affirmation. Hence, there is practically no variation in the answer to this question. It also highlights that managers generally believe in the motivating power of monetary incentives. This is also in line with a more general view that good management practices include rewarding high individual performance Bloom et al. (2014)

[9] We use the French word "*gérant*" rather than "manager", which is also frequently used, in order to avoid any confusion stemming from the possible negative connotations of the word "manager" in French (it is sometimes related to being "bossy"). "*Gérant*" is the French translation of manager and has a more neutral connotation. Moreover, the principal in our case is also an employee of the firm. Hence, using the words "employer" and "employees" could be misleading.

jects were invited in groups of 18, 21 or 24 subjects.[10] 226 participants were randomly assigned to the worker role and 113 to the principal role.

3.2 Workers

Production and cost functions Workers are invited to make consecutive effort choices for a number of piece rates. Their income is composed of a fixed share of 90 ECU (9€) and a variable share that depends on the piece rate they are paid as well as their induced production function. They are informed that an (anonymous) principal will choose a piece rate for them that will determine the variable share of their wage. The latter takes the general form $y_i = \alpha_i e_i$, where e_i is the effort level chosen by the worker and α_i is the marginal productivity which varies across workers ($\alpha \in \{\alpha_H, \alpha_L\}$). In all sessions, we define $\alpha_H = 60$ and $\alpha_L = 40$.

α_i is allocated according to the subjects' performance at an aptitude test that the workers take after receiving the instructions about the workplace setting described above, and after completing a comprehension test.[11] They are informed that performing better at the aptitude test will increase their chances of having higher productivity (a high α). Using an aptitude test to generate heterogeneity in productivity across agents in a stated effort experiment has been used in gift-exchange experiments to justify induced productivity differences (Bolton and Werner, 2016; Gross et al., 2015). The idea is to overcome a certain arbitrariness in productivity differences by creating a link between induced and real ability that would not exist under random ability allocation. We do not want principals to think that ability is completely arbitrary. The aptitude test consists of nine questions: three logic questions, three French questions and three general knowledge questions. The French and logic questions were simplified versions of TAGE MAGE, a French equivalent of GMAT. Workers have 10 minutes to complete a practice test (same format but different questions) and then have five minutes to complete a test that will define their production function.[12] Ability is determined at the pair level. We assign α_H to the worker with the best performance within the pair and α_L to the other one.

The cost function is constant across agents and it is convex in effort choices. Figure A1 in the appendix displays the production and cost function of both workers.

Workers' decisions The agents make effort choices for *all* piece rates that can be chosen by the principal. As is common in the strategy method, they are informed that the principal will only choose one of their choices as payoff-relevant.[13]

[10] Since the design of the experiment was based on a group composed of a principal matched with a pair of workers, the number of participants was a multiple of 3 in each session. Variation in participants per session stemmed from differences in the show-up rate.

[11] To ensure that all participants understand the experiment, they take an extensive comprehension test that asks them to explore the environment. The questions are designed to ensure that they understand the consequences of their decisions. Section E describes this test further and how the subjects performed.

[12] Appendix F includes the questions. The practice test is simply meant to allow them to evaluate the type of questions they will encounter and keep them occupied while principals progress through the experiment. Agents receive no feedback on this practice test.

[13] One could argue that workers may themselves form beliefs regarding which piece rate is more likely to become payoff-relevant. This is unlikely to happen in our setting since from the worker's point of view, the principal's objective function is unknown. First, they do not know that principals choose piece rates for two workers at the same time. Second, they are not informed about how principals are paid.

Piece rates range from 0.3 to 0.70 ECU (for high-ability workers) and from 0.3 to 0.75 ECU (for low-ability workers) in increments of 0.05. It is possible that workers will react differently to a certain wage if the previous piece rate was higher or lower. Nonetheless, we decided not to completely randomize the order applied to the workers because it is unfeasible to robustly identify order effects under complete randomization (81 possible combinations would need to be compared). However, we test for order effects by looking at two benchmark cases: (1) ascending order of piece rates starting at 0.3 and ending at 0.7 ECU; and (2) descending where the order is reversed. One of the order is randomly assigned to each worker.

Workers choose effort levels from a discrete set between 0 and 5 ($e \in \{0, 0.5, 1, ..., 5\}$). We elicit effort decisions for *all* piece rates. The final income of the worker is $\pi_i^w = pr^m \alpha_i e_i - c(e_i) + 9$ €, where pr^m is the piece rate chosen by the principal and $c(.)$ is the effort cost function.

A screenshot of agent B's decision can be found in the appendix, Figure A2. For each piece rate, workers can view their production table showing how each effort level translates to production, effort cost and net variable income. To ease the cognitive burden, we show them a simulation of the consequences of their decision when clicking on a particular effort level. For instance, when effort level 3 is selected, the screen shows the worker's production output (180 units), the current piece rate (0.5 ECU), the cost (48 ECU) and the net income (42 ECU) associated with such an effort level.

Workers' information set Workers are informed that the payoff-relevant piece rate will be chosen by a principal but they are not informed that this principal also chooses a piece rate for another worker. We chose this feature of the design to avoid horizontal wage comparisons among workers that could lead them to sabotage very unequal piece rates on the basis of their own fairness motives. Since we want to focus on the principals' reaction to wage inequality among workers, we want to eliminate other, possibly confusing, factors from the principal's decision, as far as possible.

Furthermore, workers are not informed how their decisions affect the principal in order to avoid vertical social preferences that have been documented in the field (Ashraf and Bandiera, 2018; DellaVigna et al., 2016). Since the remuneration of principals is our main treatment variation, we want workers' effort decisions to be orthogonal to the treatment.

3.3 Principals

Each principal is matched at the beginning of the session with two workers, and different ability levels are assigned to them on the basis of the aptitude test. Each worker is randomly assigned a neutral label – "Worker A" or "Worker B" – and we present a table summarizing both workers' characteristics in terms of productivity (how much output they can produce for a given effort level) and cost function (see Figure A1). Labels A and B are randomized and thus independent of the ability level.[14] This neutral labeling implies that we never tell the principal which subject is more

[14]However, Worker A's characteristics are always summarized in the left-hand table. Starting with Worker B on the left would have been puzzling for many subjects

productive; they can infer this on their own from the information disclosed in the tables.[15]

Belief elicitation Principals are invited to choose wages for the pair of workers they are matched with. Prior to making these decisions, we elicit their beliefs about the effort level chosen by the workers for *each* piece rate they could possibly implement. We elicit beliefs regarding each worker's effort sequentially to avoid asking too many questions at once. The workers' order of appearance (either Worker A or Worker B) is randomized at the principal level. At the end of the experiment, we randomly draw one guessed belief, and if the principal's guess is correct she receives 10 ECU (1€).[16] The drawing of the payoff-relevant piece rate in the belief elicitation is completely independent from the drawing of the payoff-relevant choice in the latter part of the experiment in order to achieve independence in the decisions across the two parts.

Belief elicitation of workers' effort choices plays a vital part in the experiment. It enables us to determine whether an egalitarian contract choice originates from normative distributive preferences or different beliefs regarding how workers should behave under each contract. Principals may believe that workers do not seek to maximize their own income and would choose different effort levels instead of the best responses. Under such a belief structure, an egalitarian contract may become optimal. In other words, eliciting beliefs enables us to determine whether our classification of output-maximizing contracts is also shared by principals or not.

Contract decisions After the belief elicitation part of the experiment, the principals make 16 binary decisions between two contracts, where each contract consists of two piece rates (one for the more productive worker $Worker_H$ and one for the less productive worker $Worker_L$). The choices are summarized in Table 2, showing the piece rates associated with each decision, as well as the distributive and productive consequences of each option (conditional on the workers best responding to the piece rate). The choices ask the principals to decide between an efficient (the total output is maximized) and an egalitarian contract ($Worker_L$ receives a higher piece rate compared to $Worker_H$), or an equal piece rate contract (both workers receive the same piece rate). The equal piece rate contract pays the workers in proportion to their output level, while the egalitarian contract assigns a higher piece rate to the lower-productivity worker to ensure that the workers are paid in proportion to their *effort level*. If workers best respond to wages, the egalitarian piece rate will either perfectly equalize income levels (Choices 1-3 and 8-10) or significantly decrease inequality (Choices 4-7). Note that we label Contract 2 as "output-maximizing", although it is not always efficient, even if workers best respond to it (see Choices 4 and 13). We also want to test also for situations in which the egalitarian or equal piece rate contract is output-maximizing to avoid positing that equality is always desirable. Some people may consider that ability-induced inequality is fair. However, Contract 2 always leads to larger inequality when workers best respond. For the

[15] In the comprehension test, we asked them to find out which worker was the most productive in a hypothetical situation (table with completely different production and cost function). See Appendix E for more details regarding the comprehension test.

[16] We are aware that this is a very simplistic way of eliciting beliefs and we measure the modal rather than the mean belief. However, we want to minimize complexity in the experiment and thus opt for a method of incentivizing beliefs that is easier for the subjects to understand.

Table 2: Set of decisions made by the principal assuming workers' choose effort to maximize their own income

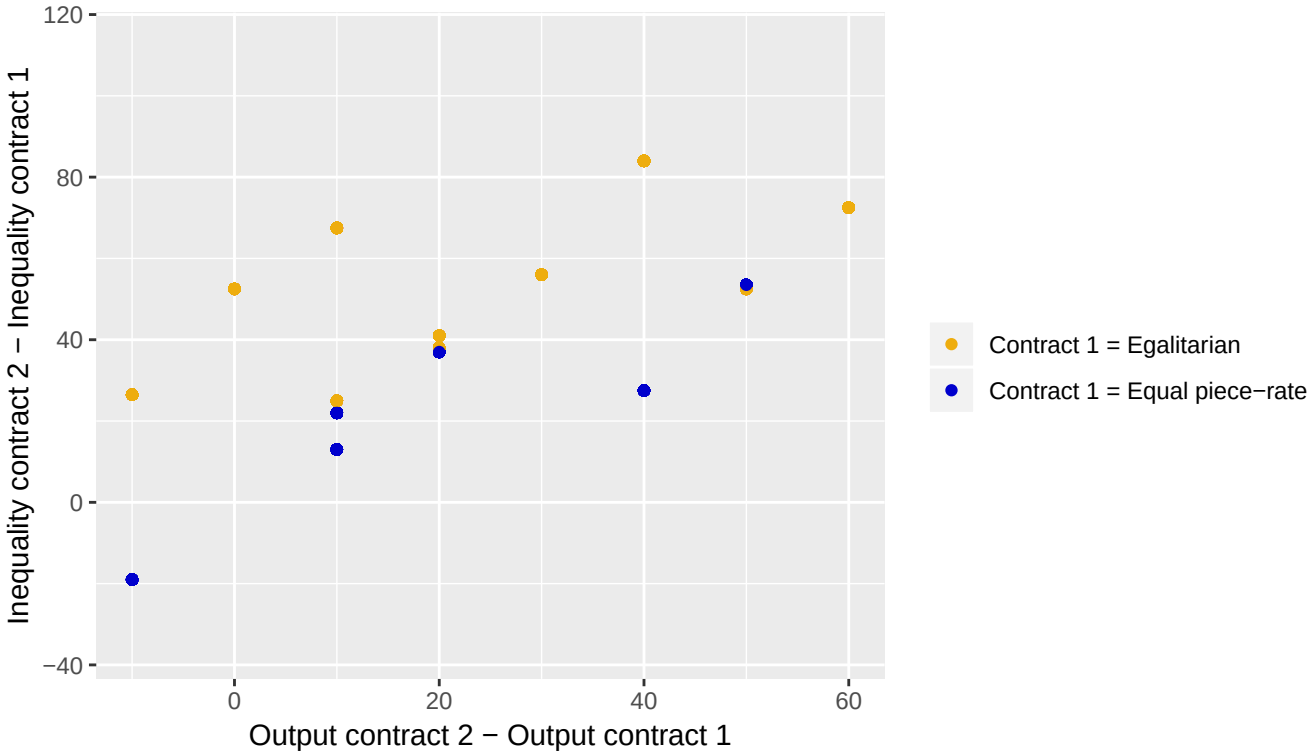
	piece rates		**Income Contract 1**		**Income Contract 2**		**Joint output**		
	Contract 1	**Contract 2**	$Worker_H$	$Worker_L$	$Worker_H$	$Worker_L$	**Contract 1**	**Contract 2**	**Δ-output**
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Egalitarian VS output-maximizing choices									
N°									
1	0.4 – 0.6	0.5 – 0.5	25.5	25.5	42	17	250	260	10
2	0.4 – 0.6	0.55 – 0.45	25.5	25.5	51.5	13.5	250	270	20
3	0.4 – 0.6	0.6 – 0.4	25.5	25.5	63	10.5	250	300	50
4	0.5 – 0.65	0.55 – 0.45	42	30.5	51.5	13.5	280	270	-10
5	0.5 – 0.65	0.6– 0.4	42	30.5	63	10.5	280	300	20
6	0.5 – 0.65	0.65 – 0.35	42	30.5	75.5	8	280	310	30
7	0.5 – 0.65	0.7 – 0.3	42	30.5	90	6	280	340	60
8	0.5 – 0.75	0.6 – 0.4	42	42	63	10.5	300	300	0
9	0.5 – 0.75	0.65 – 0.35	42	42	75.5	8	300	310	10
10	0.5 – 0.75	0.7 – 0.3	42	42	90	6	300	340	40
Equal piece rate VS output-maximizing choices									
11	0.5 – 0.5	0.55 – 0.45	42	17	51.5	13.5	260	270	10
12	0.5 – 0.5	0.6 – 0.4	42	17	63	10.5	260	300	40
13	0.55 – 0.55	0.5 – 0.65	51	20.5	42	30.5	290	280	-10
14	0.55 – 0.55	0.6 – 0.4	51	20.5	63	10.5	290	300	10
15	0.55 – 0.55	0.65– 0.35	51	20.5	75.5	8	290	310	20
16	0.55 – 0.55	0.7– 0.3	51	20.5	90	6	290	340	50

Notes: This table shows the series of decisions principals are asked to make. All units in columns (1)–(6) are in ECUs. The units in columns (7)–(9) are production quantities. The first two columns display the piece rates that are associated with each contract. The left-hand piece rate is the piece rate for the most productive worker ($Worker_H$) and right-hand piece rate is for the least productive worker ($Worker_L$) of the pair. The decisions can be split into egalitarian vs output-maximizing and equal piece rate vs output-maximizing choices. Egalitarian contracts result in outcomes proportional to effort. The equal piece rate contracts result in outcomes proportional to output. Columns (3)–(6) correspond to the variable share of income and thus exclude the 90 ECU show-up fee, common to all workers. The variable income levels (columns (3)–(6)) and the joint output for each contract (columns (7)–(8)) are *conditional on the workers best responding to the piece rate.*

sake of simplicity, we abstract from the two exceptions and retain "output-maximizing" label when referring to Contract 2.

The choices were calibrated so that both inequality and joint output vary across choices, but without a perfect positive correlation. Otherwise, it would have been impossible to disentangle their respective impacts on contract choices. Figure 1 shows how differences in inequality between Contract 2 and 1 (on the y-axis) and efficiency (difference in output between Contracts 1 and 2, on the x-axis) vary across choices. Orange dots represent each case in which Contract 1 is egalitarian and blue dots show when Contract 1 is an equal piece rate (equal procedure) contract. Choices with an egalitarian contract are naturally located at the top of the graph since they lead to a more drastic compression of wages than equal piece rate contracts. The difference in inequality ranges from -19 to 84 ECU and difference in output ranges from -10 to 60 units produced. In ECU terms, the difference in output-based income is twice as small, since each unit of output is sold at 0.5 ECU. Therefore, in ECU terms, we can say that the inequality level varies more than the output level across Choices. This calibration decision is based on pilot data showing that if output differences are too large across Contracts 1 and 2, principals quickly adopt a corner solution in which they maximize income. Consequently, if inequality and output varied on about the same scale, we would

Figure 1: Contract trade-offs assuming best responses



Notes: The Figure plots the theoretical trade-offs (assuming best responses), underlying the 16 contract choices that principals have to make. The y-axis shows the difference in inequality between both contracts, and the inequality of a contract is measured by the high-ability worker's wage minus the low-ability worker's wage. Hence, Contract 2 becomes increasingly unequal relative to Contract 1 as we move up the y-axis. The x-axis is the difference in output between contracts. Contract 2 becomes more efficient relative to Contract 1 as we move to the right-hand side of the plot. Yellow dots represent the trade-offs of equal piece rate contracts vs output-maximizing contracts, and the blue dots represent the trade-offs of egalitarian contracts vs output-maximizing contracts.

not be able to see that people also care about inequality to some extent: all principals would be mistakenly described as selfish income-maximizers. In this study, we focus on the window in which there is a trade-off between maximizing output and equality.

Figure A3 shows how we asked principals to make contract choices during the experiment. The top part of the screen shows the tables summarizing the information for Workers A and B[17], the middle part asks principals to choose between both contracts, and the bottom part simulates the consequences of such a choice, both for the workers and for the principal. This simulation part helps to ease the cognitive burden and saves computation time. This simulation is based on the effort belief elicited beforehand. We remind them of the effort level they expect their workers to choose. We then inform them about the expected production associated with such effort levels and the variable income that each worker would receive under the selected contract. The table is updated when the principal selects a different contract. We instruct them to try out both simulations before making a choice.

Since this screen must be repeated 16 times for each of the Choices, we randomize several features to avoid any anchoring biases. The 16 Choices appear in random order at the subject-level. Within

[17]Note that they are not shown the production and cost of each worker for each effort level, only their net variable income. We wanted to avoid overloading the decision table and therefore opted to omit this part from the representation. However, they are told about the composition of the worker's wage in the instructions and comprehension test, and they can access this information by clicking on the description button on the top-right corner of the screen.

a choice, the labeling of contracts as "Contract 1" or "Contract 2" is randomized. This implies that people cannot always choose Contract 2 to maximize their own income. On each occasion, they must check which contract maximizes income with a view to optimizing efficiency. The "Worker A" and "Worker B" labels are randomly assigned to the high-ability and low-ability workers and are thus independent of productivity differences.

Treatments Between subjects (and sessions), we will implement two treatments: (1) the ***spectator treatment*** and (2) the ***stakeholder treatment***.

In the spectator treatment, the principal receives a fixed wage of 20 € that is completely independent of her workers' output. The treatment enables us to identify how normative distributive preferences affect preferences over contracts without any personal and monetary cost for the principal herself. In each decision, the principal is asked to male a make a trade-off between the implementation of an egalitarian (or equal piece rate) and an output-maximizing contract, keeping her own income constant across all the decisions. The size of the trade-off is documented in column (9), if the principals believe the agents are best-responding. The treatment can be seen as analogous to a situation in which principals have no personal stake in the outcome of their organization (e.g. civil servants at the end of their career).

In the stakeholder treatment, the principal receives a fixed participation fee of 60 ECU (6 €) and a variable share from the sales of the output produced by the workers. For each unit produced, she receives 0.5 ECU. She now faces a trade-off between maximizing her own income and implementing an egalitarian (or equal piece rate) contract. By analyzing choice patterns, we can infer from this treatment the price the principals are willing to pay in order to implement an egalitarian or equal piece rate contract. The size of the trade-off depends largely on the principals' beliefs regarding whether or not they expect workers to best respond to the piece rates. This highlights the importance of the belief-elicitation part of the experiment.

3.4 Summary statistics

Table A3 shows the subjects' socio-demographic characteristics by role. Approximately 50% of the subjects are female, the average age is around 25 years old and 60% are students. There are no systematic differences in observed characteristics between workers and principals. Table A4 reports the same statistics focusing on principals only. It shows how our two treatment groups, Spectators and Stakeholders, differ along observed characteristics. Differences are non-significant, except for gender. Despite randomization across treatment groups, Stakeholders are more often female than Spectators. If anything, this bias in our sample should yield more conservative estimates of differences across treatment groups. Women are often found to be more inequality-averse in dictator games (Croson and Gneezy, 2009), which in our case, should lead to a smaller difference in contract choice between Spectators and Stakeholders. Nevertheless, we control for this variable in all our regressions.

