Experimental Economics (2023) 26:646–672
https://doi.org/10.1007/s10683-023-09791-0

ORIGINAL PAPER

Principal's distributive preferences and the incentivization of agents

Sophie Cêtre[1] · **Max Lobeck**[2]

Received: 29 November 2021 / Revised: 9 January 2023 / Accepted: 9 January 2023 /
Published online: 3 February 2023
© The Author(s) 2023

Abstract

Do principals' distributive preferences affect the allocation of incentives within firms? We run a Principal-Agent lab experiment, framed as a firm setting. In the experiment, subjects are randomized in the principal or worker position. 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-frame and the moral hazard situation, principals do hold egalitarian concerns, as principals are on average willing to trade off their firm's performance (and so their own income) for more wage equality among their workers. The willingness to reduce inequality among workers is sensitive to both extensive and intensive margin incentives, which shows that principals' choices are shaped by incentives that they face themselves.

Keywords Fairness · Distributive preferences · Principal-agent · Social preferences

JEL Classification D63 · C49 · C91

✉ Max Lobeck
max.lobeck@uni-konstanz.de

Sophie Cêtre
sophie.cetre@irsn.fr

[1] Institut de Radioprotection et de Sûreté Nucléaire, Fontenay-aux-Roses, France

[2] Cluster of Excellence "The Politics of Inequality", Thurgau Instute of Economics, University of Konstanz, Konstanz, Germany

https://doi.org/10.1007/s10683-023-09791-0 Published online by Cambridge University Press

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 (Bastos & Monteiro, 2011; Bertrand & Schoar, 2003; 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 make 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).

Establishing causality between managers' distributive preferences and firm performance is complicated. Ideally, we would need a random allocation of managers to firms—to ensure that their normative preferences vary exogenously–-and 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 incentives to maximize output than others because their pay is indexed to 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. Thus, Spectators can implement their preferred income distribution at no cost, which gives us a measure of the distribution of income principals believe is fair and their choices, thus, constitute a normative benchmark. 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.

https://doi.org/10.1007/s10683-023-09791-0 Published online by Cambridge University Press

Moreover, our framework allows us to pin down the relative importance of various fairness ideals (egalitarian, output-maximizing, 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 preference types. Choosing to reward the high ability worker with a higher piece rate is evidence of being output oriented, 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 output-equality trade-off that principals believe they face before workers start working.

We find that despite the firm-like framing and the moral hazard situation, principals do hold egalitarian concerns. On average, they are willing to accept a lower output to reduce 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 maximizing output and equality. Stakeholders are also more sensitive to these intensive margin incentives than Spectators. When the alternative to the high-inequality (output-maximizing) 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 subjects are not more willing to sacrifice output to implement equality in outcomes compared to equality in procedure. Equality in procedure as such is not seen as a particularly attractive contract characteristic and principals are more interested in distributive outcomes. Nonetheless, subjects are significantly more likely to opt for a contract that permits equality in procedure if it is posited *directly* against a contract that provides equality in outcomes compared to similar choices that do not involve such a direct trade-off. This shows that a minority of principals have a weak preference for equality in procedure over equality in outcomes.

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 main contribution to this literature is to study the trade-off that managers face between implementing wage contracts that satisfy their distributive preferences and maximizing the firm's performance.

https://doi.org/10.1007/s10683-023-09791-0 Published online by Cambridge University Press

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 & Strobel, 2007; Fisman et al., 2007). Similarly, allocation games have been used to infer whether subjects are primarily concerned about inequality, or rather inequity (Almås et al., 2020; Cappelen et al., 2007; Konow, 2000). 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 focusing on the relevance of agents' social preferences while modeling principals as profit-maximizers (Koszegi, 2014). Bartling and von Siemens (2010), Englmaier and Wambach (2010), von Siemens (2011) and Itoh (2004) incorporate workers' envy and social comparisons into the derivation of optimal contracts, and found that this affects the optimal incentive structure.

Field and lab experiments have shown that agents compare their income horizontally (e.g. Abeler et al., 2010; Bandiera et al., 2005; Breza et al., 2018; Clark et al., 2010; Cohn et al., 2014; Eisenkopf, et al., 2013; Gross et al., 2015), 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 & Shearer, 2009; Fehr et al., 1993).

Few papers study how principal's other-regarding concerns may affect the allocation of incentives within a firm. Indeed Koszegi (2014) argues in his review of the literature in behavioral contract theory that "[i]n almost all applications, researchers assume that the agent (she) behaves according to one psychologically based model, while the principal (he) is fully rational and has a classical goal (usually profit maximization)." (Koszegi, 2014, p. 1076) 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 & Schmidt, 2004; Fehr et al., 2007). Hoppe and Schmitz (2013) on the other hand study contracting under incomplete information and find no evidence that social preferences affect contract offers.[1] 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.

[1] Their setting differs from ours in many dimensions. Most importantly, we study the design of contracts that involve two, rather than one, agents. This allows us to assess the importance of a broader class of social preferences in contract calibration.

https://doi.org/10.1007/s10683-023-09791-0 Published online by Cambridge University Press

The remainder of the paper is structured as follows: Section 2 introduces the design; Section 3 presents our main results; Section 4 concludes.

2 Experimental design

2.1 Trading off distributive and incentivization concerns

This experiment studies situations in which principals have to trade off distributive concerns and the objective to maximize workers' joint production through the allocation of incentives. The underlying assumption is that principals may not only care about maximizing the firm's outcome but they may also care about the distributive properties of incentive schemes due to their attitudes towards inequality. We consider two distinct types of distributive concerns: first, the allocation of incentives itself may yield inequality in treatment and opportunity. Second, incentives affect the final distribution of income within the firm over which the principal may also have preferences. Depending on the principals' beliefs regarding the effectiveness of a given incentive scheme and the resulting distributive consequences, these motives may conflict with each other. For example, a principal may view tournament incentives as particularly effective for motivating workers but dislikes the fact that they imply large inequality in outcomes. Other principals may have strong egalitarian preferences (thereby preferring flat monthly wage distributions), while believing that an income that is proportional to effort would be the most appropriate incentivization method for maximizing firm performance. The objective of the experimental design is, thus, to study how principals handle such situations using a unified experimental framework.

