cesifo
WORKING
PAPERS

8942
2021

March 2021

Cheating in Labour Markets

Michalis Drouvelis, Martin G. Kocher

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editor: Clemens Fuest
https://www.cesifo.org/en/wp

CESifo *Working Paper No. 8942*

Cheating in Labour Markets

Abstract

Our results from a laboratory experiment offer new evidence for the detrimental effects that cheating behaviour in the workplace may have on the degree of reciprocity between firms and workers. First, we replicate existing findings showing that in the absence of monitoring (cheating is possible) workers cheat on their actual performance in a real-effort task. The extent of cheating influences how firms (managers) decide to set their wages in a subsequent gift-exchange game. Specifically, firms offer higher wages to workers who cheat and interestingly, workers expect such behaviour by firms. These higher wages are not matched by workers' performance in the gift-exchange game, where cheating is not possible, resulting in a flat wage-effort relationship. In contrast, in the presence of monitoring (cheating is not possible), we find a positive relationship between wages offered and effort provided by the workers. Our findings have implications for adopting measures at the workplace that eliminate workers' opportunities to cheat on their performance.

JEL-Codes: C720, C910, D210.

Keywords: cheating, gift-exchange, labour market, experiment.

Michalis Drouvelis
University of Birmingham / UK
M.Drouvelis@bham.ac.uk

Martin G. Kocher
Institute for Advanced Studies
University of Vienna / Austria
martin.kocher@univie.ac.at

March 2021
We are thankful for valuable comments we received at the EWEBE conference 2019, the ESA conference 2019 and many research seminars and workshops. Martin Kocher gratefully acknowledges financial support by Deutsche Forschungsgemeinschaft through CRC TRR 190.

1. Introduction

On labour markets with complete contracts, i.e. on markets on which effort levels can be fully enforced or contracted upon and on which wages lead to market clearing, economic models can comfortably abstract from potentially undesirable behaviours of workers. Firms will just offer optimal piece-rate schemes or sell shares of the firm to workers as incentive devices. As soon as contracts on labour markets are incomplete or there are limits to screening and enforcement, optimal labour market relationships between firms and workers become more complicated, but they also might become more efficient in case workers have strong enough intrinsic motivation to provide additional effort.

The fair wage-effort hypothesis (Akerlof, 1982) stipulates that workers voluntarily provide more than minimal enforceable effort levels in case they receive higher than minimum wages. The higher wage is perceived as a gift by the firm, and workers reciprocate to it with higher effort levels. A large literature has analysed whether the positive wage-effort relationship exists both in the experimental laboratory (e.g., Fehr et al., 1993; Falk and Fehr, 1998; Fehr et al., 1998; Brown et al., 2004) and in the field (e.g., Gneezy and List, 2006; Falk, 2007; Bellemare and Shearer, 2009; Dur, 2009; Kube et al., 2012). The bulk of the evidence confirms the existence of the fair wage-effort relationship. On average, reciprocity is stronger than shirking motives even in one-shot interactions, as long as the potential efficiency gain from the reciprocal firm-worker interaction is sufficiently large.

Shirking is ex post behaviour that jeopardizes the mutually beneficial firm-worker interaction, just as free riding in a social dilemma or moral hazard in an insurance context. However, there is also ex ante hazards to the positive wage-effort relationship. Consider the case of job applications. Do applicants over-state certain aspects of their experience or brag too much about their educational achievement? Do human resource departments or firms expect some cheating by applicants regarding their productivity and adjust their wage offers accordingly? How are applicants treated by firms on a labour market that is comprised of honest and cheating applicants? And does cheating on productivity have the potential to flatten the positive wage-effort relationship on labour markets with incomplete contracts?