4 Results

4.1 Effort choices and effort beliefs

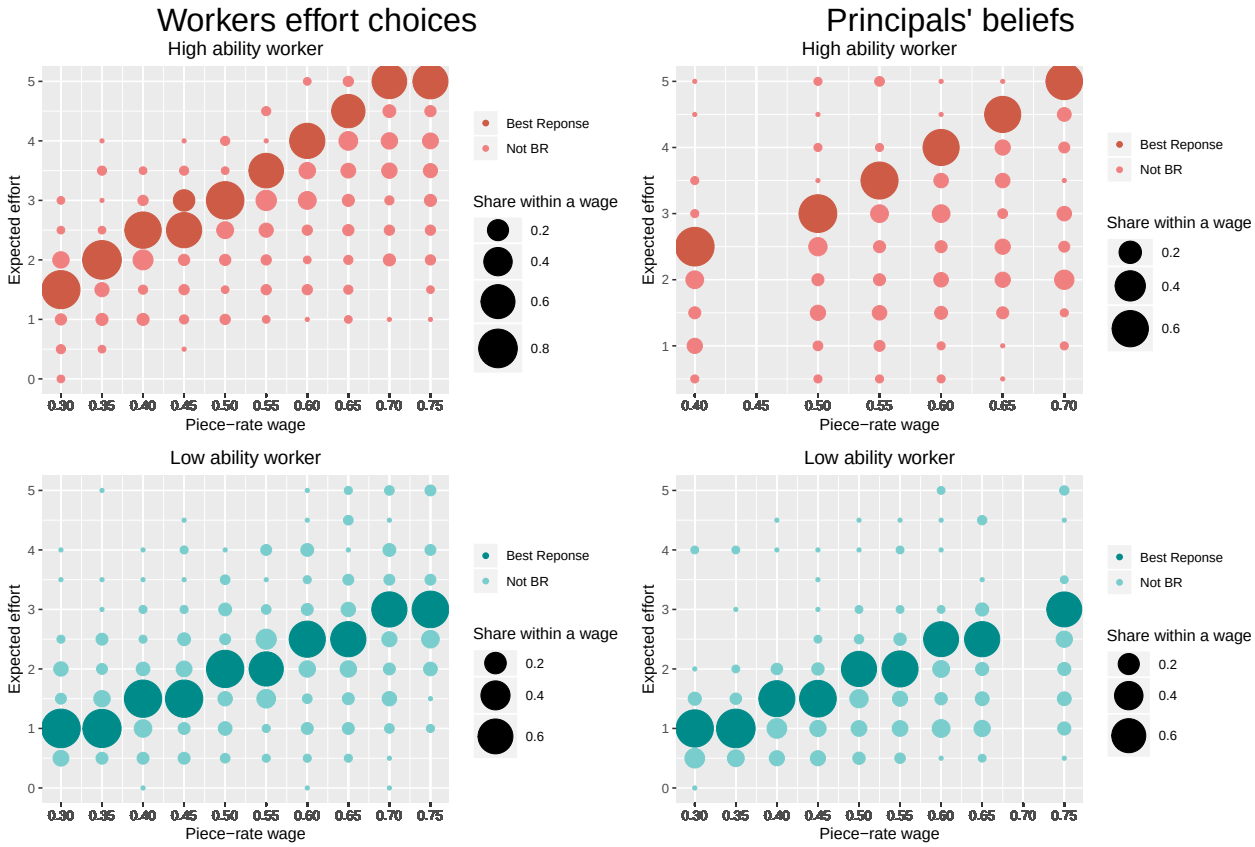
We first describe, side-by-side, the effort levels chosen by workers for each piece rate wage and principals' corresponding beliefs. Figure 2 plots workers' effort choices by ability type (high ability workers in red and low ability workers in blue) on the left-hand side, and principals' beliefs on the right-hand side. For each piece rate wage on the x-axis, we use mass points to display the share of subjects selecting each effort level. Theoretical best responses (effort levels that maximize worker's wage) are reported with a darker color. For instance, we see that around 80% of the high ability workers choose an effort level equal to 1.5 when they are offered a piece rate wage of 0.30, which also happens to be the best response. We find a clear cluster of choices around best responses, both for high-ability and low-ability workers. On average, 67% of low-ability workers and 63.5% of high-ability workers choose a best response effort level. These figures increase to 84% and 82% respectively when allowing for 0.5 deviations (+0.5 or -0.5 from the best response). Conversely, on the right-hand side of the graph, we see that principals often expect workers to best respond. They expect such behavior in 66% of the cases (87% when allowing for 0.5 effort deviation), with no significant differences in beliefs across treatment groups. Principals were also fairly accurate at predicting deviations from best responses. They correctly anticipated that high-ability workers would deviate mostly downward. They expected this type of downward bias for low-ability workers too, but these workers deviated more uniformly either up or down.

4.2 Belief-based contract trade-offs

We now show how these beliefs translate into contract characteristics. The need to create pairs of contracts requiring principals to carry out a trade-off between output maximization and egalitarian concerns guided our contract calibration. Figure 3 shows how principals' expectations regarding workers' effort choices altered these theoretical trade-offs. We interpret the results based on theoretical trade-offs as reduced-form estimates: these trade-offs are exogenous to principals' characteristics. Belief-based trade-offs show how contracts are perceived in reality by principals. This is valuable because we can rely on the true trade-offs principals believe they are facing when making their choices in order to reduce the noise in our estimations. However, these perceptions may be endogenous to principals' characteristics. For instance, certain principals may imagine that low-ability workers will decide to sabotage the experiment and choose a zero-effort level. This particular belief may be correlated to some of the principals' observed or unobserved characteristics. We thus present results using both the theoretical and the belief-based trade-offs to account for these two aspects.

On the x-axis of Figure 3, we plot the difference in output between Contract 2 (the theoretically output maximizing contract) and Contract 1 (an egalitarian or an equal piece rate Contract). On the y-axis, we plot the difference in inequality between Contract 2 and Contract 1. We measure contract inequality as the difference in wages between the high-ability worker and the low-ability worker. Hence, the y-axis is a difference of a difference and a positive number means that Contract 2 yields more inequality than Contract 1. Similarly, positive numbers on the x-axis mean that

Figure 2: Workers' stated effort and principals' expected effort by piece rate wage



Notes: The figures on the left-hand side plot the workers' choices of effort level for each piece rate (on the x-axis) by ability type. The figures on the right-hand side plot principals' beliefs regarding the effort level chosen by workers for each piece rate. High-ability workers are in red and low-ability workers are in blue. Each dot on the figures on the left-hand side represents the share of workers choosing a particular effort level at a given piece rate wage. For example, we see that around 80% of the high-ability workers choose an effort level equal to 1.5 when they are offered a piece rate wage of 0.30. The size of the dots on the figures on the right-hand side represents the corresponding shares for principals. Hence, we see that around 60% of principals expect high-ability workers to choose an effort level of 2.5 when offered a piece rate of 0.40 ECU. Best responses for each piece rate are highlighted in darker colors. Data for several of the piece rates for principals' beliefs is missing. We only elicited principals' beliefs regarding the piece rates that have a chance of being implemented. For instance, the piece rate of 0.45 is never used for the high-ability worker in any of the contracts described in Table 2.

Contract 2 yields a larger output, and therefore income, for the principal, relative to Contract 1. The small black dots represent the theoretical trade-offs, those assuming workers' best respond to piece rate wages. The red and green dots correspond to the belief-based combination of output differences and inequality differences associated with the 16 contract choices facing each principal. We can interpret these dots as the actual trade-offs that principals perceive. The size of the dots represents the frequency of observations implying the same trade-off. Figure 3 shows that many decisions are consistent with our theoretical trade-offs, as expected given the belief-elicitation results in Section 4.1.

We further classify tradeoffs into two types. In green, we identify all the belief-based contract decisions that generate a trade-off between equality and output. In red, we plot decisions for which one of the contracts yields both a larger output and a lower inequality level. 32% of the decisions fall in the red category and do not generate any particular trade-off for people who care about output and want to reduce inequality. However, we do not assume these cases to be irrelevant. For some subjects, it may be fair to over-compensate the high-ability worker. In this case, both inequality and output-maximization would be desirable outcomes and the red dots would represent a real trade-off for these subjects. The finite mixture model can be used to test whether such behavior is common in the data. For that reason, we retain the red decisions in our estimation.

That being said, certain observations remain problematic as the implied trade-offs are too large and constitute outliers. These extreme cases must be discarded in order to avoid distorting our estimates, especially with the finite mixture model. We discard observations for which the difference in output between both contracts is greater than 100 or smaller than -100 (58 out of 1808 observations are deleted). The descriptive results of Section 4.3 are barely sensitive to the inclusion or exclusion of these observations because we show mean contract choices by trade-off brackets. Extreme trade-offs only distort the mean of the far-left-hand and far-right-hand brackets, not the intermediate brackets. However, in the finite mixture model, trade-offs directly enter the objective function and the estimation is quite sensitive to these outliers, though the results remain qualitatively the same. We come back to the issue of outliers in detail in the relevant sections below.

Figure 3: Principals' belief-based contract trade-offs

Notes: The figure plots the trade-off that principals believe must be made. The y-axis shows the difference in inequality between both contracts, and the inequality of a contract is measured by the high-ability worker's wage minus the low-ability worker's wage. Hence, Contract 2 becomes increasingly unequal relative to Contract 1 as we move up the y-axis. The x-axis is the difference in output between contracts. Contract 2 becomes more efficient relative to Contract 1 as we move to the right of the plot. The size of the dots represents the frequency of choices implying the same trade-off. Black dots identify the theoretical trade-offs assuming best responses and are identical to those shown on Figure 1. Green dots show beliefs when there is a trade-off between output and equality, and red dots show cases in which one contract is both output-maximizing and egalitarian given the principal's beliefs (no trade-off).

4.3 Principals' choices

We now describe the pattern of choices across treatment groups. The y-axis of Figure 4 shows the share of cases in which the most egalitarian contract of the pair is selected. This corresponds to Contract 1 in all cases, except for Choice 13.[18] We plot this share by the size of the trade-off: Contract 2 increases in efficiency relative to Contract 1 as we move to the right of the graph. Spectator's choices are plotted with a solid blue line, while Stakeholders' choices are shown with a dotted dark blue line. The top panel shows the choices based on theoretical trade-offs (assuming workers' best responses) and the bottom panel focuses on belief-based trade-offs.

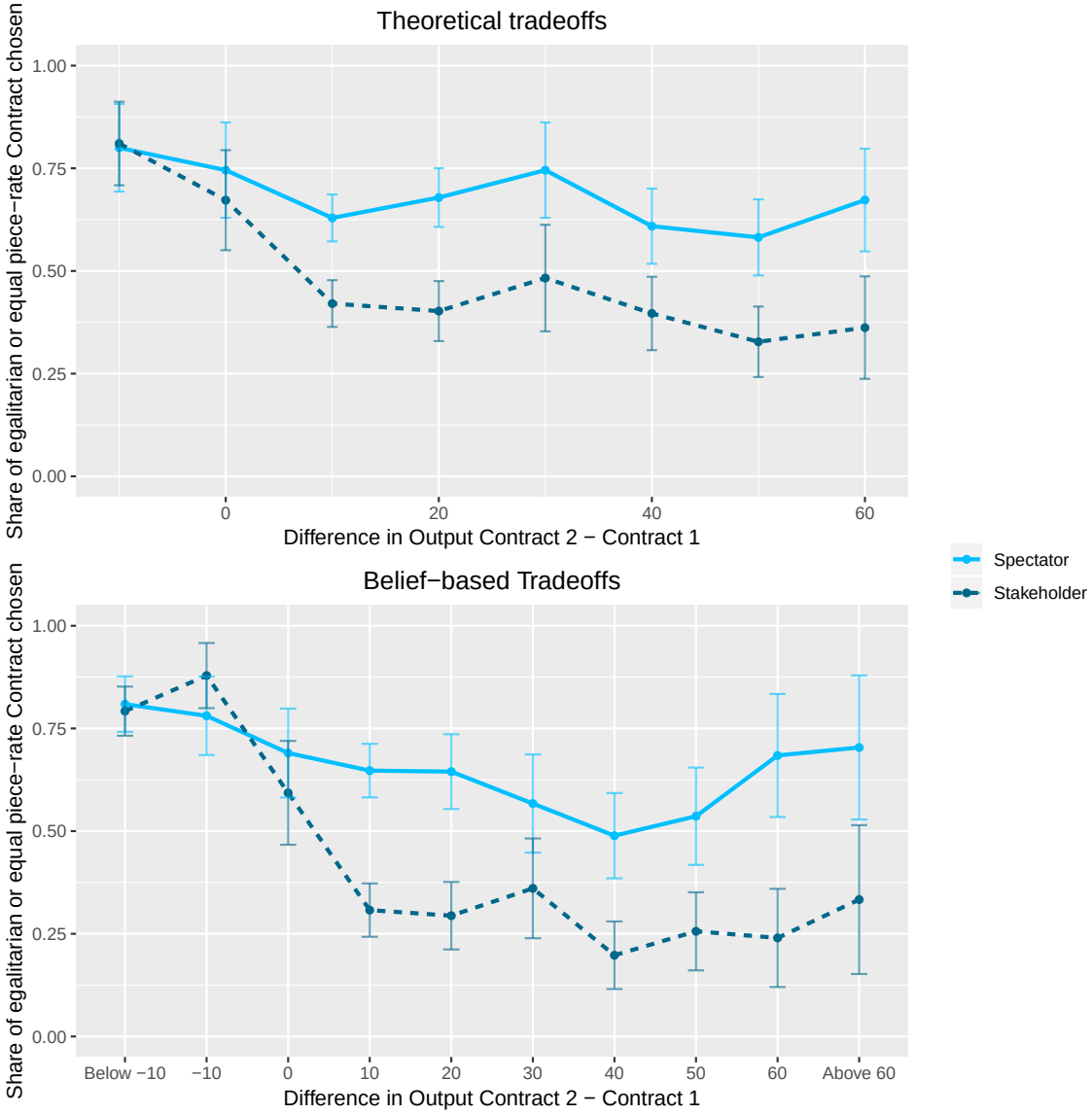
Overall, we find that, on average, both treatment groups compress wages to a certain extent, given that for all trade-offs, the share of Contract 1 decisions is always significantly different from 0. This confirms our hypothesis that, generally speaking, principals hold distributive preferences. Now turning to differences across treatment groups, we find that Spectators are more likely than Stakeholders to choose an egalitarian contract. Interestingly, when Stakeholders do not face any trade-offs (differences in output between both contracts is 0 or even negative), then the behaviors of the treatment groups become indistinguishable. This suggests that Stakeholders are sensitive to the size of the stakes. This is further confirmed when examining their choices at the intensive margin. Stakeholders are increasingly unlikely to choose an egalitarian Contract 1 as Contract 2

[18] For Choice 13, Contract 2 is the most egalitarian contract of the pair.

increases in efficiency in relation to Contract 1. On the contrary, Spectators seem less sensitive to output differences.

Note that the outliers we described in Section 4.2 can only affect the first and end points of the graph (very low and very high expected difference in output). Plotting the same graph without the outliers barely affects the results. If anything, the share of Contract 1 is lower for a high difference in output.

Figure 4: Principals' contract choices by treatment groups



Notes: the Figure shows the share of observations in which the most egalitarian contract of the pair is selected (either an egalitarian or equal piece rate contract). This corresponds to Contract 1 in all cases, except for Choice 13, where Contract 2 is the most egalitarian contract. We calculate these shares by output trade-off, i.e. the difference in output between Contract 2 and Contract 1 (except for Choice 13 where it is the difference in output between Contract 1 and Contract 2). The solid blue line represents the choices of the Spectator group and the dotted dark blue line shows the choices of the Stakeholder group. The top panel shows the choices based on theoretical trade-offs (assuming workers' best responses) and the bottom panel focuses on belief-based trade-offs. We show 95% confidence intervals for the shares.

We can also use a regression analysis to characterize principals' choices. Table 3 regresses the choice of Contract 2 on a binary Stakeholder treatment variable (characterizing the extensive margin trade-off), a dummy "1 is equal piece rate" indicating that the alternative (Contract 1) is an equal

piece rate contract (rather than an egalitarian contract), and the difference in output and inequality between both contracts. The last two variables characterize an intensive-margin trade-off between output and inequality. We interact these variables with the Stakeholder dummy to test whether the intensive margin treatment has a differential effect across Spectators and Stakeholders. Columns (2) and (4) additionally control for two dummy variables: one for whether the observation is about Choice 1 and the other for whether it is about Choice 13. We include these dummies because Choices 1 and 13 involve a *direct* choice between an equal piece rate and an egalitarian contract, and behavior in these decisions may not be captured by the difference in inequality or difference in output. Beyond that, the equal piece rate contract is Contract 2 in this case, and is not picked up by the equal piece rate dummy.

The first two columns of Table 3 calculate these trade-offs assuming that workers best respond (theoretical trade-offs), which can be interpreted as reduced-form estimates that are not biased by the heterogeneity in beliefs. The drawback of these measures is that they may be less precise given that principals may expect deviations from best responses, and therefore a quite different trade-off in reality. Columns (3) and (4) show the results using belief-based trade-offs. The fit is better for the regressions using the belief-based trade-off (the R^2 rises from about 0.1 to 0.17). This indicates that beliefs capture meaningful variations and reduce measurement error in the trade-off principals really face.

The results in Table 3 show that principals are on average significantly more willing to choose a contract if it is expected to yield a larger output relative to its alternative. The decreasing slope in Figure 4 captures this significant effect of the output gap on the Choice probability. This applies to Stakeholders and Spectators alike, but Stakeholders are even more sensitive to this trade-off relative to Spectators (positive and significant interaction term at the 1% level for belief-based regressions). The significant and positive main effect of $\frac{\Delta(\text{Output 2 and 1})}{10}$ indicates that even Spectators want to improve output, on average. Therefore, principals are intrinsically motivated to maximize output and they still respond to changes in the output gap, even after controlling for differences in inequality. We can interpret this result as a residual effect of identity: even if Spectators have no stakes in the production process, they are placed in a managerial position, which can lead them to care about output anyway. These results hold qualitatively for regressions using beliefs (Columns (3) and (4)), as well as those assuming that agents best-respond to incentives (Columns (1) and (2)).

The first row shows that stakeholders are, on average, 26 percentage points more likely to choose a high-inequality contract (coefficient positive and significant at the 5% level with theoretical trade-offs, and at the 1% level for belief-based regressions). Principals are more likely to accept inequality if they are explicitly incentivized, even after taking into account the expected cost of equality, which characterizes the shift in the intercept of the two curves in Figure 4. We will subsequently show that a significantly higher proportion of stakeholders always choose an output-maximizing contract but that no spectators do so. These individuals may characterize the extensive margin differences between the two groups.