In particular, we study whether principals forego the implementation of contracts that maximize firm output to avoid large inequality in outcomes or large procedural inequality. To that end, we generate independent variation in the magnitude of the trade-off between output-maximization and equality in opportunity, as well as equality in outcomes. Furthermore, we ask how this depends on the context in which incentivization decisions are made by studying two situations: one where principals face incentives to maximize output themselves (Stakeholder group), and another one where they have no monetary incentives (Spectator group). The latter is a control group that builds a compelling normative benchmark, where principals' distributive preferences can be implemented at no cost. The comparison of the Stakeholder group with this control group is informative on the extent to which other-regarding behavior is crowded out by incentives that they face themselves.[2]

[2] The Stakeholder group refers to a common work situation where managers are held accountable for the performance of their employees, and so their wage raises or bonuses may depend on the fulfillment of production objectives. The Spectator group is related to public sector organizations where managers' wages mostly depend on their own public servant status and level of experience, and is largely insensitive to the performance of the workers they are responsible for.

https://doi.org/10.1007/s10683-023-09791-0 Published online by Cambridge University Press

2.2 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".[3] 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 found online in the replication material.

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, 2015). Sessions were computerized using zTree (Fischbacher, 2007), average payments were 15€ and sessions lasted 90 min, on average. Overall, 339 subjects were invited in groups of 18, 21 or 24 subjects.[4] 226 participants were randomly assigned to the worker role and 113 to the principal role.

2.3 Workers

2.3.1 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, 90 ECU (9 €), and a variable share. The variable share is the product of their effort level (e_i), their marginal productivity $\left(\alpha_i \in \{\alpha_H, \alpha_L\}\right)$, and a piece rate ($w_i$) minus the cost for a given level of effort $\left(c\left(e_i\right)\right)$. As we will explain in more detail below, workers are free to choose an effort level, while they are informed that an (anonymous) principal will choose their piece rate.

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.[5] They are informed that

[3] 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.

[4] 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.

[5] 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. Appendix C in the appendix describes this test further and how the subjects performed.

https://doi.org/10.1007/s10683-023-09791-0 Published online by Cambridge University Press

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 & 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. Furthermore, we deliberately use an aptitude test that not only accounts for innate ability or willingness to work hard, but that may also depend on the education that the worker benefited from. We made this design choice to capture the fact that principals often deal with agents that have different ability levels precisely because they benefited from different education levels. This will then also be relevant for their evaluation of incentives schemes from a fairness perspective.

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 min to complete a practice test (same format but different questions) and then have 5 min to complete a test that will define their production function.[6] 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 and in all sessions we define $\alpha_H = 60$ and $\alpha_L = 40$.[7]

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.

2.3.2 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.[8]

Piece rates range from 0.30 to 0.70 ECU (for high-ability workers) and from 0.30 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

[6] Appendix D 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.

[7] Due to a bug in one session, all agents in that session were erroneously assigned α_H. Given that the experiment was conducted in strategy method, this does not affect decisions or our results.

[8] 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.

https://doi.org/10.1007/s10683-023-09791-0 Published online by Cambridge University Press

piece rates starting at 0.30 and ending at 0.70 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 ($e \in \{0, 0.5, 1, \ldots, 5\}$). We elicit effort decisions for *all* piece rates. The final income of the worker is $y_i = w_i \alpha_i e_i - c(e_i) + 90$ ECU, where $c(\cdot)$ 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.

2.3.3 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. This design feature is vital because we want to focus on the principals' reaction to wage inequality among workers; thus, we want to eliminate other, possibly confusing, factors from the principal's decision, as far as possible.[9]

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 & Bandiera, 2018; DellaVigna et al., 2016). As explained below, in the Stakeholder group agents' actions affect the principals' income. Informing agents how their actions may or may not affect the income of the principal creates an additional uncontrolled source of variation across treatment groups. It would lead to treatment differences in principals' beliefs about the workers' behavior. We shut down this mechanism by restricting the information set of workers.

2.4 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

[9] It has been shown in ultimatum games that the proposer both hold other-regarding preferences (a preference for equality) and strategic concerns (because a more equal split reduces the likelihood of punishment by rejection of the offer) (Azar et al., 2015). We anticipate that such a result could extend to the context of our experiment. It is not possible to disentangle both motives in our setting unless we added a third treatment with this horizontal comparison feature among workers. For simplicity and keep our focus on the main trade-off between personal incentives and other-regarding preferences of managers, we decided to not investigate this additional mechanism.

https://doi.org/10.1007/s10683-023-09791-0 Published online by Cambridge University Press

(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.[10] This neutral labeling implies that we never tell the principal which subject is more productive; they can infer this on their own from the information disclosed in the tables.[11]

2.4.1 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€).[12] The drawing of the payoff-relevant piece rate in the belief elicitation is completely independent of 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 contracts (the extent of output and inequality they produce) is also shared by principals or not.

2.4.2 Contract decisions

After the belief elicitation part of the experiment, 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 1, 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).

Before detailing the choices that principals face, we want to clarify the different types of contracts. Our experiment studies three types of contracts. First, we have

[10] 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.

[11] 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 C for more details regarding the comprehension test.

[12] 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.

https://doi.org/10.1007/s10683-023-09791-0 Published online by Cambridge University Press

egalitarian contracts, where the low ability worker ($Worker_L$) receives a *higher* piece rate than the high ability worker ($Worker_H$). If workers best respond to wages, the egalitarian piece rate will either perfectly equalize income levels (as in Contract 1 of Choices 1–2, 7–9, and 15) or significantly decrease inequality compared to its alternative (as in Contract 1 of Choices 3–7 and 16) because the low ability worker now receives more for each unit produced. This results in a situation where equality in *effort* yields equality in income. Second, we have equal piece rate contracts. These contracts pay an equal piece rate to each worker. They, thus, yield a situation where equality in production leads to equality in incomes. However, in this case, equal effort does not yield equal income, since the two workers differ in their productivity. Third, we have a class of choices named high-inequality contracts, where $Worker_H$ receives a higher piece rate than $Worker_L$. This contract leads to high inequalities *ex-post* but also to higher levels of joint output compared to the alternative.

In the choices that principals make, two of these contracts are posited against each other. As shown in Table 1, we have three classes of choices. First, we have choices, where principals choose between an egalitarian ($Worker_L$ receives a higher piece rate compared to $Worker_H$) and a high-inequality (output-maximizing) contract. Second, we have choices, where principals choose between an equal piece rate contract (both workers receive the same piece rate) and a high-inequality contract. Third, we have choices where the egalitarian contract is posited against an equal piece rate contract. Contract 2 yields higher output compared to Contract 1 in all decisions, except for Choice 3, where the egalitarian choice is also output-maximizing. This Choice permits us to test 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, hence its labeling "high-inequality contract".[13]

Within each class, 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 output (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

[13] Another option could have been to give the principal a choice between a continuous set of piece rates instead of 16 pairwise comparisons of piece rates. We do not follow such a design because it would make the optimization problem that a principal faces much less tractable and also harder to visualize to the subjects. In particular, for a given choice we need to take into account the contract properties (expected output and income distribution) not only for the chosen contract but for all possible alternatives that a principal faces. By having binary choices, we can observe and, thus, control for the alternatives that a principal faces and we can present the characteristics of each alternative without overcrowding the principal's decision screen. Moreover, by restricting the principal's choice set, we can force principals to make certain decisions that are important for getting a complete characterization of their preferences. For example, to disentangle a preference for equality in outcomes relative to equality in piece-rate, we need independent variation in the equality-output trade off in both dimensions. This can only be achieved by restricting the choice-set; a design that does not pose any restrictions on choices would make it impossible to identify such preferences because equal piece-rate contracts would then be by design more output-maximizing compared to egalitarian contracts.

https://doi.org/10.1007/s10683-023-09791-0 Published online by Cambridge University Press

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 income than equal piece rate contracts. The difference in inequality ranges from 13 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 eventually adopt a corner solution in which they maximize income. Consequently, if inequality and output varied on about the same scale, we would 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 in the appendix 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,[14] 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. The simulation helps to ease the cognitive burden and saves computation time. It 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 a random order at the subject-level. Within 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. 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.

2.4.3 Treatments

Between subjects, we will implement two treatments: (1) the ***spectator treatment*** and (2) the ***stakeholder treatment***. The treatment varies across sessions, meaning that principals in the same session faced the same 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

[14] Note that the production and cost of each worker for each effort level are not shown, 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.

https://doi.org/10.1007/s10683-023-09791-0 Published online by Cambridge University Press

Table 1 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 versus high-inequality									
Nr									
1	0.4–0.6	0.55–0.45	25.5	25.5	51.5	13.5	250	270	20
2	0.4–0.6	0.6–0.4	25.5	25.5	63	10.5	250	300	50
3	0.5–0.65	0.55–0.45	42	30.5	51.5	13.5	280	270	−10
4	0.5–0.65	0.6–0.4	42	30.5	63	10.5	280	300	20
5	0.5–0.65	0.65–0.35	42	30.5	75.5	8	280	310	30
6	0.5–0.65	0.7–0.3	42	30.5	90	6	280	340	60
7	0.5–0.75	0.6–0.4	42	42	63	10.5	300	300	0
8	0.5–0.75	0.65–0.35	42	42	75.5	8	300	310	10
9	0.5–0.75	0.7–0.3	42	42	90	6	300	340	40
Equal piece rate versus high-inequality									
10	0.5–0.5	0.55–0.45	42	17	51.5	13.5	260	270	10
11	0.5–0.5	0.6–0.4	42	17	63	10.5	260	300	40
12	0.55–0.55	0.6–0.4	51	20.5	63	10.5	290	300	10
13	0.55–0.55	0.65–0.35	51	20.5	75.5	8	290	310	20
14	0.55–0.55	0.7–0.3	51	20.5	90	6	290	340	50
Egalitarian versus equal piece rate and high-inequality									
15	0.4–0.6	0.5–0.5	25.5	25.5	42	17	250	260	10
16	0.5–0.65	0.55–0.55	42	30.5	51	20.5	280	290	10

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 high-inequality and equal piece rate vs high-inequality choices. Egalitarian contracts result in outcomes proportional to effort. The equal piece rate contracts result in outcomes proportional to output High-inequality contracts maximize output in most of the cases. 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

https://doi.org/10.1007/s10683-023-09791-0 Published online by Cambridge University Press

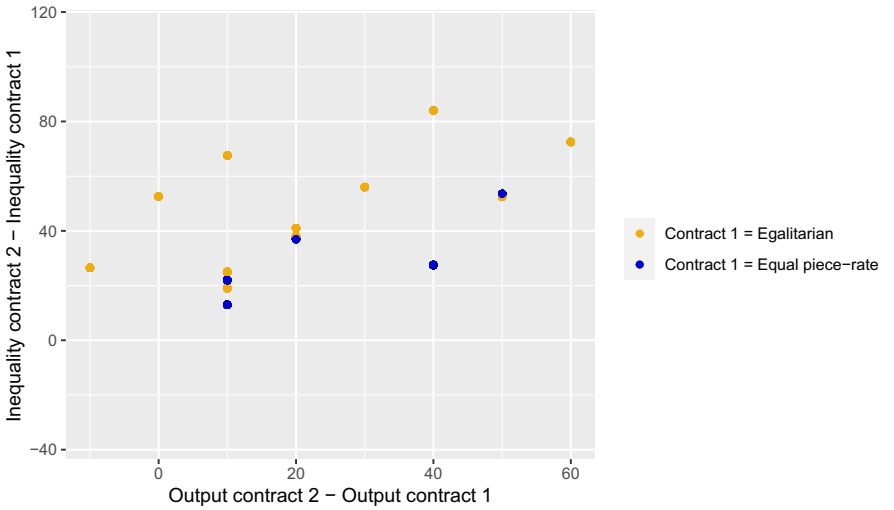


Fig. 1 Contract trade-offs assuming best responses *Notes*: The Figure plots the theoretical trade-offs (assuming best responses), underlying the 16 contract choices that principals have to make. The y-axis shows the difference in inequality between both contracts, and the inequality of a contract is 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 high-inequality contracts, and the blue dots represent the trade-offs of egalitarian contracts vs high-inequality contracts

principals trade-off output and distributive concerns without holding any monetary stake in workers' joint output themselves; it thus constitutes a normative benchmark. In each decision, the principal is asked to make a trade-off between the implementation of an egalitarian (or equal piece rate) and an high-inequality contract (which is output-maximizing in most cases), 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 also be seen as analogous to a situation in which principals have no personal stake in the outcome of their organization (e.g. most civil servants).