Relative inequality between contracts is only a significant predictor if we consider regressions (1)–(3) (significant at the 5 percent level). In these instances, principals are less likely to choose a contract that involves greater inequality after controlling for the difference in output. This effect

Table 3: Regressions that characterize Contract decisions

	Theoretical trade-offs		Belief-based trade-offs	
	(1)	(2)	(3)	(4)
Dependent variable: Contract 2 (high inequality) was chosen				
Stakeholder	0.177**	0.177**	0.266***	0.265***
	(0.0837)	(0.0838)	(0.0771)	(0.0780)
$\frac{\Delta(\text{Output 2 and 1})}{10}$	0.0467***	0.0433***	0.0316***	0.0302***
	(0.0129)	(0.0132)	(0.00667)	(0.00653)
$\frac{\Delta(\text{Output 2 and 1})}{10}$ * Stakeholder	0.0300*	0.0300*	0.0339***	0.0339***
	(0.0173)	(0.0173)	(0.0113)	(0.0112)
$\frac{\Delta(\text{Inequality 2 and 1})}{10}$	-0.0483***	-0.0366**	-0.0218**	-0.0133
	(0.0138)	(0.0144)	(0.00994)	(0.00996)
$\frac{\Delta(\text{Inequality 2 and 1})}{10}$ * Stakeholder	-0.00171	-0.00178	-0.0180	-0.0178
	(0.0184)	(0.0184)	(0.0144)	(0.0145)
1 is equal piece rate	-0.0864*	-0.0461	-0.0263	0.00466
	(0.0451)	(0.0448)	(0.0415)	(0.0407)
1 is equal piece rate * Stakeholder	0.00669	0.00619	-0.00421	-0.00462
	(0.0637)	(0.0637)	(0.0543)	(0.0545)
Choice 1 = 1		0.123***		0.102**
		(0.0431)		(0.0468)
Choice 13 = 1		0.0691		0.0792*
		(0.0427)		(0.0463)
Constant	0.476***	0.409***	0.428***	0.375***
	(0.0873)	(0.0872)	(0.0793)	(0.0787)
Control variables	Yes	Yes	Yes	Yes
Observations	1750	1750	1750	1750
R^2	0.102	0.105	0.165	0.167

Standard errors clustered on the subject level in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The specification regresses a dummy indicating the choice of Contract 2 on other Choice characteristics using a linear probability model. This samples excludes observations where the difference in *expected* output is less than or equal to 100. In columns (1) and (2), explanatory variables include a Stakeholder treatment dummy variable, the theoretical difference in output between Contract 2 and 1, the theoretical difference in inequality between Contracts 2 and 1 (both assuming workers' best responses), a dummy for whether Contract 1 constitutes an equal piece rate contract rather than an egalitarian contract, and the interactions of these variables with the Stakeholder dummy. In columns (3) and (4) principals' beliefs are used to calculate the difference variables. Columns (2) and (4) add controls for whether the current observation concerns Choice 1 (Choice 1 = 1) or Choice 13 (Choice 13 = 1). All the specifications include the following controls: female dummy, economics background dummy, whether the subject is currently a student and whether he is currently in a relationship.

becomes insignificant once we control explicitly for a choice of 1 or 13 and use belief-based trade-offs, which indicates that this may pick up a peculiarity characterized by these two choices. The interaction term between difference in inequality and the Stakeholder dummy is not significant for both theoretical trade-offs and belief-based trade-offs.

The alternative of an equal piece rate contract (rather than an egalitarian contract) is not a significant predictor of the principal's decision once we take into account the characteristics of the contract such as expected inequality and expected output. This does not mean that principals never choose the equal piece rate contract; it simply means that they are not more likely to choose an equal piece rate than an egalitarian contract after controlling for differences in output and inequality. This suggests that subjects are more interested in implementing their preferred outcome rather than treating both agents identically.

The only instance where equality in procedure seems to make a difference is when we ask subjects to choose *directly* between an egalitarian and equal piece rate contract. In this case, subjects are significantly more likely to choose the equal piece rate contract, as suggested by the positive and significant dummies of Choices 1 and 13 (significant on the 5 and 10 percent level). This shows that on average, subjects are more likely to choose an equal piece rate contract if it is posited directly against an egalitarian contract, even after taking contract characteristics into account.[19]

To sum up, the pooled results show that principals are increasingly willing to accept inequality as the cost of the egalitarian contract rises. Average sensitivity to difference in output is relatively higher for stakeholders than spectators. Furthermore, Stakeholders are significantly more likely to choose a high inequality – high output contract at any given level, suggesting a strong extensive margin effect of incentives on inequality acceptance. Although making Contract 1 an equal piece rate contract does not seem to affect how principals evaluate these contracts, they are significantly more likely to choose an equal piece rate contract if it is posited against an egalitarian contract.

Table A5 shows the results for belief-based trade-offs that control for individual fixed effects. This is an additional way to account for individual-specific heterogeneity in beliefs. The results are more or less the same.[20] Table A6 replicates Table 3 but excludes observations where the characteristics of the contracts imply that Contract 2 yields a lower output relative to Contract 1, i.e. $\frac{\Delta(\text{Output 2 and 1})}{10} \leq 0$. This enables us to show that the intensive margin results do not only reflect a threshold effect, i.e. the point at which reducing inequality (choosing Contract 1) comes at a cost to output. We can see that the coefficient of $\frac{\Delta(\text{Output 2 and 1})}{10}$ is still positive and significant for both theoretical and belief-based tradeoffs. Finally, Table A7 replicates Table 3 but includes belief-outliers, i.e. observations where the absolute difference in output is higher than 100, which constitute 3% of the total sample. The results are qualitatively very similar but the interaction term of difference in output and being a stakeholder becomes insignificant and the magnitude of the main effect is attenuated. Given the drop in the R^2 it can be assumed that these differences are largely driven by measurement error in outlier-beliefs and do not reflect systematic variations in behavior.

[19]The individual fixed effects regressions in Table A5 suggest that this effect is mainly driven by stakeholders.

[20]Note that there is no need to control for individual fixed effects with theoretical trade-offs since there is no individual-level variation in trade-offs in that case. Theoretical trade-offs are completely exogenous to individual characteristics.

5 Structural Characterization of Distributive Preferences

In this section, we will estimate the distributive preferences of principals and characterize the heterogeneity in these preferences. The goal of this exercise is to perform a counterfactual analysis that will allow us to assess *when* these preferences lead to frictions and inefficiencies. To this end, we posit a simple social preference utility function that captures several motives.

5.1 Making distributive decisions ex-ante

Before specifying the actual utility function that we want to estimate, it is worth re-emphasizing the context in which managers make decisions. While most studies on distributional preferences take the ex post perspective – dictators make distributive decisions *after* agents have worked, as in most dictator games with a preceding production stage, e.g. Cappelen et al. (2007), we are taking account of the fact that principals typically make incentivization – and hence distributive – decisions in an uncertain environment, before agents have exerted any effort.[21] It also enables testing for the importance of treating *unequal* agents equally, which has not been explored previously.

We assume that principals make decisions that maximize their expected utility $E\big(U(y_p, \pi(e_h(w_h), e_l(w_l), y_h(e_h(w_h)), y_l(e_l(w_l)), w_h, w_l)\big)$, where the principal's income is denoted by y_p and π is the agents' joint output, which is a function of w_h, w_l, the workers' piece rates, and e_h, e_l their effort levels (for the low- and high-ability agent respectively). The agents' ex post income is denoted by y_l, y_h, and also depends on the piece rates and effort level chosen. This specification enables principals to care about the distribution of income *after* workers have made their effort decisions, i.e. ex post income as a function of expected effort. It also enables principals to care about equality of *procedure*: in this case, principals dislike differences in piece rates. Note that our notion of equal procedure is somewhat different from that considered in previous work studying social preferences in a risky environment (e.g. Brock et al., 2013; Krawczyk and Le Lec, 2010), because agents are not identical to begin with (different ability levels leading to different effort levels), which implies ex post inequality even when both agents are treated equally with equal piece rates.

5.2 Utility function specification

The utility function characterizes principals' concern about their own income y_p, total output $\pi = \pi_h(e_l(w_l)) + \pi_l(e_h(w_h))$, and the distribution of income ex post between both workers $y_l(e_l(w_l)), y_h(e_l(w_h))$. The low-ability worker's income is denoted as $y_l(e_l(w_l))$ and the high-ability worker's income is denoted as $y_h(e_l((w_h)))$. In the following explanation, we will refer to these incomes as y_l and y_h with the dependence on the piece rates dropped for expositional purposes, but the reader should bear in mind that workers' income is always a function of their piece rate and their subsequent effort decision.

[21] Whether this feature matters is naturally dependent on the nature of the research question. The ex ante perspective is, for example, more relevant in our case where the principal has to decide before the workers have made their effort decisions, than in research where the question asks whether citizens perceive a realized distribution as fair.

To capture other-regarding preferences in a flexible manner that fits our framework, we assume these preferences to be non-linear for piece rates. The importance principals attach to the high- or low-ability worker's income depends on which worker is receiving a higher *piece rate* (w_h, w_l). This is captured by the indicator variables.

$$E(U) = E\Big(y_p + \gamma\pi + \big[\big(\alpha * \mathbb{1}(w_h \leq w_l) + \beta * \mathbb{1}(w_h > w_l)\big)y_l - \big(\alpha * \mathbb{1}(w_h \leq w_l) + \beta * \mathbb{1}(w_h > w_l)\big)y_h\big]\Big) \tag{1}$$

Parameters of interest are α and β, and γ. We measure the extent to which the principal values output on top of profit maximization by γ. This proxies for an intrinsic motivation to maximize profits. The α and β parameters characterize distributive preferences flexibly, by considering two cases.

- α quantifies the extent to which the principal cares about the low-ability worker relative to the high-ability worker if the latter receives a *lower* piece rate than the former.
- β quantifies the opposite scenario, i.e. how much the principal cares about the low-ability worker relative to the high-ability worker if the latter receives a *higher* piece rate than the former.

Note, that we also allow for $\alpha = \beta$. This then boils down to a more standard model of inequality aversion. We allow for this discontinuity in order to capture a distinct preference for equal procedure or the acceptance of moderate inequality.[22]

We can identify several cases:

1. **Output oriented** $\alpha = \mathbf{0}, \beta = \mathbf{0}$: This principal only cares about the maximization of output. The way income is distributed among workers is irrelevant.
2. **Equal procedure** $\alpha < \mathbf{0}, \beta > \mathbf{0}$: This principal attaches positive importance to the high ability worker's income, when his piece rate is lower than that of the low ability worker, and the principal attaches positive importance to the low-ability worker's income in the opposite case. Therefore, this principal is averse to inequality in piece rate wages and prefers to treat both agents identically.
3. **Redistributive** $\alpha > \mathbf{0}, \beta > \mathbf{0}$: This principal attaches positive importance to the low-ability worker under all circumstances facing principals in our experiment. In our setting, this implies that principals have strong preferences for redistribution from the high- to low-ability worker, and achieve equality ex post.[23]

[22] Ideally, we should also capture altruistic motives where the manager's utility increases when the sum of his agents' payoffs also increases. This would, however, be too difficult to identify along with the other motives used in our data. We have therefore decided to focus on key elements of our design, which are a preference for equal procedure and a preference for equality among agents.

[23] Our experiment does not include cases where the low-ability worker receives *higher* ex post earnings than the high-ability worker.

4. **Inequality-targeting** $\alpha = \mathbf{0}, \beta > \mathbf{0}$: This principal is focused on maximizing output if $w_h \leq w_l$ but she is willing to redistribute as soon as $w_h > w_l$. In our experiment, contracts in which $w_h \leq w_l$ are characterized by relatively low inequality, while it is relatively high for contracts in which $w_h > w_l$. Principals with such preferences can therefore be labeled as averse to high inequality but less averse to low inequality.

5. **Rewarding** $\alpha < \mathbf{0}, \beta < \mathbf{0}$: This principal attaches positive importance to the high-ability worker under all circumstances facing principals in our experiment. In our setting, this implies that principals strongly prefer giving a higher income to the high-ability agent, even when she is already paid a higher piece rate, and even if this comes at the cost of lowering total output.

How do these preferences translate into choices within our experiment? In our experimental design, principals are asked to make multiple decisions between two piece rate contracts. These contracts vary according to whether they come under the $w_h \leq w_l$ or $w_h > w_l$ domain. The choice of contract also affects the agents' (expected) income because they will subsequently work under the chosen contract. Given the evidence presented in Section 4.3, we assume that preferences are defined over distributional outcomes, i.e. the workers' expected income.

More specifically, we can run through the predicted choice patterns of each case listed above. An output-oriented principal (case 1) will always choose the contract that gives her the highest output. A principal who is interested in equal procedure (case 2) will favor a contract that helps the high- (low-) ability agent in the case of both options being characterized by $w_h \leq w_l$ ($w_h > w_l$). In the case of one option being in the $w_h \leq w_l$ domain and the other option being $w_h > w_l$, it depends on the relative strength of α and β, as well as the cost in terms of forgone output. Principals that are characterized by strong redistributive preferences (case 3) prefer the contract that minimizes ex post inequality between both workers. The willingness to forgo output for the sake of redistribution can vary according to who receives the higher piece rate, and is characterized by the magnitude of α and β. Principals who only care about the relative income of agents if $w_h > w_l$ (case 4) will choose the output-maximizing contract for all cases where $w_h \leq w_l$ in both contract options. They only take distributive consequences into consideration if a contract gives the high-ability agent a higher piece rate. In this case, they will reject contracts if the difference in payoffs becomes too great under a contract in which $w_h > w_l$. Finally, "rewarding" principals have a preference for maximizing the income of the high-ability agent relative to that of the low-ability agent. Consequently, they will always choose a contract that gives the high-ability agent a higher piece rate. In our experiment, there are two situations in which such a contract is *not* the output-maximizing contract, under the assumption that workers best-respond.

Identification of γ We use our treatment variation to identify γ. Spectators' own income was kept constant but agents' joint output varied, while both dimensions were varied for stakeholders. The parameter γ informs us how much less (if $\gamma < 1$), or more (if $\gamma > 1$) Spectators care about output relative to Stakeholders, keeping the other-regarding part of the function constant. This informs us about the relative importance of output once we take away the principals' extrinsic motives to maximize output. Intuitively, this parameter captures the intrinsic motivation to maximize output.

Principals may believe that maximizing output is the managers' job, as some kind of social norm. Even Spectators may care about output for this reason, even if they have no extrinsic (monetary) incentives to do so. This may be a consequence of the framing of the study, or an identity effect.

We can also characterize differences across treatment groups by estimating a more reduced-form model in which we do not differentiate between intrinsic and extrinsic motives to maximize agents' joint output.

$$E(U) = E\Big(\pi + \big[(\alpha * \mathbb{1}(w_h \leq w_l) + \beta * \mathbb{1}(w_h > w_l)\big)y_l \\ - (\alpha * \mathbb{1}(w_h \leq w_l) + \beta * \mathbb{1}(w_h > w_l)\big)y_h\big]\Big) \tag{2}$$

In this case, the joint output π has different meanings for Stakeholders and Spectators. For the former, it encompasses both intrinsic and extrinsic incentives, while it can only represent intrinsic incentives for the latter. Hence, in this specification, cross-treatment differences can only be evaluated in α and β.

5.3 Pooled results

Table 4: Results from a pooled specification

	Model based on equation (1)	Model based on equation (2)		
	(1) Interaction	(2) All	(3) Stakeholder	(4) Spectator
γ	0.311**	–	–	–
	(0.103)			
α	0.07	0.095	0.05	0.21
	(0.055)	(0.04)	(0.09)	(.18)
β	0.17***	0.26***	0.1*	0.55***
	(0.035)	(0.04)	(0.05)	(0.13)
σ	0.08***	0.05***	0.06***	0.033***
	(0.013)	(0.005)	(0.012)	(0.008)
N	1750	1750	898	852

The parameters are estimated using a conditional logit model. Standard errors are clustered at the subject level using the sandwich formula. Column (1) reports parameters from equation (1); column (2) reports parameters from equation (2). Columns (3) and (4) use the model based on equation (2) for the Stakeholder and Spectator sample separately. Observations are on the subject-choice level. *** denotes statistical significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level.

Table 4 focuses on average results for the entire population. Column (1) presents results from a conditional logit model that fits equation (1). The results mirror the results from Table 3, where we ran similar regressions but without assuming any underlying utility function. γ is significantly different from 0 ($p < 0.01$) but also significantly smaller than 1 (t-test, $p < 0.01$). The fact that γ is

smaller than 1 implies that monetary incentives for Stakeholders reduce their intrinsic motivation to increase output. In fact, it reduces the importance they attach to output by more than two thirds. However, the fact that γ is above 0 shows that Spectators still care about output for intrinsic motives.

We can also see that, on average, α is not significantly greater than zero but β is ($p < 0.01$). This corresponds to the behavior outlined in case 4 (inequality-targeting principals). The non-significant α suggests that principals are only willing to sacrifice output up to the point at which both agents receive the same piece rate.

Columns (2) to (4) fit a conditional logit model assuming equation (2) to be the underlying utility function. Column (2) fits the model by examining the entire sample, and columns (3) and (4) presents results that are based on the Stakeholder and Spectator sample. Comparing β across columns (2) and (3), we can observe that Stakeholders are significantly less concerned about inequality if the high-ability agent is paid a higher piece rate wage, capturing the crowding-out effect. α is non-significant for both samples, but the point estimate is larger for spectators. The point estimate is estimated relatively imprecisely.

5.4 Characterizing heterogeneity in preferences

The characterization of heterogeneity in preferences within our sample identifies which types of principals are prevalent. We can then make counterfactual analyses to determine how inequality and efficiency vary across types when making changes to the work environment. The idea is that being a redistributive principal leads to significant inefficiencies in our setting, but this may not be so true in a different context. We focus more particularly on the case in which workers stop being neutral and start comparing their own piece rate with a co-workers' piece rate. We can do this by assuming social preferences in the agents' utility function. We then simulate the efficiency of each type of principal type under this new context.

To identify principals' distributive types, we fit a finite mixture model on contract choices assuming equation (2) to be our underlying utility function. We then observe how principals are sorted into different preference classes as a function of being either a Stakeholder or a Spectator. This approach has the advantage of characterizing heterogeneity in a more comprehensive manner. Finite mixture models (FMM) can be used to characterize heterogeneity in social preferences by grouping subjects into different types. This approach has become increasingly popular in the social preference literature (e.g Cappelen et al., 2007; Bruhin et al., 2018; Sutter et al., 2018) since it is a powerful tool for summarizing the distribution of preferences and relaxing homogeneity assumptions. FMMs are less demanding in terms of data than individual-level estimations of preference parameters, and their predictive properties have been shown to be similar to those of individual estimates (Bruhin et al., 2018). Unfortunately, we cannot specify a finite mixture model assuming equation (1) to be the underlying function because we would need within-principal variations in incentives.[24] The

[24]Without this variation we have to make very strong sorting assumptions. Some Stakeholder principals always choose the contract that maximizes output and therefore their own income. Without within-subject treatment variations in individual incentives, we do not know how behavior within this class changes, i.e. we do not know whether selfish agents are also more likely to care about output for intrinsic reasons.

framework then allows us to measure how the propensity to care about the well-being of the two agents relative to output changes across treatment groups.

To estimate the parameters of the utility function posited above, we use the random utility model framework for discrete choices introduced by McFadden (1973) but assume that the population is composed of a discrete number of types. In Appendix C, we detail the derivation of the type-specific conditional density $f_k(\theta, \sigma | X_1, X_2, Choice)$ following McFadden (1973). θ is the vector of parameters in the utility function (1), σ is a choice-sensitivity parameter, X_1 (X_2) is a vector of contract characteristics associated with contract 1 (2), and $Choice$ is a dummy indicating the decision made by the agent.

The finite mixture model assumes heterogeneity in θ and σ. It posits that the population can be categorized into K preference types, where each type has a distinct parameter vector (θ_k, σ_k). Note that the true type membership is not observable. Hence, the model assumes that every subject belongs to type k with probability p_k *ex ante*. The individual contribution to the likelihood is a weighted sum over type-specific conditional densities

$$l_i(p_2, ..., p_K, \theta, \sigma | X_1, X_2, Choice) = \sum_{k=1}^{K} p_k f_i(\theta_k, \sigma_k | X_1, X_2, Choice)$$

whereby $p_1 = 1 - \sum_{k=2}^{K} p_k$.