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.

https://doi.org/10.1007/s10683-023-09791-0 Published online by Cambridge University Press

2.4.4 Principals' information set

Principals have full information about the environment they face. Principals are instructed that workers are not informed about the salary, nor the existence, of the second worker. They also know that the workers are not explicitly informed about how the principal herself is incentivized. Importantly, they know which of the workers is of low and high ability. Furthermore, they are also informed of the cost that workers have to endure for a given level of effort. This information is necessary for managers to infer the distributive consequences of their actions. Principals are specifically instructed about this in the comprehension test (see Appendix C).

2.5 Hypotheses

To derive the hypotheses that we aim to test through the experimental design, we use a simple theoretical framework that helps characterizing our predictions on choice behavior. To that end, we assume that principals have preferences over their own income, y_p (which depends on firms' level of output in the Stakeholder treatment only), and the distribution of income between workers. Principals may also care about the level of output π for intrinsic reasons. They may believe that maximizing output is the manager's job and they have to behave in this way because they are placed into this position; even Spectators may then care about output. We characterize those preferences using a simple utility function that we assume linear in income and the weight that principals put on the distributive outcomes:

$$\mathrm{E(U)} = \mathrm{E}\left[\left(y_p + \gamma\pi\left(e_H(W_H),\ e_L\left(W_L\right)\right)\right) - \beta M(y^* - Y)\right], \tag{1}$$

where y_p is the income of the principal, γ is the weight the principal puts on output maximization for intrinsic relative to extrinsic motives ($\pi\left(e_H\left(w_H\right), e_L\left(w_L\right)\right)$ is the joint output of the workers), and β is the weight that the principal puts on implementing her preferred distribution of income. To characterize such distributive concerns, we assume that her utility decreases if the implemented distribution of income y diverges from her preferred distribution of incomey^*. This loss in utility is characterized by the loss function$M(\cdot)$, which is weakly convex in the difference of the two. Our experimental design covers two distinct types of distributive concerns: (1) equality in workers' income, where $y^* = \left(y_h\left(e\left(w_h\right)\right), y_l\left(e\left(w_l\right)\right)\right)$ such that $y_h = y_l$ and (2) equality in opportunity, where workers' piece rates are equal: $y^* = \left(y_h\left(e\left(w_h\right)\right), y_l\left(e\left(w_l\right)\right)\right)$ such that $w_h = w_l$.

The binary choices in our experimental design force principals to trade-off the motive to maximize output and distributive concerns. If $\beta = 0$ then principals do not assign any importance towards minimizing $M(\cdot)$ and they are not willing to trade-off output to implement an alternative distribution of income. In this case, they will prefer the output-maximizing contract in all 16 decisions. If β is very large, then the utility gain from minimizing the loss-function $M(\cdot)$ will always be greater than choosing contracts that maximize joint output and principals choose the contract that minimizes $M(\cdot)$ in all 16 decisions. Finally, if $\beta > 0$ but not too large, principals are sensitive to the size of the trade-off between inequality and output maximization

https://doi.org/10.1007/s10683-023-09791-0 Published online by Cambridge University Press

and they will choose an output-maximizing contract once it gets too costly to choose the contract minimizing $M(\cdot)$.

The parameter γ reflects how much Spectators care about output relative to Stakeholders, keeping the other-regarding part of the function constant. Thus, it characterizes crowding out of distributive concerns in contract calibration through financial incentives. If $\gamma = 0$ we have a situation where spectators put no weight on maximizing output and there is no normative appeal for maximizing output. They would, hence, always choose to minimize $M(\cdot)$ and never choose the output-maximizing contract if it conflicts with $M(\cdot)$. If $\gamma > 0$, output maximization has an intrinsic value and spectators are willing to trade-off output maximizing and distributive concerns; if $\gamma = 1$ intrinsic motives to maximize output are as important as extrinsic concerns.[15]

Finally, our design sheds light on what constitutes y^* by studying whether subjects are more or less willing to trade off output to minimize inequality in outcomes or procedural inequality. If subjects care more about procedural equality, they should be more willing to forgo output when facing a choice between a high-inequality and an equal piece rate contract. To be more precise, since the equal piece rate contract is also inequality minimizing when posited against a high inequality contract, a preference for procedural equality is identified as a higher willingness to choose an inequality minimizing contract when it features equal piece rates than piece rates that favor the low ability worker.

This yields the following hypotheses:

Hypothesis 1 *Existence of distributive concerns* If principals hold significant distributive concerns ($\beta > 0$), they will be willing to forgo output in order to minimize inequality in incomes or inequality in procedure among their agents. Hence, output-maximizing contracts are not always chosen.

Hypothesis 2 *Crowding out of distributive concerns* If extrinsic motives to maximize output crowd out distributive concerns in contract choice ($\gamma < 1$), then a larger share of principals will choose the high-inequality contract if they are in the Stakeholder condition compared to the Spectator condition.

Hypothesis 3 *Trade-off between output maximization and distributive concerns* If there exists an intensive-margin trade-off between minimizing inequality and maximizing output, principals will be more likely to choose an output maximizing contract, the more costly the inequality minimizing or equal piece rate contract gets in terms of output forgone. If $\gamma > 0$, such a trade-off exists for both Spectators and Stakeholders.

Hypothesis 4 *Equality in procedure versus equality in outcomes* If principals are more concerned about equality in procedure than equality in outcomes, they are

[15] Note that all parameters characterize the importance of a given motive relative to output maximization, which is normalized to 1.

https://doi.org/10.1007/s10683-023-09791-0 Published online by Cambridge University Press

more willing to forgo output to implement an equal piece rate contract compared to an egalitarian contract.

For a derivation of these hypotheses using a random utility model, see Appendix B.

2.6 Summary statistic

Table A1 in the appendix shows the subjects' sociodemographic characteristics by role. Approximately 50% of the subjects are female, the average age is around 25 years old and 60% are students.[16] There are no systematic differences in observed characteristics between workers and principals. Table A2 in the appendix 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 & 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.

3 Results

3.1 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 70% of high-ability workers choose the 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 67% of the cases (81% when allowing for 0.5 effort deviation), with

[16] 20% are either employed or doing an internship.

https://doi.org/10.1007/s10683-023-09791-0 Published online by Cambridge University Press

no significant differences in beliefs across treatment groups. Principals were also fairly accurate at predicting deviations from the 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.

3.2 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 2 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. In the regressions, 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 Fig. 3, we plot the difference in output between Contract 2 (the high-inequality and 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 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 3.1.

We further classify trade-offs 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

https://doi.org/10.1007/s10683-023-09791-0 Published online by Cambridge University Press

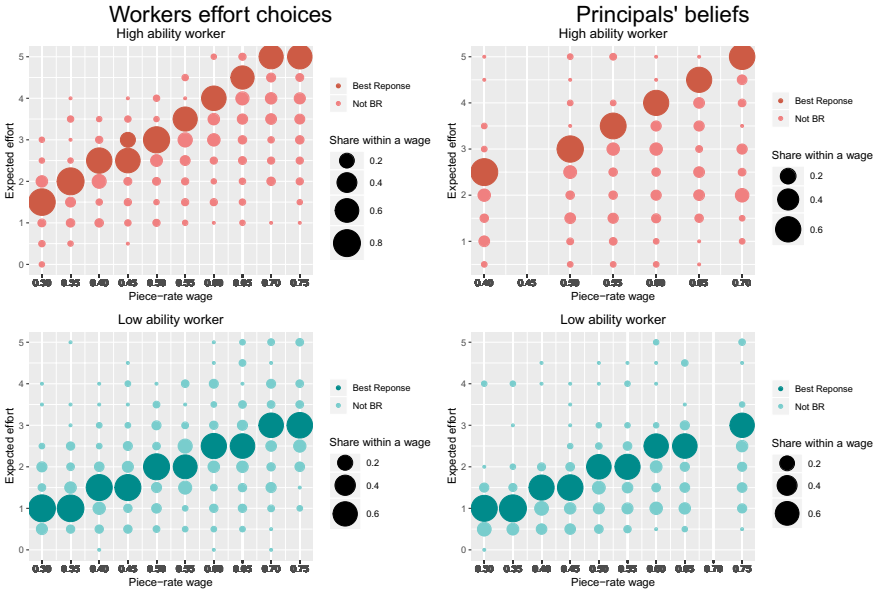


Fig. 2 Workers' stated effort and principals' expected effort by piece rate wage. *Notes*: The figures on the left-hand side plot the workers' choices of effort level for each piece rate (on the x-axis) by ability type. The figures on the right-hand side plot principals' beliefs regarding the effort 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 1. Principals' tasks during the experiment were longer and more demanding than the ones of workers. Hence, we decided to avoid a too large cognitive burden by showing them only the piece rates that would be relevant to their decisions

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. 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. 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 3.3 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

https://doi.org/10.1007/s10683-023-09791-0 Published online by Cambridge University Press



Fig. 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 Fig. 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)

the far-left-hand and far-right-hand brackets, not the intermediate brackets but our regression analysis may be sensitive to such outliers. We come back to the issue of outliers in detail in the relevant sections below.

3.3 Principal's choices

We now describe the pattern of choices across treatment groups. The y-axis of Fig. 4 shows the share of cases in which the high-inequality contract of the pair is selected. We plot this share by the size of the trade-off: Contract 2 increases in output 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.

Overall, we find that, on average, both treatment groups compress wages to a certain extent, given that for all trade-offs, the share of the high-inequality contract is significantly different from 1. Alternatively, this means that the share of inequality-minimizing Contract 1 decisions is always significantly different from 0. This confirms our Hypothesis 1 that, generally speaking, principals are willing to forego output in order to reduce inequality across workers.

Now turning to differences across treatment groups, we find that Stakeholders are more likely than Spectators to choose a high-inequality contract, which confirms Hypothesis 2.

https://doi.org/10.1007/s10683-023-09791-0 Published online by Cambridge University Press

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 likely to choose a high-inequality Contract 2 as Contract 2 increases in output in relation to Contract 1, which confirms Hypothesis 3. On the contrary, Spectators seem less sensitive to output differences. The difference between spectators and stakeholders indicates that, on average, principals are not interested in rewarding high ability agents for doing well in the task.[17]

Furthermore, the figure captures a concave relationship, indicating that when reaching a difference in output of about 40, the share of high-inequality contracts is not further increasing. This can have two reasons: First, by design, contracts that feature a high difference in output are also characterized by high levels of inequality. This concern may lead to a rejection of contracts that have high output differentials, due to concerns for the large inequality they induce. In the regressions discussed below, we find evidence that our principals are indeed attentive to inequality differentials after controlling for output differences across contracts. Second, some principals may prefer to redistribute income at all costs. This is particularly the case for Spectators: about 24% of them always choose an inequality-minimizing contract in at least 14 out of the 16 choices, which is consistent with the level of the light-blue Spectator curve at the left-hand side of the graph. With respect to Stakeholders, the corresponding figure is less than 9%, suggesting that the flattening of the curve after the initial increase is mostly due to principals that try to balance inequality and output decision by decision. About 69% of Stakeholders are characterized by this kind of intermediate trading-off behavior (they choose the high-inequality contract between 3 and 13 times out of the 16 choices). The pure output-maximizers, for whom $\beta = 0$ and that always choose the output-maximizing contract as a corner solution of their optimization problem, constitutes 19% of the Stakeholder group.