The overall log-likelihood function takes the logarithm over l_i and sums across all N individuals.

$$ll(p_2, ..., p_K, \theta, \sigma | X_1, X_2, Choice) = \sum_{i=1}^{N} \log \Big(\sum_{k=1}^{K} p_k f_i(\theta_k, \sigma_k | X_1, X_2, Choice) \Big) \tag{3}$$

In our estimation of type-specific parameters of the utility function (2), we are interested in documenting how the classes are divided across treatment groups. In other words, we want to characterize the *ex ante* class probability as a function of the treatment group T_i. This shows how treatment groups are sorted differently into types. To do this, we specify the probability of being a member of class $k > 1$ using a logit specification where $\alpha_{i,k}$ determines how much more (or less) likely a subject in the Spectator sample is to be in class k, relative to being in the Stakeholder sample.

$$p_k = \frac{\exp(\alpha_{0,k} + \alpha_{1,k} T_i)}{1 + \sum_{k=2}^{K} \exp(\alpha_{0,k} + \alpha_{1,k} T_i)}$$

The number of types must be determined by the researcher and should accurately describe the heterogeneity of the data, without over-specifying the model. We follow Bruhin et al. (2018) in using the normalized entropy criterion (NEC) to determine the optimal number of types.[25]

[25] We refer to the discussion and summary of the econometric literature by (Bruhin et al., 2018, p.16) on which criterion is best to determine the optimal K in a very similar setting.

The NEC measures ambiguity in the *ex post* assignment of individuals to types. We can use Bayes' rule to estimate the *ex post* probability $\tau_{i,k}$ that subject i is in class k.

$$\tau_{i,k} = \frac{\hat{p}_k f_i(\hat{\theta}_k, \hat{\sigma}_k | X_1, X_2, C)}{\sum_{m=1}^{K} \hat{p}_m f_i(\hat{\theta}_m, \hat{\sigma}_m | X_1, X_2, C)}$$

Ideally, the aim is to obtain an unambiguous mapping of subjects into types. This implies that $\tau_{i,k}$ should be either close to 0 or close to 1. The NEC normalizes entropy, $E(k)$ is close to 0 if all $\tau_{i,k}$'s are close to 0 or 1. If the number of classes leads to an ambiguous mapping of subjects into types, $\tau_{i,k}$'s are closer to 0.5 and $E(K)$ increases.

$$E(K) = -\sum_{k=1}^{K} \sum_{i=1}^{N} \tau_{i,k} \log \tau_{i,k}$$
$$NEC(K) = \frac{E(K)}{ll(K) - ll(1)}$$

To determine the optimal number of types, we compare the NEC for different K values ($K > 1$) and select the model with the lowest NEC. Note that this cannot exclude the possibility of a model with only one class performing better. Since the NEC cannot be calculated for $K = 1$ we will fit a model with only one class and then examine whether there is *clear* evidence that a non-negligible proportion of subjects follow a decision rule that is inconsistent with the model implied by the parameters (e.g. selfish behavior, although the model implies strong inequality aversion).

To estimate the finite mixture model, we use the expectation-maximizing (EM) algorithm. The EM-algorithm is a numerical method used to maximize the likelihood function but does not yield standard errors (see McLachlan and Peel, 2004, chapter 2 for a detailed description of how to use the EM algorithm to fit finite mixture models). It is frequently used in the estimation of finite mixture models because gradient-based algorithms tend to suffer from convergence problems due to the non-linearity of the likelihood function. We follow McLachlan and Peel (2004), p.64, in their procedure for calculating standard errors by bootstrapping them parametrically using 1000 iterations and clustering at the individual level.

5.5 Results from the finite mixture model

The FMM characterizes heterogeneity using the value function specified in equation (2). It should be interpreted as reduced-form because we bundle intrinsic and extrinsic motivation to maximize output. This approach has the advantage that we can characterize the crowding out of inequality concerns by incentivizing principals based on sorting into classes, conditional on their treatment.

As mentioned above, we use the NEC to select the optimal number of types. The number of classes that yield the lowest NEC is 3. The NEC for the model with two classes is 0.03; it is 0.02 for the model with three classes and 0.07 for the model with four classes. The specification with K = 4 performs clearly worse than the other two specifications, and the specification with K =

3 performs better than the specification with K = 2.[26] Figure A4 shows that nearly all subjects can be unambiguously assigned to one of the classes based on their behavior, confirming that class-assignment is relatively straightforward under this specification. The FMM results are shown in Table 5.

Table 5: Results from the finite mixture model with three classes

	Output maximizers (1)	Intermediate (2)	Strong redistributors (3)
Parameters			
α	-0.01	0.04	0.49
	$[-0.1, 0.06]$	$[-0.11, 0.33]$	$[0.37, 0.71]$
β	0.00	0.27	0.63
	$[-0.02, 0.03]$	$[0.2, 0.33]$	$[0.58, 0.77]$
σ	0.47	0.03	0.27
	$[0.36, 0.78]$	$[0.026, 0.04]$	$[0.19, 0.45]$
Shares			
Full sample	0.21	0.64	0.15
if Stakeholder	0.42	0.49	0.09
	$[0.35, 0.43]$	$[0.465, 0.56]$	$[0.065, 0.125]$
if Spectator	0.00	0.79	0.21
	$[0.0, 0.0]$	$[0.77, 0.83]$	$[0.17, 0.23]$

Bootstrapped 95% confidence intervals in squared brackets clustered at the individual level using 1000 iteration (McLachlan and Peel, 2004, p.64). One observation is at the subject-choice level (N = 1750). The NEC is 0.02 for a mixture model with three classes. This table presents results from a finite mixture model outlined in section 5.2. The model uses three discrete classes. The columns separate preferences across the three classes. The first panel displays the parameter across classes and the second panel displays class shares. We only use observations where the difference in output based on elicited beliefs is lower than 100. Table B1 replicates this table using the full sample.

The model yields three classes that can be easily interpreted. The first class attaches no importance to agents' well-being, irrespective of whether one agent is better or worse off. This class makes up 21% of the overall sample but is exclusively composed of Stakeholders. These principals are *not* willing to pay for a reduction in inequality; they only care about maximizing output and – given that this group is completely composed of Stakeholders – their own income.[27]

The second class of subjects (Intermediate type) has a positive and significant β. This means that they are willing to increase the income of the low ability worker when she receives a lower piece rate than the high-ability worker. The point estimate is significantly lower than that of group (3), therefore their willingness to redistribute in these situations is limited. α is indistinguishable from 0 but it is estimated relatively imprecisely. However, what we can conclude from this group is (i) that they do care about the distributive consequences of their decisions and (ii) that they are concerned about situations with a very high degree of inequality – situations in which the low-ability agent

[26] Results for the specifications with K=2 or K=4 are available on request.

[27] We can make this statement because none of the Spectators are sorted into this group.

is strongly disadvantaged relative to the high-ability agent. This group of principals constitutes around 65% of the overall sample. Most of the Spectators (79%) can be classified as Intermediate types and around half of the Stakeholders fall into this category.

Finally, the third class (Strong redistributors) attaches considerable importance to the income of the low-ability worker when her piece rate is higher than that of the high-ability agent, and similarly for cases in which the low-ability agent receives a higher piece rate than the high-ability agent. This group always seeks to increase the low-ability agent's income. In our framework, this boils down to a model in which the principal has strong redistributive concerns and wants to minimize inequality as far as possible.[28]

Comparing the class shares across treatment groups, we observe that the results show clear crowding out. While virtually none of the Spectators are characterized by the output-maximizing class, we find that 42% of Stakeholders are sorted into this group. We thus show that monetary incentives completely crowd out other-regarding behavior for 42% of principals.

5.6 Counterfactual analyses

So far, we have assumed that *workers* do not have social preferences and are thus neutral to piece rate differences relative to their co-worker. This is a mechanical feature of our design since we did not inform workers that they were forming pairs. We wanted to isolate the principals' normative preferences, abstracting from strategic concerns arising when workers compare themselves. Our structural estimation enables us to simulate what would have happened in situations where workers dislike inequality (with varying definitions of inequality). In these situations, we show that egalitarian principals' choices become more optimal from an output-maximization perspective. The intuition is that egalitarian principals tend to treat workers more equally and are thus able to avoid sabotage situations that may arise due to undesired inequality.

The simulations are based on a simple principal-agent model in which principals maximize expected utility and their income is the profit made by the firm.[29] Agents hold a power-cost function (see e.g. DellaVigna and Pope, 2018) $c(e) = \frac{ke^{1+s}}{1+s}$, where we vary the curvature of the effort function, s, across high- and low-ability agents such that $s_h < s_l$. This characterizes the idea that high-ability agents tire less quickly as they increase their effort level.[30]

The extent to which principals are able to fully anticipate their workers' social preferences is unclear. Even though the majority of principals believed that agents best respond to incentives in our experiment, this was not the case for all principals. We will study two natural benchmarks for

[28]Note that redistributive contracts do not allow for situations in which the low-ability agent is better off ex post than the high-ability agent. Consequently, we can readily interpret these decisions as redistributive.

[29]We depart from the design in this case because we need to introduce a budget constraint, and therefore letting the principals bear wage costs introduces a budget constraint. Otherwise, if principals only maximized output without any budget constraint, they would choose $w_h = w_l \longrightarrow \infty$ which minimizes inequality while maximizing output. In our experiment, since choices are binary, it is not necessary to introduce a constraint, and this would have overly complicated the design.

[30]This is another departure from the approach we adopted in the experiment, where agents are heterogeneous in terms of their marginal productivity. We diverge from this approach because the ability term is canceled out in the principal's maximization problem and *always* yields equal piece rate contracts in equilibrium. Note that we can generate similar results by assuming that agents differ linearly in their productivity as in the experiment, but that high-ability agents have higher bargaining power due to their higher ability. However, we prefer the above approach as it does not require the modeling of the labor market.

all three distributive preference types: (1) sophisticated principals who correctly anticipate agents' other-regarding concerns; (2) naive principals who falsely believe that agents are not other-regarding and so do not adapt their contracts' choices as agents' other-regarding concerns grow stronger. The two benchmarks show how profits change across the three distributive preference types as agents become more other-regarding, with principals' expectations remaining constant.

5.6.1 Including social comparisons by agents

Social comparisons among agents matter in the field (Breza et al., 2017; Card et al., 2012) and also, but to a lesser extent, in the lab (e.g. Gagnon et al., 2020; Gross et al., 2015; Charness and Kuhn, 2007). The standout finding from these studies is that agents generally accept inequality that makes them better off or that reflects differences in productivity. Only one study performs horizontal comparisons under differences in piece rates (Gagnon et al., 2020). One of their findings is that agents are averse to being treated differently. Nonetheless, we will assign several utility functions to agents and compare them across the following four different scenarios, covering a broad spectrum of social preferences: (1) caring about differences in piece rates; (2) caring about receiving higher piece rates; (3) caring about differences in potential income; and (4) caring about being better off in terms of potential income.[31]

We follow the general framework laid out by Breza et al. (2017). Workers do not only care about their own wage but also about the reference wage. Reference wage is hereby assumed to be her colleague's wage. We posit that workers' payoffs are denoted as

$$V(y_i, y_R, w_i, w_r, e_i) = y_i(e_i, w_i) + M(w_i, w_R, y_i(e_i, w_i), y_R(e_R, w_R))e_i \tag{4}$$

where $y_i(e_i, w_i)$ is the ex post income of agent i, which depends on the effort level e_i she exerts and the piece rate she receives. $M(.)$ is the social preferences function, which depends on the agent's piece rate w_i, the reference piece rate w_R and the reference ex post income y_R. We thus assume that workers may not only care about other workers' ex post income but *also* about their colleagues' piece rates. We will vary the precise structure of $M(.)$ across the scenarios.

(1) Agents care about differences in piece rates If agents have a distinct preference for equal treatment, we can model the agent as being averse to differences in piece rates:

$$M(w_i, w_R) = -\alpha_a(w_R - w_i | w_i < w_R) - \beta_a(w_i - w_R | w_R \leq w_i) \tag{5}$$

For a given level of α_a and β_a they will reduce their effort level if the dispersion of piece rates become too high.[32] Indeed their optimal effort decision becomes:

[31] We are not claiming that agents necessarily hold these exact preferences, we are merely generating hypothetical situations that give us an idea of what would have happened if agents held these preferences.

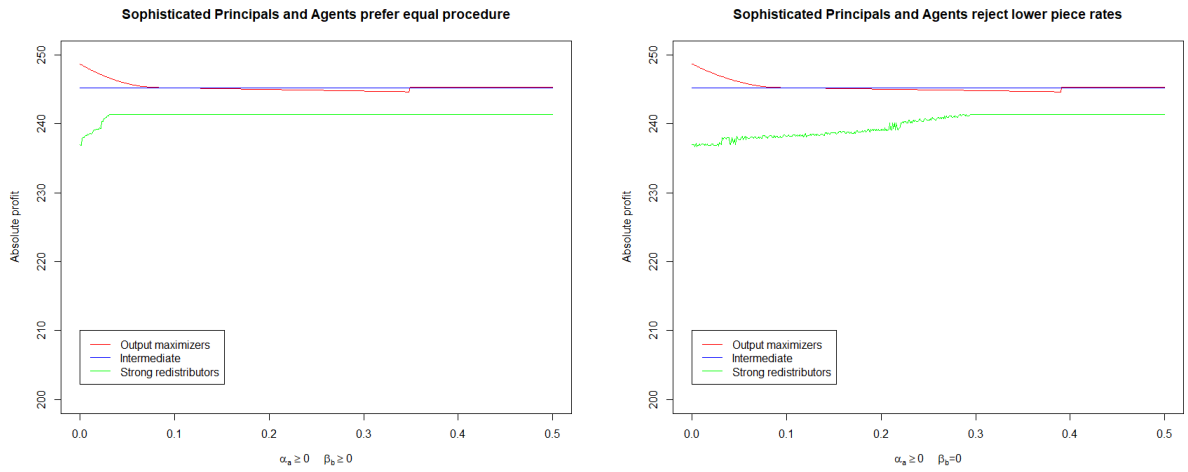
[32] This result is illustrated by the derivative of the agent's utility with respect to effort: $V'_{e_i}(.) = y'_i(e_i) - \big(\alpha_a(w_i - w_R | w_i < w_R) + \beta_a(w_R - w_i | w_R \leq w_i)\big)$. At high levels of piece rate inequality or strong other-regarding motives, a marginal increase in effort will may reduce utility even if $y'(e_i) > 0$. Note that $y(e_i) = w_i * e_i$.

$$e = \left[\frac{w_i - \alpha_a(w_R - w_i | w_i < w_R) - \beta_a(w_i - w_R | w_R \leq w_i)}{k} \right]^{1/s_i} \tag{6}$$

We study two distinct cases: (1a) agents caring about differences in piece rates as such ($\alpha_a = \beta_a \geq 0$) and (1b) agents caring about receiving a slightly *higher* piece rate, i.e. they are only upset when they get a lower piece rate relative to the other agent ($\alpha_a \geq 0$ and $\beta_a = 0$).

Note that principals are now continuously choosing between all possible piece rate contracts according to their expectations of agents' responses. For each level of α_a and β_a we have one pair (w_H, w_L) that is payoff-maximizing for principals. There is no closed-form solution for the principal's problem. We can therefore present the results of the numerical simulations.

Figure 5: Firm's profit if agents care about equal procedure and principals anticipate it correctly



(a) Agents reject inequality in piece rates (b) Agents dislike having a lower piece rate

Notes: The graphs display simulated profit. Principals choose piece rates using the three preference types identified in the previous section. Agents hold preferences characterized by equation (5). The y-axis displays absolute profit. The x-axis displays variation in α_a and β_a. Figure 5a simulates agents with $\alpha_a = \beta_a \geq 0$, while Figure 5b simulates agents with $\alpha_a \geq 0$ and $\beta_a = 0$. Principals correctly anticipate their agents' behavior.

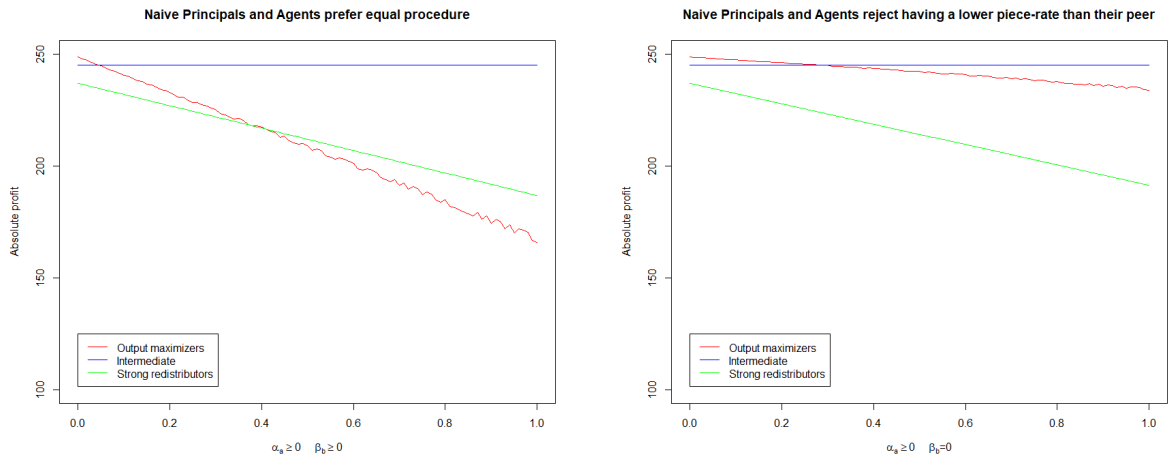
Sophisticated Principals Figure 5 plots the profits associated with each distributive type for different values of α_a and β_b. Here, we make the assumption that principals correctly anticipate the agents' social preferences. Figure 5a considers a case where $\alpha_a = \beta_b$ and agents care about differences in piece rates symmetrically. While profits change for output-maximizers and egalitarian principals, they are constant for intermediate principals. This stems from the fact that they already implement an equal piece rate contract if agents do not hold any social preferences because they are averse to inequality once the high ability agent is paid at a higher piece rate. In addition, we see that as agents become more other-regarding, the output-maximizing principals' profits decrease because they now face retaliation if there is a difference in piece rates. For low levels of α_a and β_a the gap in profits between intermediates and output maximizers shrinks as α_a and β_a increase. For high levels of α_a and β_a, the gap eventually closes, and the two types prefer the same contract, which gives the same piece rate to both agents. Turning to the behavior of egalitarian principals, we can observe

that profits rise as α_a and β_a increase. This comes from the fact that it is now even more costly to implement redistributive contracts because they misallocate incentives, and they are disliked by agents because they do not pay equal piece rates. Egalitarian principals react to this pressure by issuing contracts that become more equal in piece rates and less distorting (yielding higher profits). For high levels of α_a and β_a, egalitarian principals assign the same piece rates to both agents. However, this contract has a lower piece rate *level* than that preferred by the intermediate and output-maximizing types. This piece rate will indeed generate less inequality than the piece rate proposed by intermediate principals, but is Pareto inferior for the agents as they could *both* be better off under the other equal piece rate contract.

Figure 5b examines the case in which agents dislike receiving a lower piece rate than their co-worker but do not mind receiving a higher piece rate. In this case, convergence is slower. This stems from the fact that the *high-ability* agent does not reject this contract whereas the low-ability agent does. This makes equality much more costly for the redistributive principal because he cannot reduce the high-ability agent's piece rate without seeing a drop in his effort.

Naive Principals Figure 6 plots simulated profits for the three types of principals if they are naive about agents' social preferences. A first look at the graphs reveals stark differences relative to Figure 5. Figure 6a considers the case in which agents dislike any difference in piece rates. Strikingly, intermediate principals do not incur any losses by wrongly anticipating that agents are averse to differences in piece rates because they already implement an equal piece rate contract if agents do not hold any social preferences. On the contrary, strong redistributors incur substantial losses because they implement a contract in which the high-ability agent receives a lower piece rate than the low-ability agent. Naive output maximizers also incur large losses as agents become more other-regarding. At some point, they are even less efficient than egalitarian principals because their preferred piece rate spread is too high.

Figure 6: Firm's profit if agents care about equal procedure and principal is naive



(a) Agents reject inequality in piece rates

(b) Agents dislike having a lower piece rate

Notes: The graphs display simulated profit. Principals choose piece rates according to the three preference types identified in the previous section. Agents hold preferences characterized by equation (5). The y-axis displays absolute profit. The x-axis displays the variation in α_a and β_a. Figure 6a simulates agents with $\alpha_a = \beta_a \geq 0$, while Figure 6b simulates agents with $\alpha_a \geq 0$ and $\beta_a = 0$. Principals believe that agents do not hold any social preferences.

Turning to Figure 6b, where agents are only averse to differences in piece rates that make them worse-off, we can observe that egalitarians still perform worse as agents become more other-regarding. Output-maximizers are, however, nearly as efficient as intermediate principals. This is because the high-ability agent does not retaliate to receiving a higher piece rate whereas the low-ability agent does. However, given his low ability, this is not very costly. The opposite is true for the egalitarian principal who pays a higher piece rate to the low-ability agent and the high-ability agent retaliates. This has a significant effect on profits, as characterized by the graph.

(2) Agents have a preference for ex post equality These agents can be modeled as being difference-averse in their expectations. Hence, they care about inequality in the income that individuals are able to attain – their potential income – under a given piece rate. In other words, they care about the inequality of outcomes that would occur if both agents best-responded to incentives $(y_i(w_i, e_i^*), y_R(w_R, e_R^*))$:[33]

$$M(w_i, w_R) = -\alpha_a(y_R(w_R, e_R^*) - y_i(w_i, e_i^*)|y_i^* < y_R^*) - \beta_a(y_i(w_i, e_i^*) - y_R(w_R, e_R^*)|y_R^* \leq y_i^*) \quad (7)$$

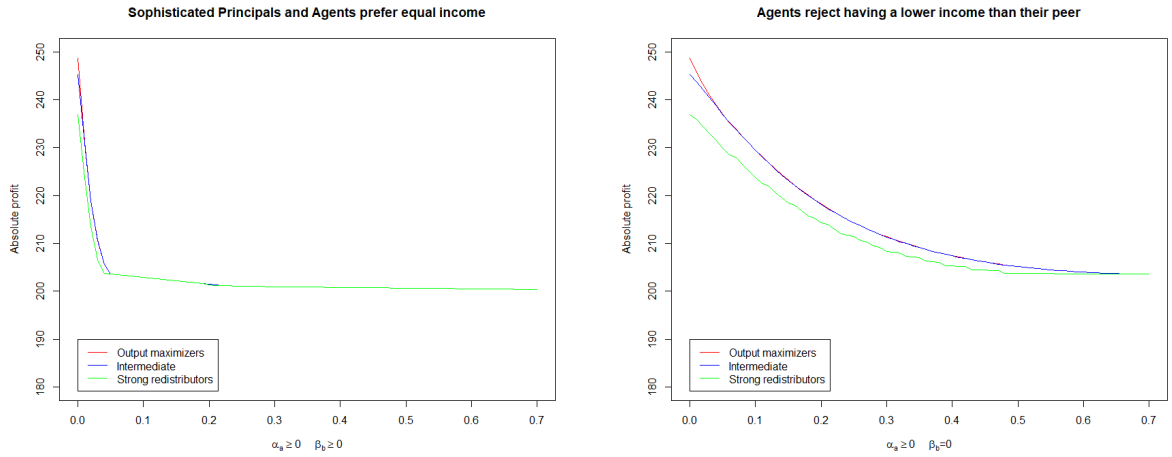
For a given level of α_a and β_a agents will reduce their effort level if the dispersion of *potential* income becomes too high. Indeed their optimal effort decision becomes

$$e_i = \left[\frac{w_i - \alpha_a(y_R(w_R, e_R^*) - y_R(w_i, e_i^*)|y_i^* < y_R^*) - \beta_a(y_i(w_i, e_i^*) - y_R(w_R, e_R^*)|y_R^* \leq y_i^*)}{k}\right]^{1/s_i} \quad (8)$$

We will further examine principals' decisions when they correctly anticipate the agents' behavior, and when they are naive about agents' other-regarding preferences.

[33]This modelling choice takes an ex ante perspective and argues that agents care more about what they would have earned if both had exerted their optimum effort levels. We prefer this approach to one assuming that agents care about equilibrium levels of inequality, i.e. the distribution of income after reacting to the choice of contract *and* its distributive consequences. We make this decision because it reflects the idea that agents care about b eing able to earn the same ex-post income. If agents care about equilibrium levels of inequality, then we would end up with multiple equilibria, including cases in which the low-ability agent increases his effort to compensate for having a low piece rate. This would amount to rewarding the principal for her unequal treatment. We do not consider this to be a realistic situation as it does not capture the fact that agents mostly care about the principal's intentions.

Figure 7: Firm's relative profit if agents care about potential ex-post income and principals are sophisticated



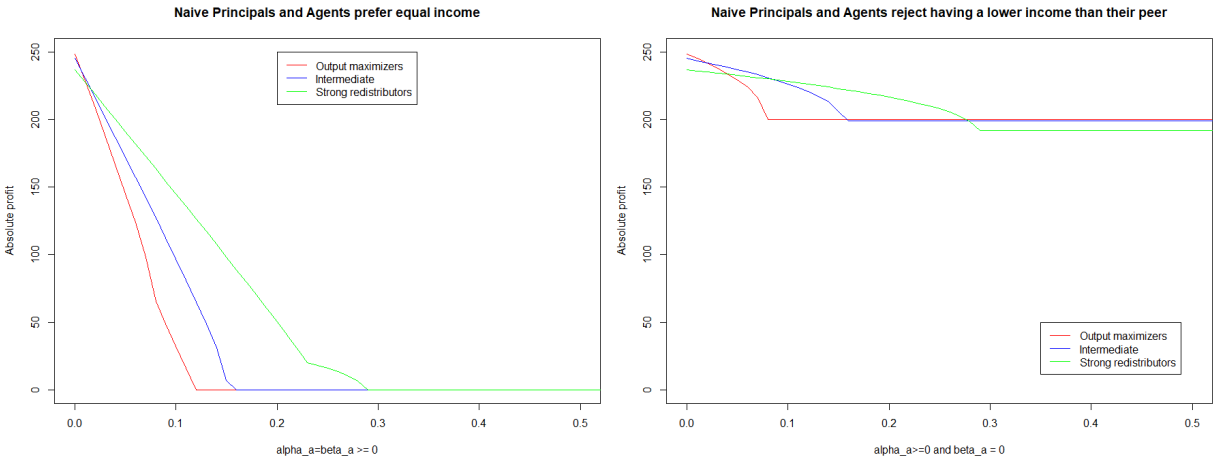
(a) Agents reject inequality in potential income

(b) Agents dislike having lower potential income and principal anticipates correctly

Notes: The graphs display the simulated profit for each of the three preference types. Principals choose piece rates based on the preferences identified in the previous section. Agents hold preferences characterized by equation (7). The y-axis displays the absolute profit.The x-axis displays the variation in α_a and β_a. Figure 7a simulates agents with $\alpha_a = \beta_a \geq 0$, while Figure 7b simulates agents with $\alpha_a \geq 0$ and $\beta_a = 0$. Principals correctly anticipate their agents' behavior.

Sophisticated Principals For each distributive type, Figure 7 plots the absolute profits for different values of α_a and β_b, now assuming that agents hold preferences as in equation (7). Figure 7a considers a case in which $\alpha_a = \beta_b$ and agents care about differences in potential income symmetrically. We can see a similar convergence in behavior to the previous case but the egalitarian principals now become indistinguishable from the output maximizers as other-regarding concerns grow stronger. This stems from the fact that agents punish deviations more severely for higher α_a and β_b. This becomes very costly for all principals and consequently, the optimal behavior now becomes egalitarian with the low-ability worker receiving a higher piece rate in order to harmonize *ex post* the workers' income. This is why the output-maximizing principals are behaving like egalitarian ones. Even though we still see the same convergence as in Figure 7b, it occurs more slowly if agents only care about being *worse* off than their peers. Taking a closer look at the income levels, we can see that output-maximizing principals are quick to reduce inequality by giving a lower piece rate to the high-ability worker and a higher piece rate to the low-ability worker. This is due to the fact that the *low-ability* worker's rejection of inequality becomes a much stronger response than the gain from giving marginally higher incentives to the high-ability worker.

Figure 8: Firm's relative profit if agents care about potential ex post income and principals are naive



(a) Agents reject inequality in potential income

(b) Agents dislike having lower potential income and principal anticipates correctly

Notes: The graphs display simulated profit for each of the three preference types. Principals choose piece rates using the preferences identified in the previous section. Agents hold preferences characterized by equation (7). The y-axis displays absolute profit.The x-axis displays the variation in α_a and β_a. Figure 8a simulates agents with $\alpha_a = \beta_a \geq 0$, while Figure 8b simulates agents with $\alpha_a \geq 0$ and $\beta_a = 0$. Principals believe that agents do not hold any social preferences.

Naive Principals Figure 8 simulates profits for naive principals and inequality-averse agents. Figure 8a assumes that agents care equally about disadvantageous and advantageous inequality. As agents become more inequality-averse, strong redistributors become the most efficient type in relative terms. This is intuitive because the contract they prefer remains that which equalizes ex post incomes, even if they expect agents to be neutral with respect to their co-workers. At some point, other-regarding concerns become so strong that agents do not work at all, even under a contract chosen by a naive egalitarian principal.[34] Hence, agents with high α_a and α_b will eventually retaliate in response to even a small gap in potential income.

Figure 8b assumes that agents dislike being worse-off than their peers. For modest levels of other-regarding concerns, the naive strong redistributors do better than the other two types. However, at some point, low-ability agents no longer exert any effort at all, even under a contract preferred by the egalitarian principal. As in the previous figure, this is due to the fact that there is a small difference in potential income, even in contracts implemented by naive egalitarians. Low-ability agents, who receive a slightly lower potential income, will eventually sabotage this contract if α_a becomes too large. Then, only high-ability agents will work (because they do not care about advantageous inequality) and we return to the situation in which the naive output-maximizing principal is the most efficient.

The simulations have shown that intermediate principals become indistinguishable from output-maximizing principals as we increase workers' distaste for piece rate inequality. Egalitarian principals, however, still prefer suboptimal incentives that, ex post, yield lower inequality. However, if we

[34]The egalitarian principals that we identified in our data are not "perfect" egalitarians and still face a residual trade-off.

assume that agents dislike inequality in ex post income, we find that all three types become indistinguishable in equilibrium. If we assume that principals are naive about workers' social preferences and falsely believe that workers will best respond to incentives, we observe that intermediate principals are more efficient if agents only care about differences in piece rates, and egalitarian principals become more efficient if workers are egalitarian in expected income. These results demonstrate that the manner in which other-regarding preferences held by principals conflict with optimality is crucially dependent on the setting in which the principal operates.

6 Conclusion

Our results suggest that we should rethink how social preferences affect labor market interactions by modeling them under the assumption that other-regarding preferences are important not only to agents, but also to principals. Managers are the decision-makers for wage-allocation schemes and should therefore be a more frequent focus of research, in order to develop a better understanding of the determinants of wage inequality. Our highly encouraging survey evidence shows that even after controlling for a wide array of firm and manager-level characteristics, a significant correlation remains between the implementation of performance pay within firms and managers' fairness beliefs. Although the existence of other-regarding preferences is well-established in the behavioral economics literature, we show that its realm extends even to situations where output-maximization should be key to survival in a competitive economy.

Our experiment, in a controlled setting, establishes that such a relationship is causal, at least in the context of our experiment, and that principals hold normative distributive preferences that are partially crowded-out by incentive concerns. Extensive margins (irrespective of whether the principal has a monetary stake in the production process) are crucial to understanding wage contract choices. Intensive margins (the size of the trade-off between output and equality) also matter, but to a lesser extent.

Future research should generate experimental evidence from the field by eliciting managers' other-regarding preferences and their beliefs in an incentivized manner, and link them to firm outcomes. Furthermore, it would be of great interest to document how managers sort into different sectors or firms based on their other-regarding preferences.

References

Abeler, Johannes, Steffen Altmann, Sebastian Kube, and Matthia Wibral. Gift exchange and workers' fairness concerns: When equality is unfair. *Journal of the European Economic Association*, 8(6):1299–1324, 2010.

Alekseev, Aleksandr, Gary Charness, and Uri Gneezy. Experimental methods: When and why contextual instructions are important. *Journal of Economic Behavior and Organization*, 2017.

Almås, Ingvild, Alexander W Cappelen, and Bertil Tungodden. Cutthroat capitalism versus cuddly socialism: Are Americans more meritocratic and efficiency-seeking than Scandinavians? *Mimeo*, 2018.

Ashraf, Nava and Oriana Bandiera. Social Incentives in Organizations. *Annual Review of Economics*, 2018.

Balafoutas, Loukas, Martin G. Kocher, Louis Putterman, and Matthias Sutter. Equality, equity and incentives: An experiment. *European Economic Review*, 2013.

Bandiera, O, I Barankay, and I Rasul. Incentives for Managers and Inequality among Workers: Evidence from a Firm-Level Experiment. *The Quarterly Journal of Economics*, 122(2):729–773, 2007.

Bandiera, Oriana, Iwan Barankay, and Imran Rasul. Social Preferences and the Response to Incentives: Evidence from Personnel Data. *Quarterly Journal of Economics*, 120(3):917–962, 2005.

Bartling, Björn and Ferdinand A Von Siemens. The intensity of incentives in firms and markets: Moral hazard with envious agents. *Labour Economics*, 17(3):598–607, 2010.

Bastos, Paulo and Natália P. Monteiro. Managers and wage policies. *Journal of Economics and Management Strategy*, 2011.

Bellemare, Charles and Bruce Shearer. Gift giving and worker productivity: Evidence from a firm-level experiment. *Games and Economic Behavior*, 2009.

Bertrand, Marianne and Antoinette Schoar. Managing with style: The effect of managers on firm policies. *Quarterly Journal of Economics*, 2003.

Bloom, Nicholas, Renata Lemos, Raffaella Sadun, Daniela Scur, and John Van Reenen. Jeea-fbbva lecture 2013: The new empirical economics of management. *Journal of the European Economic Association*, 12(4):835–876, 2014.

Bolton, Gary E and Peter Werner. The influence of potential on wages and effort. *Experimental Economics*, 19:535–561, 2016.

Brandts, Jordi, José M Ortiz, and Carles Solà. Distributional concerns in managers' compensation schemes for heterogeneous workers: Experimental evidence. 2019.

Breza, Emily, Supreet Kaur, and Yogita Shamdasani. The morale effects of pay inequality. *The Quarterly Journal of Economics*, 2017.

Brock, J Michelle, Andreas Lange, and Erkut Y Ozbay. Dictating the risk: Experimental evidence on giving in risky environments. *American Economic Review*, 103(1):415–37, 2013.

Bruhin, Adrian, Ernst Fehr, and Daniel Schunk. The many faces of human sociality: Uncovering the distribution and stability of social preferences. *Journal of the European Economic Association*, 2018.

Cabrales, Antonio, Raffaele Miniaci, Marco Piovesan, and Giovanni Ponti. Social preferences and strategic uncertainty: An experiment on markets and contracts. *American Economic Review*, 100 (5):2261–2278, 2010.

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

Card, David, Alexandre Mas, Enrico Moretti, and Emmanuel Saez. Inequality at work: The effect of peer salaries on job satisfaction. *The American Economic Review*, 102(6):2981–3003, 2012.

Charness, Gary and Peter Kuhn. Does Pay Inequality Affect Worker Effort? Experimental Evidence. *Journal of Labor Economics*, 2007.

Clark, Andrew E, David Masclet, and Marie Claire Villeval. Effort and comparison income: Experimental and survey evidence. *ILR Review*, 63(3):407–426, 2010.

Cohn, Alain, Ernst Fehr, Benedikt Herrmann, and Frédéric Schneider. Social comparison and effort provision: Evidence from a field experiment. *Journal of the European Economic Association*, 12 (4):877–898, 2014.

Cronqvist, Henrik and Frank Yu. Shaped by their daughters: Executives, female socialization, and corporate social responsibility. *Journal of Financial Economics*, 2017.

Croson, Rachel and Uri Gneezy. Gender differences in preferences. *Journal of Economic literature*, 47(2):448–474, 2009.

DellaVigna, Stefano and Devin Pope. What motivates effort? evidence and expert forecasts. *The Review of Economic Studies*, 85(2):1029–1069, 2018.

DellaVigna, Stefano, John A List, Ulrike Malmendier, and Gautam Rao. Estimating social preferences and gift exchange at work. *Mimeo*, 2016.

Eisenkopf, Gerald and Sabrina Teyssier. Envy and loss aversion in tournaments. *Journal of Economic Psychology*, 34:240–255, 2013.

Engelmann, Dirk and Martin Strobel. Inequality aversion, efficiency, and maximin preferences in simple distribution experiments. *The American Economic Review*, 94(4):857–869, 2004.

Englmaier, Florian and Achim Wambach. Optimal incentive contracts under inequity aversion. *Games and Economic Behavior*, 69(2):312–328, 2010.

Fehr, E., G. Kirchsteiger, and A. Riedl. Does Fairness Prevent Market Clearing? An Experimental Investigation. *The Quarterly Journal of Economics*, 1993.

Fehr, Ernst and Klaus M Schmidt. Fairness and incentives in a multi-task principal–agent model. *scandinavian Journal of Economics*, 106(3):453–474, 2004.

Fehr, Ernst, Alexander Klein, and Klaus M Schmidt. Fairness and contract design. *Econometrica*, 75(1):121–154, 2007.

Fischbacher, Urs. Z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, 2007.

Fisman, Raymond, Shachar Kariv, and Daniel Markovits. Individual preferences for giving. *American Economic Review*, 2007.

Gagnon, Nickolas, Kristof Bosmans, and Arno Riedl. The effect of unfair chances and gender discrimination on labor supply. *Working Paper*, 2020.

Greiner, Ben. An online recruitment system for economic experiments. 2004.

Gross, Till, Christopher Guo, and Gary Charness. Merit pay and wage compression with productivity differences and uncertainty. *Journal of Economic Behavior and Organization*, 117:233–247, 2015.

Kocher, Martin G., Ganna Pogrebna, and Matthias Sutter. Other-regarding preferences and management styles. *Journal of Economic Behavior and Organization*, 2013.

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

Krawczyk, Michal and Fabrice Le Lec. 'give me a chance!'an experiment in social decision under risk. *Experimental economics*, 13(4):500–511, 2010.

McFadden, D. Conditional logit analysis of qualitative choice behaviour. In Zarembka, P., editor, *Frontiers in Econometrics*, pages 105–142. Academic Press New York, New York, NY, USA, 1973.

McLachlan, Geoffrey and David Peel. *Finite mixture models*. John Wiley & Sons, 2004.

Sutter, Matthias, Francesco Feri, Daniela Glätzle-Rützler, Martin G Kocher, Peter Martinsson, and Katarina Nordblom. Social preferences in childhood and adolescence. A large-scale experiment to estimate primary and secondary motivations. *Journal of Economic Behavior and Organization*, 146:16–30, 2018.

von Siemens, Ferdinand A. Heterogeneous social preferences, screening, and employment contracts. *Oxford Economic Papers*, 63(3):499–522, 2010.