Note that the outliers we described in Sect. 3.2 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.

The first two columns of Table 2 characterize these trade-offs assuming that workers best respond (theoretical trade-offs), which can be interpreted as reduced-form estimates and, importantly, they are robust to any endogeneity in beliefs. The drawback of these measures is that they may be less precise given that principals may expect deviations from the 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 2 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

[17] If this were the case, we should see a much higher willingness by spectator principals to give a higher piece rate to the high-ability agent.

https://doi.org/10.1007/s10683-023-09791-0 Published online by Cambridge University Press

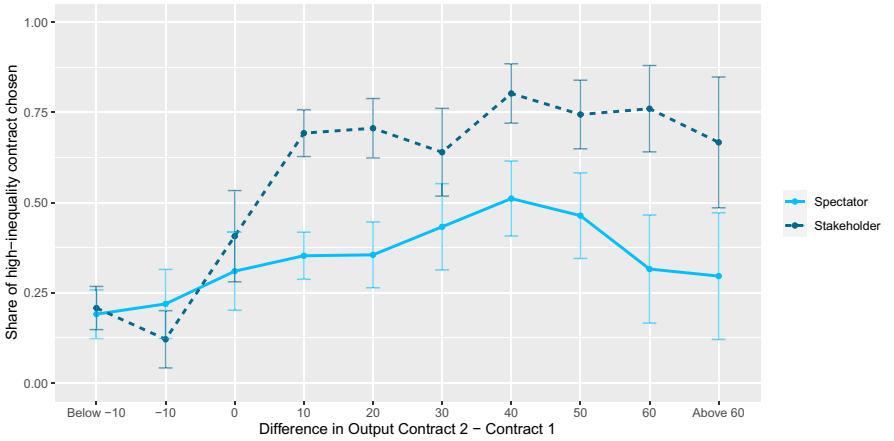


Fig. 4 Principals' contract choices by treatment groups. *Notes*: The Figure shows the share of observations in which the high-inequality contract 2 of the pair is selected. We calculate these shares by output trade-off, i.e. the difference in output between Contract 2 and Contract 1. The solid blue line represents the choices of the Spectator group and the dotted dark blue line shows the choices of the Stakeholder group. The measures are calculated using principals' beliefs on workers' behavior. The same figure using belief-based data is Figure A4 in the appendix. We show 95% confidence intervals for the shares

alternative. The increasing slope in Fig. 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 $\frac{\Delta(Output\ 2\ and\ 1)}{10}$ 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 not explicitly incentivized, even after taking into account the expected cost of equality, which characterizes the shift in the intercept of the two curves in Fig. 4.

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, and this further explains the convexity shown in the plots of Fig. 4. The average effect becomes insignificant once we control explicitly for a decision being an equal piece rate versus egalitarian choice and use belief-based trade-offs, which indicates that this may pick up a peculiarity characterized

https://doi.org/10.1007/s10683-023-09791-0 Published online by Cambridge University Press

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.

Next, we ask whether equal piece rate contracts are considered as more attractive than egalitarian contracts by the principals. Our data allows us to study this from two angles. First, we ask whether subjects are more willing to trade-off output for a reduction in *piece rate* inequality compared to their willingness to trade-off output for a reduction in ex-post *income* inequality. We test for this by including a dummy that indicates that Contract 1 was an equal piece rate contract rather than a egalitarian contract (1 is equal piece rate). The coefficient of this variable indicates whether subjects are more or less likely.

to choose a contract with higher inequalities if the alternative is an equal piece rate contract rather than a egalitarian contract after controlling for differences in output and inequality. We further interact this variable with the stakeholder dummy to test whether this sensitivity differs across treatment groups. Second, we ask whether subjects are more or less likely to embrace an equal piece rate contract if they face a direct choice between an egalitarian and an equal piece rate contract after controlling for differences in inequality reduction and output. This is the case for Choices 15 and 16, where subjects have the choice between an egalitarian and an equal piece rate contract. We capture this through the "egalitarian versus equal piece rate" dummy.

The low-inequality alternative being 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 equally willing to forego output to implement a redistributive and equal piece rate contract; thus, they do not have a strict preference for equal piece rate contracts.

This assessment changes for some subjects if we posit an equal piece rate contract directly against an egalitarian contract. In this case, subjects are indeed significantly more likely to choose the equal piece rate contract, as suggested by the positive and significant effect of facing an egalitarian versus an equal piece rate contract. On average, subjects are 10 percentage points more likely to choose a high-inequality contract if this contract also provides equality in piece rates compared.

Putting both results together, we can conclude that equal piece rates are indeed attractive for some principals from a fairness perspective if promoted as a direct alternative to an egalitarian contract, but their willingness to implement an equal piece rate contract is not different from their willingness to implement an egalitarian contract, as suggested by the insignificant equal piece rate dummy presented above. Note also that the egalitarian contract remains attractive for around half of

https://doi.org/10.1007/s10683-023-09791-0 Published online by Cambridge University Press

Table 2 Regression 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)	(0.0837)	(0.0771)	(0.0780)
$\frac{\Delta(Output\ 2\ and\ 1)}{10}$	0.0467**	0.0429***	0.0316***	0.0302***
	(0.0129)	(0.0131)	(0.00667)	(0.00654)
$\frac{\Delta(Output\ 2\ and\ 1)}{10}$*Stakeholder	0.0300*	0.0299*	0.0339***	0.0339***
	(0.0173)	(0.0173)	(0.0113)	(0.0112)
$\frac{\Delta(Inequality\ 2\ and\ 1)}{10}$	−0.0483***	−0.0358**	−0.0218**	0.0132
	(0.0138)	(0.0142)	(0.00994)	(0.00994)
$\frac{\Delta(Inequality\ 2\ and\ 1)}{10}$ * Stakeholder	-0.00171	-0.00165	-0.0180	-0.0177
	(0.0184)	(0.0184)	(0.0144)	(0.0145)
1 is equal piece rate	−0.0864*	-0.0444	-0.0263	0.00490
	(0.0451)	(0.0446)	(0.0415)	(0.0408)
1 is equal piece rate * Stakeholder	0.00669	0.00660	-0.00421	-0.00447
	(0.0637)	(0.0637)	(0.0543)	(0.0545)
Egalitarian versus Equal piece Rate		0.0977***		0.0907**
		(0.0325)		(0.0388)
Constant	0.476***	0.406***	0.428***	0.375
	(0.0873)	(0.0871)	(0.0793)	(0.0788)
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 choice was an equal egalitarian versus equal piece rate choice. 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

the subjects in either case.[18] Hence, while some principals do indeed have a weak preference for an equal piece rate contract, this is rather a minority as most subjects

[18] The individual fixed effects regressions in Table A3 in the appendix suggest that this effect is mainly driven by stakeholders.

https://doi.org/10.1007/s10683-023-09791-0 Published online by Cambridge University Press

are either strictly output maximizing or are primarily interested in equalizing ex-post incomes rather than piece rates. Overall, Hypothesis 4 is thus only weakly validated.

To sum up, the regression 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 A3 in the appendix 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.[19] Table A4 in the appendix replicates Table 2 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.

4 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. 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.

[19] 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.

https://doi.org/10.1007/s10683-023-09791-0 Published online by Cambridge University Press

External validity is an obvious concern in such kind of experiments. We can worry about the fact that, in real situations, individuals are partly self-selected into managerial positions and their distributive preferences may be one of the factors determining their access to such positions. In our experiment, individuals are randomly selected into the manager position. Our particular problem amounts to the larger issue in the experimental economics literature about whether experienced professionals behave in a similar way as traditional lab samples (mostly students), in firm-like experimental games. Fréchette (2015, 2016) reviews this literature and finds that overall, those two types of populations don't behave too differently in experimental games such as bargaining games, signaling games and other-regarding games. Cooper et al. (1999) find no differences in the long-run for repeated signaling games across real managers and students in China. Fehr and List (2004) find that CEOs are more trusting and more trustworthy than students in a trust game, but they react in a similar way to the features of the experimental design. This suggests that if the magnitude of the treatment effects may differ across students and managers, the direction of the treatment is probably the same.

Future research could start from our experimental design and add more complexity to the decision environment, for example a selection phase for managers, in order to tackle more precisely the issue of selection (or self-selection) based on other-regarding preferences. Another avenue could also be to generate experimental evidence from the field by eliciting managers' other-regarding preferences and their beliefs in an incentivized manner, and link them to real firm outcomes.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10683-023-09791-0.

Acknowledgements The research leading to the results of this study has received funding from the CEPREMAP, EUR Grant ANR-17-EURE-0001 and the Deutsche Forschungsgemeinschaft (DFG—German Research Foundation) under Germany's Excellence Strategy—EXC-2035/1—390681379. We are immensely grateful to the editor and two anonymous referees, as well as 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, Alvin Roth and Andreas Ziegler 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 SMYEE meeting 2021, 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. The replication material for the study is available at https://doi.org/10.17605/OSF.IO/JBX92.

Funding Open Access funding enabled and organized by Projekt DEAL.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

https://doi.org/10.1007/s10683-023-09791-0 Published online by Cambridge University Press

References

Abeler, J., Altmann, S., Kube, S., & Wibral, M. (2010). Gift exchange and workers' fairness concerns: when equality is unfair. *Journal of the European Economic Association, 8*(6), 1299–1324.

Alekseev, A., Charness, G., & Gneezy, U. (2017). Experimental methods: when and why contextual instructions are important. *Journal of Economic Behavior and Organization, 134*, 48–59. https://doi.org/10.1016/j.jebo.2016.12.005

Almås, I., Cappelen, A. W., & Tungodden, B. (2020). Cutthroat capitalism versus cuddly socialism: are Americans more meritocratic and efficiency-seeking than Scandinavians? *Journal of Political Economy, 128*(5), 1753–1788. https://doi.org/10.1086/705551

Ashraf, N., & Bandiera, O. (2018). Social incentives in organizations. *The Annual Review of Economics, 10*, 439–463.

Azar, O., Lahav, Y., & Voslinsky, A. (2015). Beliefs and social behavior in a multi-period ultimatum game. *Frontiers in Behavioral Neuroscience, 9*, 29.

Balafoutas, L., Kocher, M. G., Putterman, L., & Sutter, M. (2013). Equality, equity and incentives: an experiment. *European Economic Review, 60*, 32–51. https://doi.org/10.1016/j.euroecorev.2013.01.005

Bandiera, O., Barankay, I., & Rasul, I. (2005). Social preferences and the response to incentives: evidence from personnel data. *The Quarterly Journal of Economics, 120*(3), 917–962.

Bandiera, O., & Iwan. Barankay, and Imran. Rasul. (2007). Incentives for managers and inequality among workers: evidence from a firm-level experiment. *Quarterly Journal of Economics, 122*(2), 729–773.

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

Bastos, P., & Monteiro, N. P. (2011). Managers and wage policies. *Journal of Economics and Management Strategy, 20*(4), 957–984. https://doi.org/10.1111/j.1530-9134.2011.00310.x

Bellemare, C., & Shearer, B. (2009). Gift giving and worker productivity: evidence from a firm-level experiment. *Games and Economic Behavior, 67*(1), 233–244. https://doi.org/10.1016/j.geb.2008.12.001

Bertrand, M., & Schoar, A. (2003). Managing with style: the effect of managers on firm policies. *Quarterly Journal of Economics, 118*(4), 1169–1208. https://doi.org/10.1162/003355303322552775

Bloom, N., Lemos, R., Sadun, R., Scur, D., & Van Reenen, J. (2014). Jeea-Fbbva lecture 2013: the new empirical economics of management. *Journal of the European Economic Association, 12*(4), 835–876. https://doi.org/10.1111/jeea.12094

Bolton, G. E., & Werner, P. (2016). The influence of potential on wages and effort. *Experimental Economics, 19*, 535–561.