A Tables and Figures

Table A1: Description of the main variables used in the REPONSE survey

Variable name	Original question in the survey	Scale used in the analysis
Dependent variables		
White-collar individualized wage raise	*Did white-collar workers benefited from individualized wage raises*	0 = No; 1 = Yes
White-collar performance-based bonus	*Did white-collar workers benefited from bonuses related to individual performance?*	0 = No; 1 = Yes
Blue-collar individualized wage raise	*Did non-white collar workers benefited from individualized wage raises*	0 = No; 1 = Yes
Blue-collar performance-based bonus	*Did non-white collar workers benefited from bonuses related to individual performance?*	0 = No; 1 = Yes
Main explanatory variables	*With regard to individualized wages (regardless of whether it is implemented in your firm), what do you think about the following assertions? Individualized wages designate a one-off wage-raise or bonus policy that is differentiated across employees and depends on individual assessments*	
Individualized wage raises are unfair.	*They are fairer than undifferentiated increases. 1= Completely agree; 2=Somewhat agree; 3=Somewhat disagree; 4=Completely disagree*	1=Disagree ; 0 = Agree
Individualized wage raises create tension	*They create tensions that could undermine collective functioning. 1= Completely agree; 2=Somewhat agree; 3=Somewhat disagree; 4=Completely disagree*	1=Disagree ; 0 = Agree
Individualized wage raises are subjective	*They cannot be based on objective criteria. 1= Completely agree; 2=Somewhat agree; 3=Somewhat disagree; 4=Completely disagree*	1=Disagree ; 0 = Agree
Individualized wage raises motivate	*They motivate employees. 1= Completely agree; 2=Somewhat agree; 3=Somewhat disagree; 4=Completely disagree*	1=Disagree ; 0 = Agree

The control variables we use can be classified in two types:

- **Individual controls:** they correspond to the individual-level characteristics of the managers who answered the survey. We control for gender, two education dummies (whether the respondent has at least a high-school diploma, and whether the respondent has partially or totally completed undergraduate studies), the position held by the manager within the firm (executive manager, local manager or human resources manager).[35]

[35] As age information is missing for the 2011 wave, we do not control for it. Our results hold true for the 2017 wave with age dummy controls.

- **Firm controls**: five dummies for the size of the plant (below 30 employees, 20-49, 50-99, 100-199, 200-499), four dummies for the age of the plant (under 5 years old, 5-9, 10-19, 20-49), four dummies for the main type of employee working in the firm (blue-collar worker, employee, technicians, sales, white-collar is omitted), the proportion of people on short-term contracts, whether the firm uses interim contracts, whether the firm follows a 35-hour-per-week system, whether it has an independent status (i.e. not belonging to a larger firm), four dummies for the share of unionized people in the firm (0%, 1 to 5%, 5 to 10%, 11% to 20%).[36]

Table A2: Performance-based bonuses and managers' distributive preferences

	White-collar workers			Blue-collar workers		
	(1)	(2)	(3)	(4)	(5)	(6)
Dep var = Did white/blue-collar workers benefited from bonuses based on individual performance?						
Individualized wage raises are unfair	-0.181***	-0.111***	-0.0834***	-0.131***	-0.0810***	-0.0606***
	(0.0135)	(0.0149)	(0.0160)	(0.0150)	(0.0163)	(0.0182)
Individualized wage raises create tension		-0.110***	-0.0613***		-0.0624***	-0.0445***
		(0.0110)	(0.0119)		(0.0117)	(0.0133)
Individualized wage raises motivate		0.0382**	0.0104		0.0421**	0.0298
		(0.0181)	(0.0190)		(0.0194)	(0.0213)
Individualized wage raises are subjective		-0.0940***	-0.0489***		-0.0712***	-0.0775***
		(0.0128)	(0.0136)		(0.0139)	(0.0155)
Wave dummy	Yes	Yes	Yes	Yes	Yes	Yes
Individual controls	No	No	Yes	No	No	Yes
Firm controls	No	No	Yes	No	No	Yes
Observations	7689	7587	5785	8152	8046	6162
Pseudo R^2	0.020	0.042	0.140	0.009	0.016	0.040

Robust standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
The Table shows marginal effects from logit specifications. We regress a binary variable for whether white-collar workers benefited from bonuses based on individual performance in columns (1) to (3) (blue-collar workers in columns (4) to (6)) on binary variables for whether the manager answering the survey thinks that individualized wage raises are unfair, whether they create tensions, motivate, or are subjective. All regressions include a 2017 wave dummy. We additionally control for individual and firms controls in columns (3) and (6). See Appendix Table A1 for a description of all the variables.

[36]We use the data reported by the manager answering the survey. This information is sometimes missing, hence the drop in observations when we add firm controls. Our results hold true when we remove these union dummies.

Table A3: Summary statistics Agents vs principal

Variable	(1) Workers	(2) principals	(3) Diff.	(4) Obs.
Female	0.500	0.434	-0.066	339
	(0.501)	(0.498)	(0.058)	
Age	25.468	25.514	0.046	325
	(5.356)	(4.154)	(0.593)	
In a relationship	0.346	0.343	-0.003	325
	(0.477)	(0.477)	(0.056)	
Student	0.615	0.596	-0.018	327
	(0.488)	(0.493)	(0.057)	
Econ student	0.314	0.310	-0.004	339
	(0.465)	(0.464)	(0.054)	
Master or PhD education level	0.438	0.434	-0.004	339
	(0.497)	(0.498)	(0.057)	
Observations	226	113		339

Table A4: Summary statistics principals Spectator vs Stakeholder

Variable	(1) Spectators	(2) Stakeholders	(3) Diff.	(4) Obs.
Female	0.345	0.517	0.172	113
	(0.480)	(0.504)	(0.093)*	
Age	25.420	25.600	0.180	105
	(3.923)	(4.387)	(0.815)	
In a relationship	0.377	0.309	-0.068	108
	(0.489)	(0.466)	(0.092)	
Student	0.667	0.527	-0.139	109
	(0.476)	(0.504)	(0.094)	
Econ student	0.364	0.259	-0.105	113
	(0.485)	(0.442)	(0.087)	
Master or PhD education level	0.436	0.431	-0.005	113
	(0.501)	(0.500)	(0.094)	
Observations	55	58		113

Table A5: Regressions that characterize contract decisions using belief-based trade-offs and individual fixed effects

	Stakeholders		Spectators	
	(1)	(2)	(3)	(4)
Dependent variable: Contract 2 (high inequality) was chosen				
A is equal piece rate	-0.0245	0.0204	-0.0310	0.000299
	(0.0343)	(0.0363)	(0.0451)	(0.0444)
$\frac{\Delta(\text{Expected Output 2 and 1})}{10}$	0.0700***	0.0674***	0.0299***	0.0283***
	(0.00854)	(0.00836)	(0.00794)	(0.00774)
$\frac{\Delta(\text{Expected Inequality 2 and 1})}{10}$	-0.0334***	-0.0194*	-0.0241**	-0.0150
	(0.00887)	(0.00997)	(0.0112)	(0.0115)
Choice 1 = 1		0.131**		0.107
		(0.0566)		(0.0644)
Choice 13 = 1		0.114**		0.0658
		(0.0545)		(0.0614)
Constant	0.609***	0.528***	0.410***	0.356***
	(0.0413)	(0.0495)	(0.0508)	(0.0538)
Fixed effects	Yes	Yes	Yes	Yes
N	898	898	852	852

Standard errors clustered on the subject level in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The specification regresses a dummy indicating the choice of Contract 2 on other Choice characteristics using a linear probability model. Explanatory variables include the expected difference in output between Contract 2 and 1, the expected difference in inequality between Contracts 2 and 1 (both based on principals' beliefs) and a dummy for whether Contract 1 constitutes an equal piece rate contract rather than an egalitarian contract. All columns include individual fixed effects.

Table A6: Regressions that characterize contract decisions excluding observations where Contract 2 yields a lower output relative to Contract 1

	Theoretical trade-offs		Belief-based trade-offs	
	(1)	(2)	(3)	(4)
Dependent variable: Contract 2 (high inequality) was chosen				
Stakeholder	0.230**	0.230**	0.247***	0.247***
	(0.0909)	(0.0909)	(0.0807)	(0.0813)
$\frac{\Delta(\text{Output 2 and 1})}{10}$	0.0417***	0.0422***	0.0320***	0.0308***
	(0.0127)	(0.0128)	(0.00674)	(0.00661)
$\frac{\Delta(\text{Output 2 and 1})}{10}$ * Stakeholder	0.00721	0.00720	0.0317***	0.0317***
	(0.0170)	(0.0170)	(0.0114)	(0.0113)
$\frac{\Delta(\text{Inequality 2 and 1})}{10}$	-0.0499***	-0.0480***	-0.0234**	-0.0159
	(0.0138)	(0.0147)	(0.0109)	(0.0108)
$\frac{\Delta(\text{Inequality 2 and 1})}{10}$ * Stakeholder	0.00584	0.00572	-0.0111	-0.0110
	(0.0180)	(0.0180)	(0.0150)	(0.0151)
A is equal piece rate	-0.0975**	-0.0888*	-0.0241	0.00373
	(0.0458)	(0.0463)	(0.0435)	(0.0430)
A is equal piece rate * Stakeholder	-0.00454	-0.00507	0.00103	0.000447
	(0.0640)	(0.0641)	(0.0560)	(0.0562)
Choice 1 = 1		0.0506		0.0955**
		(0.0424)		(0.0474)
Choice 13 = 1		-0.00848		0.0704
		(0.0473)		(0.0507)
Constant	0.514***	0.498***	0.429***	0.382***
	(0.0907)	(0.0922)	(0.0835)	(0.0837)
Control variables	Yes	Yes	Yes	Yes
Observations	1532	1532	1620	1620
R^2	0.103	0.104	0.171	0.173

Standard errors clustered on the subject level in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The specification regresses a dummy indicating the choice of Contract 2 on other Choice characteristics using a linear probability model. The sample excludes observations where the difference in output is lower than or equal to zero or higher than a 100. In columns (1) and (2), explanatory variables include a Stakeholder treatment dummy variable, the theoretical difference in output between Contract 2 and 1, the theoretical difference in inequality between Contracts 2 and 1 (both assuming workers' best responses), a dummy for whether Contract 1 constitutes an equal piece rate contract rather than an egalitarian contract and the interactions of these variables with the Stakeholder dummy. In columns (3) and (4), principals' beliefs are used to calculate the difference variables. Columns (2) and (4) add controls for whether the current observation is about Choice 1 (Choice 1 = 1) or Choice 13 (Choice 13 = 1). All the specifications include the following controls: female dummy, economics background dummy, whether the subject is currently a student and whether he is currently in a relationship.

Table A7: Regressions that characterize contract decisions including outliers

	Theoretical trade-offs		Belief-based trade-offs	
	(1)	(2)	(3)	(4)
Dependent variable: Contract 2 (high inequality) was chosen				
Stakeholder	0.190**	0.190**	0.228***	0.208***
	(0.0830)	(0.0831)	(0.0632)	(0.0675)
$\frac{\Delta(\text{Output 2 and 1})}{10}$	0.0452***	0.0419***	0.0194***	0.0180***
	(0.0127)	(0.0129)	(0.00655)	(0.00651)
$\frac{\Delta(\text{Output 2 and 1})}{10}$ * Stakeholder	0.0317*	0.0317*	0.0148	0.0151
	(0.0167)	(0.0167)	(0.0111)	(0.0109)
$\frac{\Delta(\text{Inequality 2 and 1})}{10}$	-0.0468***	-0.0356**	-0.0191**	-0.0112
	(0.0134)	(0.0138)	(0.00921)	(0.00954)
$\frac{\Delta(\text{Inequality 2 and 1})}{10}$ * Stakeholder	-0.00727	-0.00727	-0.00268	0.00182
	(0.0180)	(0.0180)	(0.00998)	(0.0111)
A is equal piece rate	-0.0876**	-0.0493	-0.0226	0.0118
	(0.0437)	(0.0436)	(0.0425)	(0.0429)
A is equal piece rate * Stakeholder	0.0105	0.0105	0.0283	0.0337
	(0.0621)	(0.0621)	(0.0519)	(0.0531)
Choice 1 = 1		0.117***		0.122**
		(0.0432)		(0.0497)
Choice 13 = 1		0.0673		0.110**
		(0.0415)		(0.0549)
Constant	0.483***	0.419***	0.433***	0.379***
	(0.0856)	(0.0855)	(0.0836)	(0.0864)
Control variables	Yes	Yes	Yes	Yes
Observations	1808	1808	1808	1808
R^2	0.100	0.103	0.121	0.125

Standard errors clustered on the subject level in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

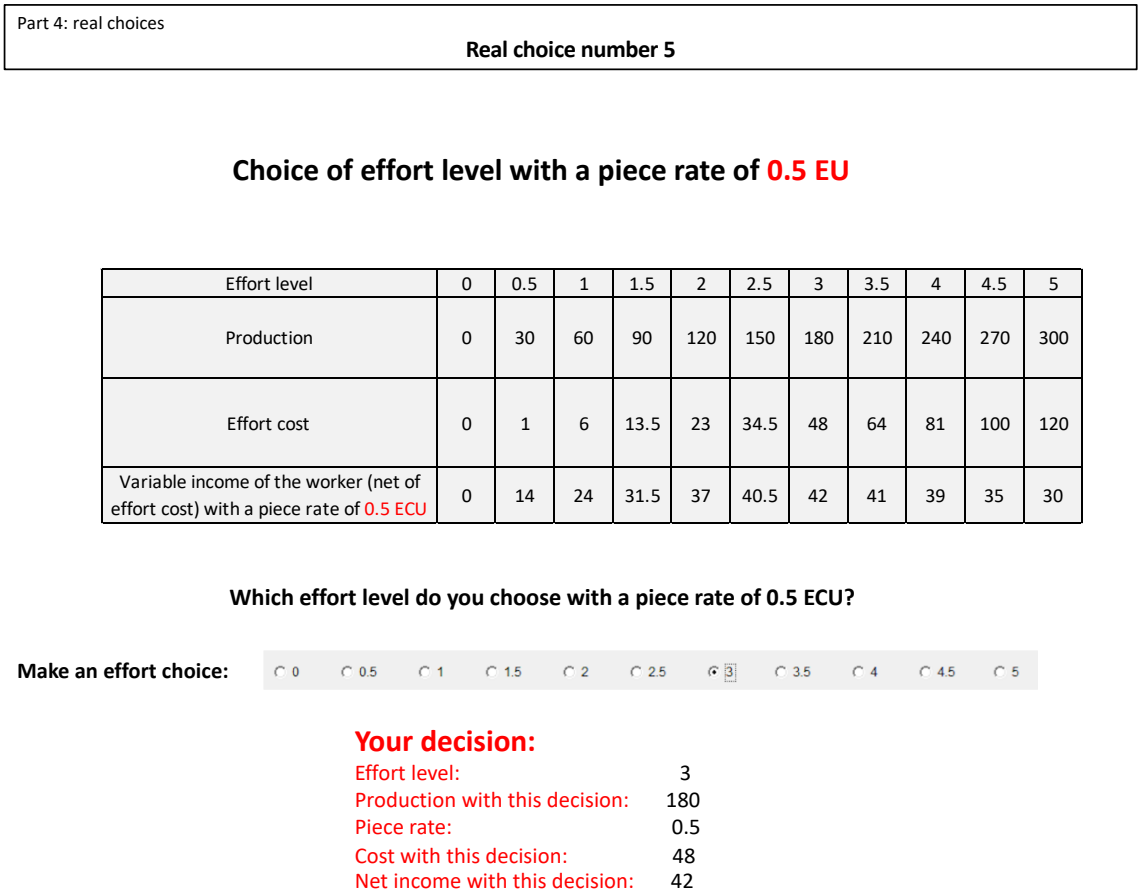
The specification regresses a dummy indicating the choice of Contract 2 on other Choice characteristics using a linear probability model. This sample includes all observations. In columns (1) and (2), explanatory variables include a Stakeholder treatment dummy variable, the theoretical difference in output between Contract 2 and 1, the theoretical difference in inequality between Contracts 2 and 1 (both assuming workers' best responses), a dummy for whether Contract 1 constitutes an equal piece rate contract rather than an egalitarian contract and the interactions of these variables with the Stakeholder dummy. In columns (3) and (4) the difference variables are computed using principals' beliefs. Columns (2) and (4) add controls for whether the current observation is about Choice 1 (Choice 1 = 1) or Choice 13 (Choice 13 = 1). All the specifications include the following controls: female dummy, economics background dummy, whether the subject is currently a student and whether he is currently in a relationship.

Figure A1: The production and cost function per effort level and agent.

Worker A											
Effort level	0	0.5	1	1.5	2	2.5	3	3.5	4	4.5	5
Production	0	30	60	90	120	150	180	210	240	270	300
Effort cost	0	1	6	14	23	35	48	64	81	100	120

Worker B											
Effort level	0	0.5	1	1.5	2	2.5	3	3.5	4	4.5	5
Production	0	20	40	60	80	100	120	140	160	180	200
Effort cost	0	1	6	14	23	35	48	64	81	100	120

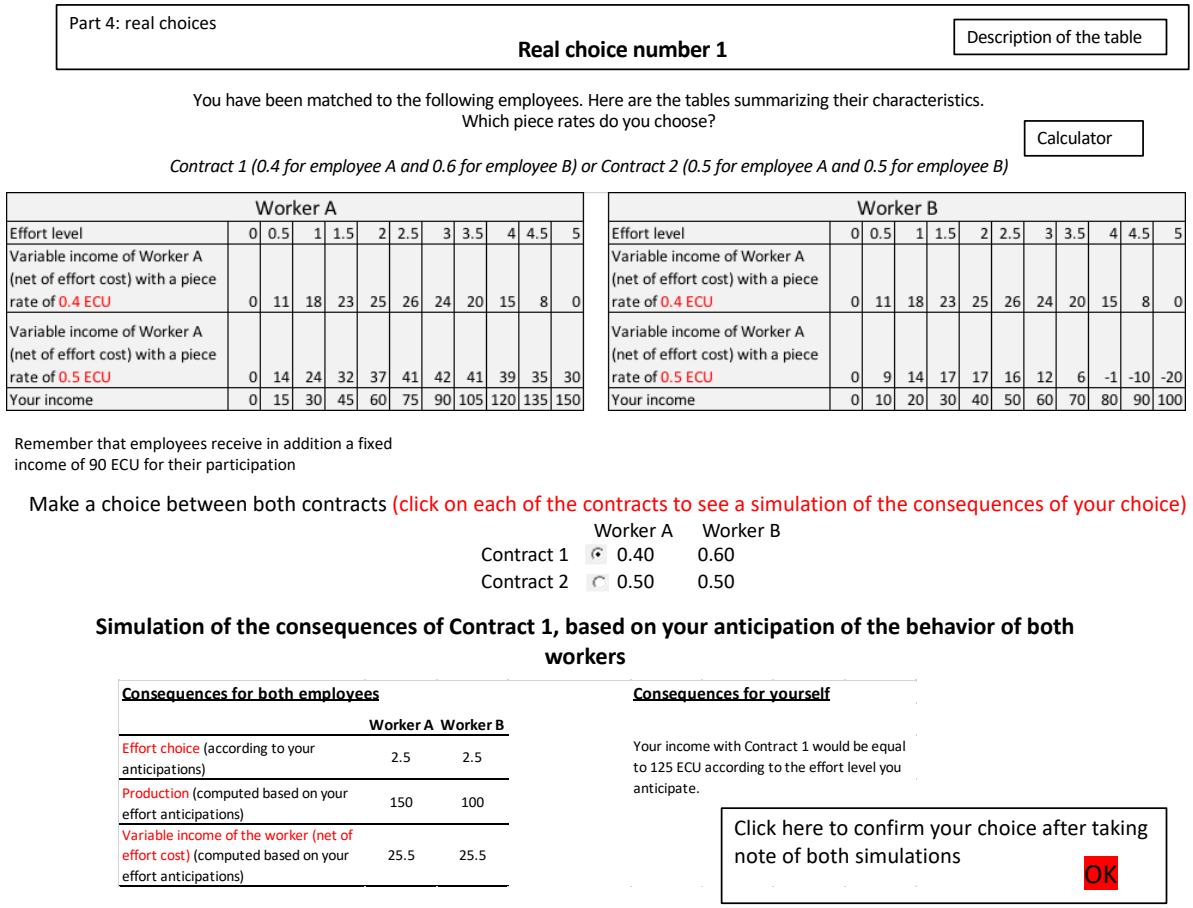
Figure A2: Screenshot of a decision made by agent B.



Effort level	0	0.5	1	1.5	2	2.5	3	3.5	4	4.5	5
Production	0	30	60	90	120	150	180	210	240	270	300
Effort cost	0	1	6	13.5	23	34.5	48	64	81	100	120
Variable income of the worker (net of effort cost) with a piece rate of 0.5 ECU	0	14	24	31.5	37	40.5	42	41	39	35	30

Notes: this is a translated version of the experiment. Original screenshots are available upon request. We recreated the exact same display as the French version.