Brandts, J., Ortiz, J. M., & Belda, C. S. (2019). Distributional concerns in managers' compensation schemes for heterogeneous workers: experimental evidence. *Review of Behavioral Economics, 6*(3), 193–218. https://doi.org/10.1561/105.00000107

Breza, E., Kaur, S., & Shamdasani, Y. (2018). The morale effects of pay inequality. *The Quarterly Journal of Economics, 133*(2), 611–663.

Cabrales, A., Miniaci, R., Piovesan, M., & Ponti, G. (2010). Social preferences and strategic uncertainty: an experiment on markets and contracts. *American Economic Review, 100*(5), 2261–2278. https://doi.org/10.1257/aer.100.5.2261

Cappelen, A. W., Hole, A. D., Sørensen, E. Ø., & Tungodden, B. (2007). The pluralism of fairness ideals: an experimental approach. *The American Economic Review, 97*(3), 818–827.

Clark, A. E., Masclet, D., & Villeval, M. C. (2010). Effort and comparison income: experimental and survey evidence. *Industrial and Labor Relations Review, 63*(3), 407–426.

Cohn, A., Fehr, E., & Goette, L. (2014). Fair wages and effort provision: combining evidence from a choice experiment and a field experiment. *Management Science, 61*(8), 1777–1794.

Cooper, D. J., Kagel, J. H., Lo, W., & Qing Liang, Gu. (1999). Gaming against managers in incentive systems: experimental results with Chinese students and Chinese managers. *American Economic Review, 89*(4), 781–804. https://doi.org/10.1257/aer.89.4.781

Cronqvist, H., & Yu, F. (2017). Shaped by their daughters: executives, female socialization, and corporate social responsibility. *Journal of Financial Economics, 126*(3), 543–562. https://doi.org/10.1016/j.jfineco.2017.09.003

Croson, R., & Gneezy, U. (2009). Gender differences in preferences. *Journal of Economic Literature, 47*(2), 448–474.

https://doi.org/10.1007/s10683-023-09791-0 Published online by Cambridge University Press

DellaVigna, S., List, J. A., Malmendier, U., & Rao, G. (2016). Estimating social preferences and gift exchange at work. *Mimeo*.

Eisenkopf, G., Fischbacher, U., & Föllmi-Heusi, F. (2013). Unequal opportunities and distributive justice. *Journal of Economic Behavior and Organization*. https://doi.org/10.1016/j.jebo.2013.07.011

Engelmann, D., & Strobel, M. (2007). Preferences over income distributions experimental evidence. *Public Finance Review, 35*(2), 285–310.

Englmaier, F., & Wambach, A. (2010). Optimal incentive contracts under inequity aversion. *Games and Economic Behavior, 69*(2), 312–328.

Fehr, E., Kirchsteiger, G., & Riedl, A. (1993). Does fairness prevent market clearing? An experimental investigation. *The Quarterly Journal of Economics, 108*(2), 437–459. https://doi.org/10.2307/2118338

Fehr, E., Klein, A., & Schmidt, K. M. (2007). Fairness and contract design. *Econometrica, 75*(1), 121–154.

Fehr, E., & List, J. A. (2004). The hidden costs and returns of incentives—trust and trustworthiness among ceos. *Journal of the European Economic Association, 2*(5), 743–771. https://doi.org/10.1162/1542476042782297

Fehr, E., & Schmidt, K. M. (2004). Fairness and incentives in a multi-task principal-agent model. *The Scandinavian Journal of Economics, 106*(3), 453–474.

Fischbacher, U. (2007). Z-tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*. https://doi.org/10.1007/s10683-006-9159-4

Fisman, R., Kariv, S., & Markovits, D. (2007). Individual preferences for giving. *American Economic Review, 97*(5), 1858–1876. https://doi.org/10.1257/aer.97.5.1858

Fréchette, G. R. (2015). Laboratory experiments: professionals versus students. In G. R. Fréchette & A. Schotter (Eds.), *Handbook of experimental economic methodology* (pp. 360–390). Oxford University Press.

Fréchette, G. R. (2016). Experimental economics across subject populations. In J. H. Kagel & A. E. Roth (Eds.), *the handbook of experimental economics, volume two*. Princeton University Press.

Gagnon, N., Bosmans, K., and Riedl, A. (2020). The effect of unfair chances and gender discrimination on labor supply. Working Paper.

Greiner, B. (2015). Subject pool recruitment procedures: organizing experiments with ORSEE. *Journal of the Economic Science Association, 1*(1), 114–125.

Gross, T., Guo, C., & Charness, G. (2015). Merit pay and wage compression with productivity differences and uncertainty. *Journal of Economic Behavior and Organization, 117*, 233–247. https://doi.org/10.1016/j.jebo.2015.06.009

Hoppe, E. I., & Schmitz, P. W. (2013). Contracting under incomplete information and social preferences: An experimental study. *The Review of Economic Studies, 80*(4), 1516–1544. https://doi.org/10.1093/restud/rdt010

Itoh, H. (2004). Moral hazard and other-regarding preferences. *The Japanese Economic Review, 55*(1), 18–45. https://doi.org/10.1111/j.1468-5876.2004.00273.x

Kocher, M. G., Pogrebna, G., & Sutter, M. (2013). Other-regarding preferences and management styles. *Journal of Economic Behavior and Organization, 88*, 109–132. https://doi.org/10.1016/j.jebo.2013.01.004

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

Koszegi, B. (2014). Behavioral contract theory. *Journal of Economic Literature, 52*(4), 1075–1118. https://doi.org/10.1257/jel.52.4.1075

von Siemens, F. A. (2011). Heterogeneous social preferences, screening, and employment contracts. *Oxford Economic Papers, 63*(3), 499–522. https://doi.org/10.1093/oep/gpq028

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

https://doi.org/10.1007/s10683-023-09791-0 Published online by Cambridge University Press