Figure A3: Screenshot of a decision made by the principal.



Part 4: real choices

Real choice number 1

Description of the table

You have been matched to the following employees. Here are the tables summarizing their characteristics.
Which piece rates do you choose?

Calculator

Contract 1 (0.4 for employee A and 0.6 for employee B) or Contract 2 (0.5 for employee A and 0.5 for employee B)

Worker A											
Effort level	0	0.5	1	1.5	2	2.5	3	3.5	4	4.5	5
Variable income of Worker A (net of effort cost) with a piece rate of 0.4 ECU	0	11	18	23	25	26	24	20	15	8	0
Variable income of Worker A (net of effort cost) with a piece rate of 0.5 ECU	0	14	24	32	37	41	42	41	39	35	30
Your income	0	15	30	45	60	75	90	105	120	135	150

Worker B											
Effort level	0	0.5	1	1.5	2	2.5	3	3.5	4	4.5	5
Variable income of Worker A (net of effort cost) with a piece rate of 0.4 ECU	0	11	18	23	25	26	24	20	15	8	0
Variable income of Worker A (net of effort cost) with a piece rate of 0.5 ECU	0	9	14	17	17	16	12	6	-1	-10	-20
Your income	0	10	20	30	40	50	60	70	80	90	100

Remember that employees receive in addition a fixed income of 90 ECU for their participation

Make a choice between both contracts (click on each of the contracts to see a simulation of the consequences of your choice)

	Worker A	Worker B
Contract 1	0.40	0.60
Contract 2	0.50	0.50

Simulation of the consequences of Contract 1, based on your anticipation of the behavior of both workers

Consequences for both employees

	Worker A	Worker B
Effort choice (according to your anticipations)	2.5	2.5
Production (computed based on your effort anticipations)	150	100
Variable income of the worker (net of effort cost) (computed based on your effort anticipations)	25.5	25.5

Consequences for yourself

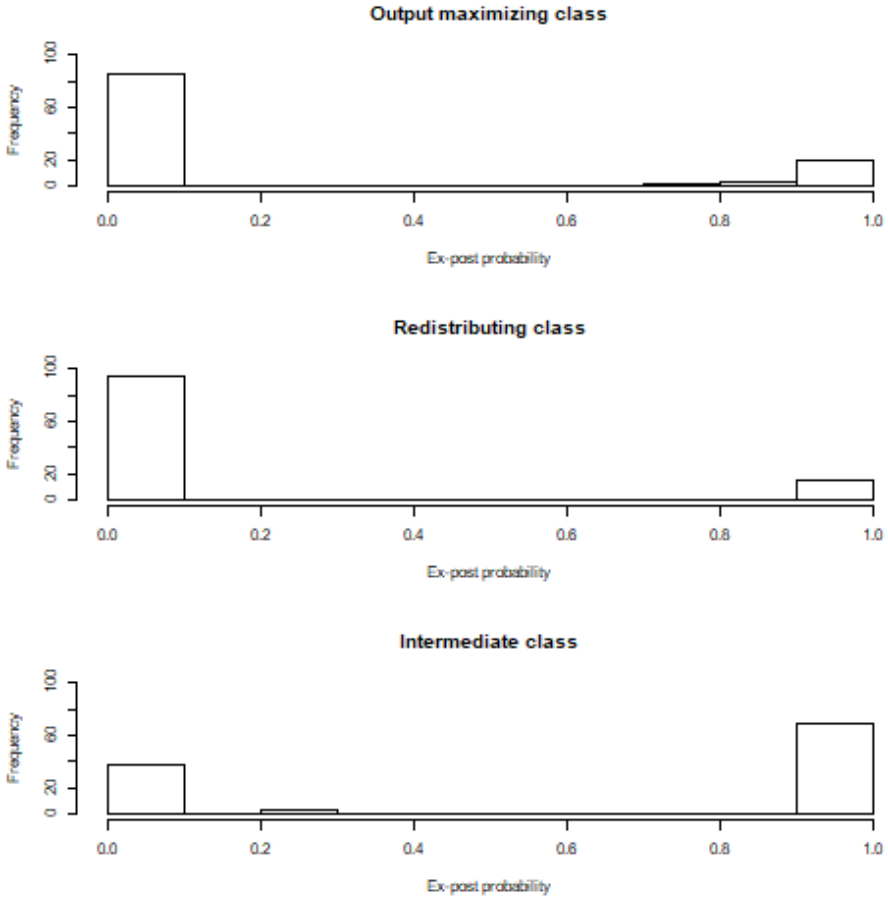
Your income with Contract 1 would be equal to 125 ECU according to the effort level you anticipate.

Click here to confirm your choice after taking note of both simulations

OK

Notes: this is a translated version of the experiment. Original screenshots are available upon request. We recreated the exact same display as the French version.

Figure A4: Distribution of individual ex post probabilities to be part of a given class



Notes: These histograms characterize the distribution of ex post class probabilities on the individual level. The x-axis characterizes the ex post probability and the y axis reports the number of subjects within a bin. The first histograms shows the distribution of the ex post probability to be in the output maximizing class, the second to be in the redistributing class, and the third to be in the intermediate class.

B Robustness Checks

Table B1: Results from the finite mixture model with three classes including outlier beliefs

	Output maximizers	Intermediate	Strong redistributors
	(1)	(2)	(3)
Parameters			
β	0.00	0.36	10.38
	[0.07]	[0.06]	[2.07]
σ	0.28	0.03	0.003
	[0.07]	[0.012]	[0.04]
Shares			
Full sample	0.22	0.63	0.145
if Stakeholder	0.36	0.44	0.19
	[0.02]	[0.04]	[0.03]
if Spectator	0.00	0.47	0.53
	[0.01]	[0.07]	[0.07]

Standard errors from gradient based estimation in parentheses.
Bootstrapped standard errors (1000 iterations) in squared brackets following (McLachlan and Peel, 2004, p.64) .
This table presents results from a finite mixture model outlined in section 5.2. The model uses three discrete classes. Columns separate preferences across the three classes. The first panel displays the parameter across classes and the second panel displays class shares. Includes belief outliers.

Table B1 replicates Table 5 using all 1808 subject-choice observations. Compared to the table posted previously, we have a strong redistributor class that became extremely noisy given a low σ. Furtermore, the NEC became significantly higher (0.09), indicating a worse fit by the data.

What explains these differences given that we have only deleted 58 observations, i.e. 3% of the overall sample? As mentioned above, trade-offs directly enter the objective function in the framework of a conditional logit model and the estimation is moderately sensitive to these outliers because they may imply a relatively high willingness to pay for the reduction of inequalities. Given that we are certainly measuring these beliefs with noise, the subjects may not have always behaved in accordance with the model because we do not observe the "true" incentives these subjects faced. The likelihood of this being the case is higher for large deviations from the best-response benchmark and will generally imply a very high or low willingness to pay for equality that may not always be in accordance with the behavior in the other choices, thus generating a noisy strong redistributor group.

C Deriving the density of the likelihood on the individual level

The model assumes that utility has a deterministic (u) component characterized by equation (2) and a random component (ε) that is stochastic.

$$U^c(X_c, \theta, \sigma) = u(x_c, \theta) + \varepsilon_{X_c} \tag{9}$$

Let $c \in \{A, B\}$ be the contract chosen by the principal, let $X_c = (y_c, \pi_{1,c}, \pi_{2,c}, x_{1,c}, x_{1,c})$ be a vector of the contract's characteristics (own income, worker's production, worker's income); let θ be a vector of parameters e.g. (β, σ) if we fit the model posited in equation (2); ε is an idiosyncratic error in the valuation of u that is assumed to follow a type-I extreme value distribution with a scale parameter $\frac{1}{\sigma}$.

The random component allows us to identify the probability that a principal chooses a given contract within his choice set (Contract 1 or 2). We assume that any principal will choose Contract 1 over 2 if $U^1(X_c, \theta, \sigma) \geq U^2(X_c, \theta, \sigma)$. This can be re-expressed as a probability and yields:

$$\begin{aligned} Pr\big(Choice_t = 1\big) &= Pr\big(u^1(X_1, \theta) - u^2(X_c, \theta) \geq \varepsilon_2 - \varepsilon_1\big) \\ &= \frac{\exp\big(\sigma u^1(X_1, \theta)\big)}{\exp\big(\sigma u^1(X_1, \theta)\big) + \exp\big(\sigma u^2(X_2, \theta)\big)} \end{aligned}$$

If σ is equal to zero, the probability that we choose any contract is equal to 0.5, and the deterministic part of the utility function does not affect her decision and the parameters are uninformative.

The subject's contribution to the conditional density at the choice level will therefore be

$$f_{i,t}(\theta, \sigma | X_1, X_2, Choice) = Pr\big(Choice_{i,t} = 1\big)^{\mathbb{1}(Choice_{i,t}=1)} Pr\big(Choice_{i,t} = 2\big)^{\mathbb{1}(Choice_{i,t}=2)}$$

where t denotes one of the $T = 16$ individual decisions between two contracts. Taking the product over all the decisions the subject makes, we have the subject's overall contribution to the density.

$$f_i(\theta, \sigma | X_1, X_2, Choice) = \prod_{t=1}^{T} f_{i,t}(\theta, \sigma | X_1, X_2, Choice)$$

If we assume that heterogeneity is constant within a type, we can rewrite this density function as a type-specific contribution to the density. Therefore, this represents the contribution of an individual of type k to the density:

$$f_k(\theta_k, \sigma_k | X_1, X_2, Choice) = \prod_{t=1}^{T} f_{i,t}(\theta_k, \sigma_k | X_1, X_2, Choice)$$

D Instructions

Thank you for participating in this experiment. Please read the following instructions carefully. Your answers will remain anonymous throughout the experiment. Please refrain from talking to

your neighbors, and turn off your cellphones. If you choose your answers carefully, you may earn a substantial payoff.

The currency used in this experiment is the ECU. At the end of the experiment, you will be paid in euros using the following conversion rate: 1 euro = 10 ECU.

D.1 Principals

This experiment takes place in a firm. There are two possible roles: being the principal of the firm or being one of the two employees. Your role has been drawn randomly; you are the principal of the firm.

The employees As the principal of the firm, you have to choose the wage paid to both employees. These two people are also participating in this experiment at the same time as you. Although you are in the same room, you will never know who they are, and they will never know who you are. Your identity and their identity will remain anonymous throughout the experiment.

[Stakeholder treatment] You will receive compensation of 60 ECU for your participation. In addition, you will obtain a variable wage that will depend on the production level of both employees. You will obtain the revenues generated by the sales of the units produced by the employees. You will also have the opportunity to earn more money if you correctly guess your employees' behavior.

[Spectator treatment] You will receive a fixed wage of 200 ECU for your participation. You will also have the opportunity to earn more money if you correctly guess the behavior of your employees.

Both employees' wages are paid in two parts. They first receive a fixed participation fee of 90 ECU. The second part is variable and depends on the number of units they produced. Your task is to choose how this variable part is calculated.

Employees' effort level and production Both employees will have to choose their effort levels for the performance of their jobs. Each effort level is associated with a production level. The higher the effort level chosen by the employees, the more they will produce.

[Stakeholder treatment] The more they produce, the more money you will earn. Each unit produced by the employees will earn you 0.5 ECU.

[Spectator treatment] Your own wage is completely independent of their performance. You will receive a fixed wage of 200 ECU.

Example of an effort-production table:

Figure D1: Effort-production table *[Stakeholder treatment]*

Niveau d'effort	0	0.5	1	1.5	2	2.5	3	3.5	4	4.5	5
Production	0	25	50	75	100	125	150	175	200	225	250
Votre revenu	0	13	25	38	50	63	75	88	100	113	125

Figure D2: Effort-production table *[Spectator treatment]*

Niveau d'effort	0	0.5	1	1.5	2	2.5	3	3.5	4	4.5	5
Production	0	25	50	75	100	125	150	175	200	225	250
Votre revenu	200	200	200	200	200	200	200	200	200	200	200

Choice of the wage compensation scheme You will have to define the details of both employees' employment contracts. You will have to decide on the piece-rate wage that each employee will receive. We will show you several examples at the end of the instructions.

Your employees' ability You will obtain information about the ability of both employees. One of them will be more productive than the other. In other words, for the same effort level, one of them will produce a larger quantity than the other.

We will show you a table for each employee describing how their efforts translate into units produced for both employees. You will be able to refer to these tables when you make your wage compensation choices.

The employees' ability will be determined by an aptitude test that they will take at the beginning of the experiment. The higher their grade in the test, the higher their productivity.

This test is a multiple-choice questionnaire consisting of 3 French questions, 3 logic questions and 3 general knowledge questions. They will have 5 minutes to complete the test.

At the end of the instructions, you will also have the opportunity to answer the questions of this test in order to better understand how your employees' productivity has been determined.

Individual choices The employees choose their effort level in complete independence; they will never communicate with each other, nor with you, during the experiment.

They will know the piece rate you chose for them but will be unaware that you have hired another employee. They will not know which piece rate you chose for the other employee. They are not informed that there is another employee.

Effort cost Employees choose an effort level after they have each discovered their piece-rate wage. The higher the effort level they choose, the more it will cost them. Each effort level is associated with a cost in ECU. Therefore, if they choose a high effort level, they will have a higher effort cost to deduct from the wage you will pay them. The cost of the effort is identical for both employees.

Example of an effort-production table

Figure D3: Effort-production-cost table

Niveau d'effort	0	0.5	1	1.5	2	2.5	3	3.5	4	4.5	5
Production	0	25	50	75	100	125	150	175	200	225	250
Coût de l'effort	0.0	1.3	5.0	11.3	20.0	31.3	45.0	61.3	80.0	101.3	125.0

Hence, if this employee chooses an effort level of 1.5, it will cost 11.3 ECU. If she chooses an effort level of 5, it will cost 125 ECU.

Impact of your choices You will choose between several employment contracts for the two employees chosen randomly from among the participants in this experiment today. Your choices have real consequences for both participants. One of your wage choices for both employees will be drawn randomly and will be implemented. You will be the sole decision-maker for both employees.

[Stakeholder treatment] Your own income will correspond to the sales of the unit produced by both employees. Each unit produced will earn you 0.5 ECU. You may additionally earn money for guessing the effort level that your employees will choose in response to various piece-rate wages.

[Spectator treatment] On top of your fixed wage of 200 ECU, you may earn money for guessing the effort levels that your employees will choose when confronted with various piece-rate wages.

D.2 Workers

This experiment takes place in a firm. There are two possible roles: being the principal of the firm or being an employee. Your role has been drawn randomly: you are an employee. You will receive a fixed wage of 90 ECU for participating. You can also obtain an additional wage that will depend on your decisions.

Firm You work in a firm. A principal who has been drawn at random from the people present in this room will offer you a work contract describing your wage for each unit you will produce (piece-rate wage). You must choose an effort level that will be associated with a quantity of units produced. The higher the effort level you choose, the more you will produce. The more you produce, the higher your income will be.

The table below illustrates hypothetically how effort may translate into production for several different effort levels.

Here is an example of an effort-production table:

Figure D4: Effort-production table

Niveau d'effort	0	0.5	1	1.5	2	2.5	3	3.5	4	4.5	5
Production	0.0	25.0	50.0	75.0	100.0	125.0	150.0	175.0	200.0	225.0	250.0

Ability You will have the opportunity to influence how your choice of effort level translates into the quantity produced. You will take an aptitude test that will determine your ability level. This test is a multiple-choice questionnaire consisting of 3 French questions, 3 logic questions and 3 general knowledge questions. Participants will have 5 minutes to complete the test. The higher your performance at this test, the higher your production level will be for a given effort level. You will have an opportunity to familiarize yourself with this type of test by answering 9 other similar questions for 10 minutes.

Effort cost If you choose a high effort level, you will produce more but this will be more costly for you as well. Each effort level is associated with a cost in ECU. Therefore, if you choose a high effort level, you will have a higher effort cost to deduct from your income.

Example of an effort-production-cost table

Figure D5: Effort-production-cost table

Niveau d'effort	0	0.5	1	1.5	2	2.5	3	3.5	4	4.5	5
Production	0.0	25.0	50.0	75.0	100.0	125.0	150.0	175.0	200.0	225.0	250.0
Coût de l'effort	0.0	1.3	5.0	11.3	20.0	31.3	45.0	61.3	80.0	101.3	125.0

Therefore, if you choose an effort level of 1.5, it will cost you 11.3 and you will produce 75 units. If you choose an effort level of 4, it will cost you 80 and you will produce 200 units.

Your income You will be paid a fixed amount for each unit produced. You will be informed of this piece rate before choosing your effort level. In the example below, we show you your variable income (net of effort cost) for a piece rate of 0.4 ECU. You net variable income corresponds to the production multiplied by the piece rate minus the effort cost. In summary, your net variable income = production x piece-rate - effort cost.

Figure D6: effort-production-cost-income table

Niveau d'effort	0	0.5	1	1.5	2	2.5	3	3.5	4	4.5	5
Production	0.0	25.0	50.0	75.0	100.0	125.0	150.0	175.0	200.0	225.0	250.0
Coût d'effort	0.0	1.3	5.0	11.3	20.0	31.3	45.0	61.3	80.0	101.3	125.0
Votre rémunération variable (nette du coût de l'effort) avec un salaire à la pièce de 0.4	0.0	18.7	15.0	18.8	20.0	18.7	15.0	8.7	0.0	-11.3	-25.0

Impact of your choices You will be asked to choose effort levels for several employment contracts. At the same time, the principal of the firm will choose one of these contracts. You will be paid according to the choice made by the principal. The principal will choose a contract without knowing which effort level you chose. You will be unable to communicate with the principal of the firm during the experiment, and will not know his or her identity. Therefore, the principal will be unable to influence your choices. You are completely free to choose your effort level and the principal will be unaware of your choice when making his or her employment contract decision.

E Comprehension test

The principals' comprehension tests were composed of 3 sets of questions of increasing difficulty (tests 1, 2 and 3). Each test consisted of 2 to 6 questions. For each of the three tests, subjects could take three trial tests with hints and feedback on each question to improve their understanding. After the three tests, they had to answer simple True-False questions in order to assess their overall understanding of the rules of the experiment.

Workers also had to take a comprehension test based on the same format, but the questions were adapted to their own choice environment.

Workers and principals were given different tests since their choices were very different. The workers' test ensured that workers were capable of reading the effort-cost-income table (as in Figure A2). We asked them to determine how much income they would obtain under various piece-rate wages and effort choices. The principals' comprehension tests ensured that they were capable of reading the double table describing the characteristics of Workers A and B (as in Figure A3). We asked them to determine the differences between worker A and B (which is the more productive?) and to determine their output, how much income each worker would receive, and their own income in various situations. The Spectators' test was slightly easier since their income is 20 euros in all cases.

E.1 Questions principals

Before moving on to your final choices, we will first ask you a few questions in order to assess your understanding. This test will have no impact on the rest of the experiment. We just want to make sure that you fully understand how the experiment works. You can raise your hand at any time, and someone will come to answer your questions.

E.1.1 Test 1

Let's take the following example. Here is the information about your employees A (first table) and B (second table). The left-hand columns show the production, cost of effort and your income for low effort levels, and the right-hand columns give this information for higher effort levels.

[Stakeholder treatment] Therefore, for employee A, we can see that if he or she chooses an effort level of 2, he or she will produce 100 units. It will cost him or her 20 ECU. For employee B, if he or she chooses an effort level equal to 2, he or she will produce 50 units. It will cost him or her 20 ECU.

[Spectator treatment] Therefore, for employee A, we can see that if he or she chooses an effort level of 2, he or she will produce 100 units. It will cost him or her 20 ECU and you will earn income of 50 ECU. Indeed, each unit produced is sold at 0.5 ECU. For employee B, if he or she chooses an effort level equal to 2, he or she will produce 50 units. It will cost him or her 20 ECU and you will earn income of 25 ECU. Your total income from the sales of the units produced will thus be equal to $50 + 25 = 75$ ECU

Figure E1: Effort-production-cost table *[Stakeholder treatment]*

Employé A

Niveau d'effort	0	0.5	1	1.5	2	2.5	3	3.5	4	4.5	5
Production	0.0	25.0	50.0	75.0	100.0	125.0	150.0	175.0	200.0	225.0	250.0
Coût de l'effort	0.0	1.3	5.0	11.3	20.0	31.3	45.0	61.3	80.0	101.3	125.0
Votre revenu	0.0	12.5	25.0	37.5	50.0	62.5	75.0	87.5	100.0	112.5	125.0

Employé B

Niveau d'effort	0	0.5	1	1.5	2	2.5	3	3.5	4	4.5	5
Production	0.0	12.5	25.0	37.5	50.0	62.5	75.0	87.5	100.0	112.5	125.0
Coût de l'effort	0.0	1.3	5.0	20.0	20.0	31.3	45.0	61.3	80.0	101.3	125.0
Votre revenu	0.0	6.3	12.5	18.8	25.0	31.3	37.5	43.8	50.0	56.3	62.5

Figure E2: Effort-production-cost table *[Spectator treatment]*

Employé A

Niveau d'effort	0	0.5	1	1.5	2	2.5	3	3.5	4	4.5	5
Production	0.0	25.0	50.0	75.0	100.0	125.0	150.0	175.0	200.0	225.0	250.0
Coût de l'effort	0.0	1.3	5.0	11.3	20.0	31.3	45.0	61.3	80.0	101.3	125.0

Employé B

Niveau d'effort	0	0.5	1	1.5	2	2.5	3	3.5	4	4.5	5
Production	0.0	12.5	25.0	37.5	50.0	62.5	75.0	87.5	100.0	112.5	125.0
Coût de l'effort	0.0	1.3	5.0	20.0	20.0	31.3	45.0	61.3	80.0	101.3	125.0

Question 1: Which employee is of higher ability (who is the more productive employee)?

Imagine that employee A chose an effort level of 0.5 and employee B an effort level of 3.

Question 2: What is the total production?

Question 3 ***[Stakeholder treatment only]*****:** How much income do you earn from employee B (what is your income due to the production of employee B)?

Question 4 ***[Stakeholder treatment only]*****:** What is your total income? (add up the income that you earn from both employee A and employee B)

E.1.2 Test 2

You clearly understand how production works in your firm. Now we are going to show you wage simulations to help you make your choices. These examples will have no impact on the rest of the experiment. Let's consider a first choice between two employment contracts. Contract 1 pays employee A 0.6 ECU per unit produced and employee B 0.4 ECU per unit produced. Contract 2 pays employee A 0.4 ECU and employee B 0.6 ECU. We have added two lines to the table, which show the variable wage (net of effort cost) of your employees for both contracts. We have deleted the lines showing your employees' production and effort cost in order to simplify the tables. Remember that the variable wage (net of effort cost) is equal to the production multiplied by the piece-rate wage minus the effort cost.

Figure E3: Effort-production-cost table *[Stakeholder treatment]*

Employé A

Niveau d'effort	0	0.5	1	1.5	2	2.5	3	3.5	4	4.5	5
rémunération variable (nette du coût de l'effort) de l'employé A avec un salaire à la pièce de 0.6	0.0	13.8	25.0	33.8	40.0	43.8	45.0	43.8	40.0	33.8	25.0
rémunération variable (nette du coût de l'effort) de l'employé A avec un salaire à la pièce de 0.4	0.0	18.7	15.0	18.8	20.0	18.7	15.0	8.7	0.0	-11.3	-25.0
Votre revenu	0.0	12.5	25.0	37.5	50.0	62.5	75.0	87.5	100.0	112.5	125.0

Employé B

Niveau d'effort	0	0.5	1	1.5	2	2.5	3	3.5	4	4.5	5
rémunération variable (nette du coût de l'effort) de l'employé B avec un salaire à la pièce de 0.4	0.0	3.8	5.0	3.8	0.0	-6.3	-15.0	-26.3	-40.0	-56.3	-75.0
rémunération variable (nette du coût de l'effort) de l'employé B avec un salaire à la pièce de 0.6	0.0	6.3	10.0	11.3	10.0	6.3	0.0	-8.8	-20.0	-33.8	-50.0
Votre revenu	0.0	6.3	12.5	18.8	25.0	31.3	37.5	43.8	50.0	56.3	62.5

Figure E4: Effort-production-cost table *[Spectator treatment]*

Employé A

Niveau d'effort	0	0.5	1	1.5	2	2.5	3	3.5	4	4.5	5
rémunération variable (nette du coût de l'effort) de l'employé A avec un salaire à la pièce de 0.6	0.0	13.8	25.0	33.8	40.0	43.8	45.0	43.8	40.0	33.8	25.0
rémunération variable (nette du coût de l'effort) de l'employé A avec un salaire à la pièce de 0.4	0.0	18.7	15.0	18.8	20.0	18.7	15.0	8.7	0.0	-11.3	-25.0

Employé B

Niveau d'effort	0	0.5	1	1.5	2	2.5	3	3.5	4	4.5	5
rémunération variable (nette du coût de l'effort) de l'employé B avec un salaire à la pièce de 0.4	0.0	3.8	5.0	3.8	0.0	-6.3	-15.0	-26.3	-40.0	-56.3	-75.0
rémunération variable (nette du coût de l'effort) de l'employé B avec un salaire à la pièce de 0.6	0.0	6.3	10.0	11.3	10.0	6.3	0.0	-8.8	-20.0	-33.8	-50.0

Imagine that you choose Contract 1, hence a rate of 0.4 ECU for employee A and 0.6 ECU for employee B.

Imagine that employee A chose an effort level of 2.5 and employee B an effort level of 1.

Question 1: What is the variable wage (net of the effort cost) of employee A?

Question 2: What is the variable wage (net of the effort cost) of employee B?

Question 3 *[Stakeholder treatment]*: What is your own total income?

E.1.3 Test 3

[Same tables as in Test 2] Imagine that you choose Contract 1, hence a rate of 0.4 ECU for employee A and 0.6 ECU for employee B.

Question 1: For this piece-rate wage of 0.6 ECU, which effort level would employee A choose if he or she wanted to make as much money as possible?

Question 2: For this piece-rate wage of 0.4 ECU, which effort level would employee B choose if he or she wanted to make as much money as possible?

Question 3 *[Stakeholder treatment]*: What would be your total income if both employee A and employee B chose the effort levels that maximize their revenues?

Imagine that you choose Contract 2, hence a rate of 0.6 ECU for employee A and 0.4 ECU for employee B.

Question 4: For this piece-rate wage of 0.4 ECU, which effort level would employee A choose if he or she wanted to make as much money as possible?

Question 5: For this piece-rate wage of 0.6 ECU, which effort level would employee B choose if he or she wanted to make as much money as possible?

Question 6 *[Stakeholder treatment]*: What would be your total income if both employee A and employee B chose the effort levels that maximize their revenues?

E.1.4 True-False

To make sure that you understand the general rules of the experiment, here are several assertions. You have to determine which ones are correct and which ones are wrong.

1. You are matched with 3 employees.
2. Employees choose their effort level according to the piece-rate wages you offer them. You cannot force your employees to choose a particular effort level.
3. Your employees obtain compensation of 90 ECU for their participation.
4. Your employees will not know the piece-rate that you offered the other employee.
5. Both employees are identical.
6. *[Spectator treatment]*: You will earn a fixed wage of 200 ECU. You can earn more money by correctly guessing your employees' reactions.
7. A contract giving the highest piece-rate to the higher ability employee leads to a higher production level but implies larger variable wages differences relative to productivity differences.

8. A contract giving the same piece-rate to both employees causes variable wages to become proportional to the quantity that the employees respectively produce.

9. A contract giving a higher piece-rate to the low-ability employee leads to a lower production level but reduces the differences in the variable wages of both employees.

E.2 Workers' questions

E.2.1 Test 1

Imagine that you can transform effort into production according to the table below. The left-hand columns indicate production and the cost of effort for low effort levels, and the right-hand columns give this information for higher effort levels.

Figure E5: Effort-production-cost table

Niveau d'effort	0	0.5	1	1.5	2	2.5	3	3.5	4	4.5	5
Production	0.0	25.0	50.0	75.0	100.0	125.0	150.0	175.0	200.0	225.0	250.0
Coût de l'effort	0.0	1.3	5.0	11.3	20.0	31.3	45.0	61.3	80.0	101.3	125.0

Question 1: How much would you produce if you chose effort level 2?

Question 2: What is the cost associated with effort level 2?

Question 3: How much would you produce if you chose effort level 4?

Question 4: What is the cost associated with effort level 4?

E.2.2 Test 2

Now imagine that we pay you 0.4 ECU per unit produced. The table below has an additional line compared to the previous one. This line describes your variable wage (net of effort cost) for each production level. Your variable wage (net of effort cost) corresponds to the production multiplied by the piece-rate minus the effort cost.

Figure E6: Effort-production-cost table

Niveau d'effort	0	0.5	1	1.5	2	2.5	3	3.5	4	4.5	5
Production	0.0	25.0	50.0	75.0	100.0	125.0	150.0	175.0	200.0	225.0	250.0
Coût de l'effort	0.0	1.3	5.0	11.3	20.0	31.3	45.0	61.3	80.0	101.3	125.0
Votre rémunération variable (nette du coût de l'effort) avec un salaire à la pièce de 0.4	0.0	18.7	15.0	18.8	20.0	18.7	15.0	8.7	0.0	-11.3	-25.0

Question 1: How much would you produce if you chose effort level 3?

Question 2: What is the cost associated with effort level 3?

Question 3: What effort level allows you to obtain the highest variable wage (net of effort cost)?

Question 4: What effort level allows you to obtain the lowest variable wage (net of effort cost)?

E.2.3 Test 3

Now imagine that we pay you 0.6 ECU per unit produced. The last line of the table below describes your variable wage (net of effort cost) for each production level with this piece-rate wage.

Figure E7: Effort-production-cost table

Niveau d'effort	0	0.5	1	1.5	2	2.5	3	3.5	4	4.5	5
Production	0.0	25.0	50.0	75.0	100.0	125.0	150.0	175.0	200.0	225.0	250.0
Coût de l'effort	0.0	1.3	5.0	11.3	20.0	31.3	45.0	61.3	80.0	101.3	125.0
Votre rémunération variable (nette du coût de l'effort) avec un salaire à la pièce de 0.6	0.0	13.8	25.0	33.8	40.0	43.8	45.0	43.8	40.0	33.8	25.0

Question 1: What effort level allows you to obtain the highest variable wage (net of effort cost)?

Question 2: What effort level allows you to obtain the lowest variable wage (net of effort cost)?

E.2.4 True-False

To make sure that you understand the general rules of the experiment, here are several assertions. You have to determine which ones are correct and which ones are wrong.

1. Each effort level costs the same in ECU.

2. You must choose the preferred effort level of the firm's principal.

3. You must choose effort levels for several employment contracts, but in the end, only one employment contract will be implemented so that you can be paid.

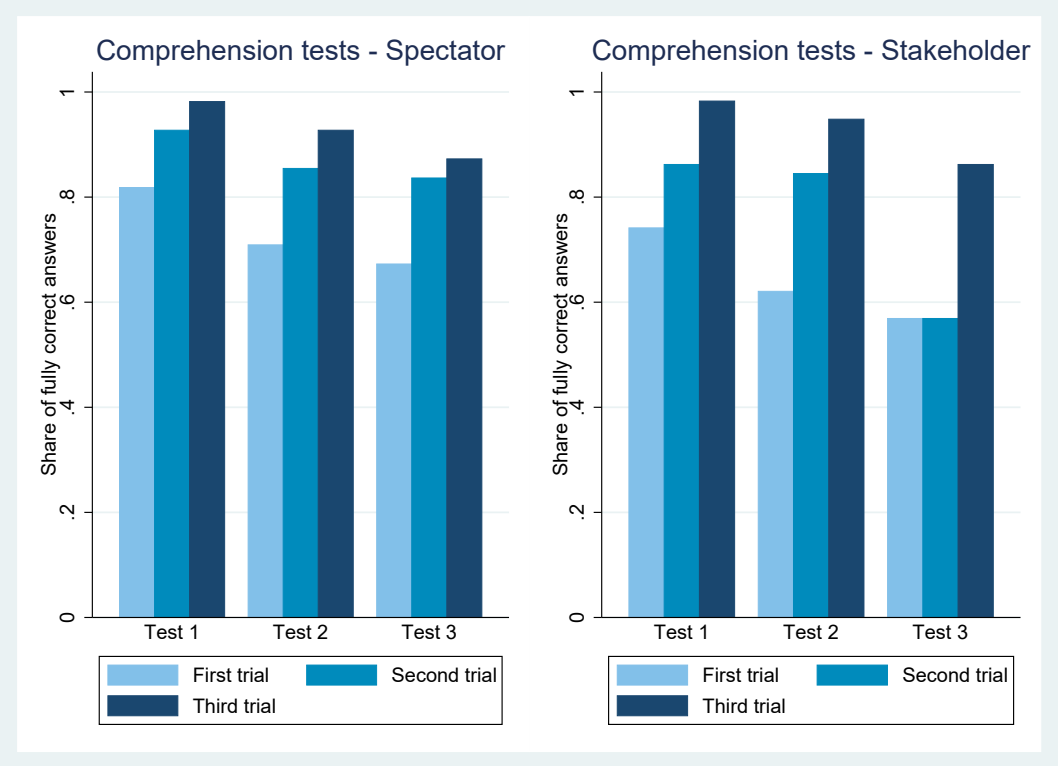
4. You receive a fixed wage of 90 ECU on top of your variable wage.

5. Your fixed wage of 90 ECU will be paid to you once only.

E.3 Comprehension test performance

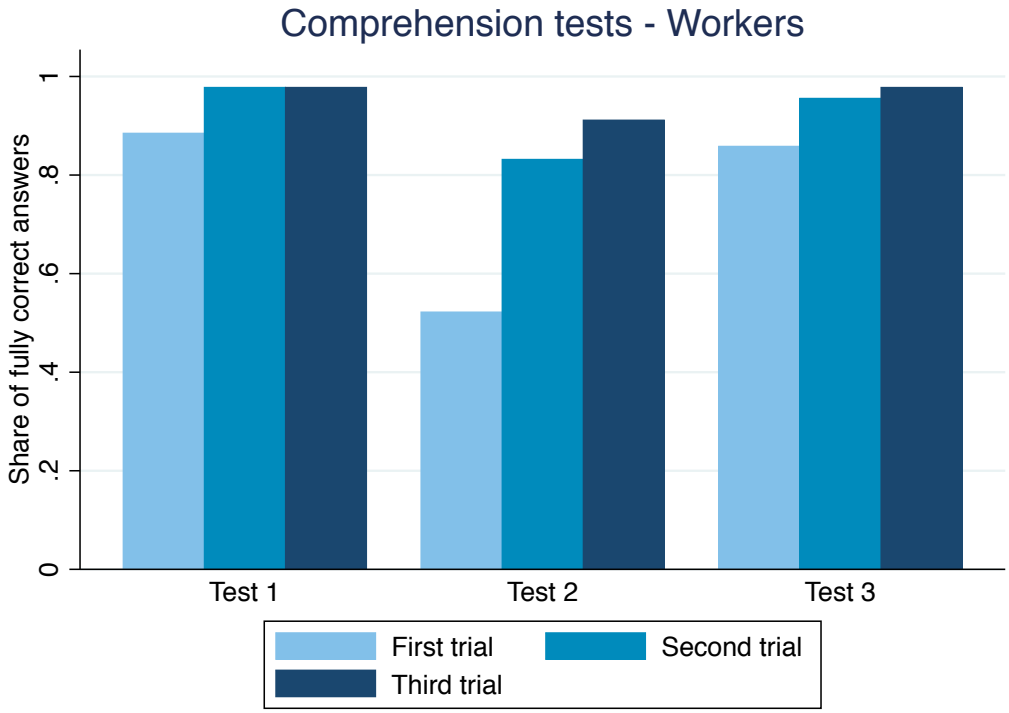
Overall, subjects managed to complete the comprehension tests without any major difficulty and obtained fairly high scores. For each test, the majority of the subjects' answers were completely correct at the first try. Subsequent attempts with feedback improved scores substantially. For the last trials, the share of completely correct answers was always above 83% for all three tests. There were minor variations across Spectators and Stakeholders: principals in the Spectator treatment tended to perform slightly better. This can be easily explained by the fact that the comprehension test for Stakeholders had a few more questions and was harder because we also asked them to compute their own income under various scenarios, which was not necessary for Spectators.

Figure E8: Principals' comprehension tests



Notes:

Figure E9: Workers' comprehension tests



Notes: each bar displays the share of principals achieving a perfect score for each Test and trial. There are three trials per test. The first test has 2 (4) questions for Spectators (Stakeholders), the second test has 2 (3) questions for Spectators (Stakeholders) and the third test has 4 (6) questions for Spectators (Stakeholders)

Table E1: True-False average score

Variable	Obs	Mean	Std. Dev.	Min	Max
Average score Stakeholders	58	.877	.111	.5	1
Average score Spectators	55	.907	.0998	.556	1
Average score Workers	226	.857	.186	.2	1

Notes: The average score is calculated as follows. We create binary variables for each question of the True-False test that are equal to 1 if the subject answered correctly. The average score is the mean of these binary variables.

F Aptitude test

Translated from French to English by the Authors.

F.1 French Questions

Question 1: A hyperbole is a figure of speech in which the expression of an idea or reality is exaggerated in order to highlight it (example: this man is as handsome as an angel). Among the five sentences below, only one does not include hyperbole. Which one?

1. I've been waiting for you for an eternity!
2. Your story is as old as the hills: surely you don't expect anyone to believe you?
3. He came in soaked to the bones because of the storm that was raging outside.

4. **I finished this book in three hours, I devoured it.**

Question 2: Which of the following assertions is the odd one out?[37]

1. All his work is just a drop in the ocean of the work that remains to be done.
2. His explanation was as clear as a mountain stream.[38]
3. **There is a chasm between the world champion and his rivals.**
4. The sea is your mirror, you contemplate your soul in its infinitely rolling waves.

Question 3 Which of the following words is a synonym of eminent?[39]

1. **Remarkable**
2. Immediate
3. Indiscreet
4. Boaster

F.2 Logic questions

Question 4: David has capital of 10,000 euros that he decides to invest in a savings account. After withdrawing his investment with interest two years later, he has total capital of 12,100 euros. What is the annual interest rate on the savings account?

1. 7%
2. **10%**
3. 11%
4. 13%

Question 5: The group formed by the words "triangle", "glove", "clock", "bicycle", corresponds to the group formed by the following numbers:

1. 1,2,3,4
2. 10,4,7,2
3. 4,8,10,12

[37]Subjects had to realize that all sentences except one uses a water-related semantic field. Sentences are translated word for word to make this clearer but obviously, these French expressions using water elements do not always have an exact English counterpart.

[38]Crystal-clear would be the correct translation but then this sentence would be an intruder too

[39]In French immediate can be translated by "imminent" and thus many people are confused about the difference between "éminent" and "imminent"

4. **3,5,12,2**

Question 6: Complete the following series 5V - 4Q - 3L - 2G -?

1. 1A
2. 1B
3. 1C
4. 1D

F.3 General knowledge

Question 7: Simone Veil[40]

1. Was an attorney
2. Had been convicted for anti-Semitic statements
3. **Was the first woman President of the European Parliament**
4. Entered the Panthéon in September 2017

Question 8 The Schengen Agreement is treaty about:

1. The European flag
2. The introduction of the Euro
3. The project of European Constitution
4. **The free movement of people**

Question 9: NASDAQ is a stock market located:

1. **In the United States**
2. In Asia
3. In the United Kingdom
4. In Germany

[40]Simone Veil was a judge but not a lawyer. She entered the French Panthéon in 2018