While ex post behaviour has been studied extensively in the context of principle-agent (gift-exchange) games in the experimental laboratory, ex ante problems such as cheating have not yet been analysed, to the best of our knowledge, in this context. This paper looks at a parsimonious experimental labour market with real effort provision in which the option to cheat on one's productivity is either possible or not. Firms see potentially distorted productivity signals when they make wage offers to workers. After matching of firms and workers, workers provide effort and cannot cheat anymore. In such a setup cheating might affect the wage offer relationship and lead to adverse selection dynamics that crowd out the well-known positive relationship between wage and non-enforceable effort levels.

Our research design is related to a quickly growing strand of the literature that analyses the trade-off between being honest and forgo some profit, on the one hand, and being dishonest and earn more, on the other hand. At least since Becker (1968) economists have been interested in the way decision makers trade off benefits and costs of a decision that involve a moral component such as filing a tax declaration, paying for public transport, voting in a general election or a referendum, etc. While Becker mentions potential psychological costs of breaking a moral norm, he neglects them in his analysis. Recent evidence (e.g., Gneezy, 2005; Mazar et al., 2008; Sutter, 2009; Fischbacher, 2013) shows that there are indeed psychological costs of lying, cheating, and other violations of everyday norms. Rosenbaum et al. (2014) survey the literature, and Abeler et al. (2019) provide a meta study with the main result that two motivations can explain the consistent finding of decisions makers to follow moral norms at a high cost to them, even if there are no consequences such as fines or punishment for norm violation: a genuine preference for complying with the norm and a preference for being seen as compliant with the norm.

Interestingly, the vast majority of existing empirical assessments that rely mostly on laboratory experiments as a consequence of the difficulty to observe some norm violations outside the laboratory and to abstract from existing punishment schemes have so far focused on individual decisions. Only, recently several studies on cheating or lying in groups or teams – in comparison to cheating or lying as an individual – have been published (e.g., Weisel and Shalvi, 2015; Kocher et al., 2018). A recent discussion in the literature also looks at whether markets crowd out morality (Falk and Szech, 2013; Kirchler et al., 2016). However, there seems a void in research on the market consequences of individual cheating that we want to start to fill with this paper.

More specifically, we use a variant of the task in Mazar et al. (2008) that has been implemented first by Klimm (2019). Participants in the experiment have to find to numbers with two decimals in a matrix that add up to exactly 10.00. Subjects are instructed to solve as many matrices as possible within three minutes, and we paid a piece rate for every correctly solved matrix. In the monitoring treatment participants know that the experimenter will check their solutions (cheating is not possible). In the non-monitoring treatment we ask them to report the number of matrices that they solved correctly themselves, without checking the solutions. In order to make the absence of monitoring credible, participants are told, before they report the number of correctly solved matrices, that they will be requested to shred their solution sheets. The experimenter was not able to verify the actual number of correctly solved matrices on the individual level. However, by comparing the means of the number of correctly solved matrices in the treatments, cheating on the population level can be detected by the researcher.

After the matrix task participants were randomly assigned roles, either “manager” or “worker”. Managers acted as firms that were asked to make fixed wage offers to all workers on the market. They saw how many matrices the workers (reported to) have solved correctly – depending on the treatment – and also the gender of the worker. We hypothesized that there might be a gender difference in cheating and perhaps also in the expectation regarding the inclination to cheat in such a setup. Each manager was randomly paired with a worker and the wage offer for this specific worker by the relevant manager was implemented for real. Afterwards, workers worked, and the manager received a multiple of the number of correctly solved matrices, but of course they had to pay the wage to the worker. Hence, the market is a bilateral gift-exchange game, with no enforcement option for the managers. Performance in this second part was monitored; thus, cheating was not possible anymore in all treatments. We implemented treatments with detailed information (instructions) on the second section part already before the first part (hence, potential spill-overs from cheating in the first part to the labour market could be taken into account by workers) and without detailed information on the second part. Interestingly, there was very little difference between the two information conditions regarding the second part. Our general setup is reminiscent of a paper by Englmaier et al. (2014) in which they study the relative importance of productivity and trust as well as trustworthiness for a subsequent labour market with a gift-exchange nature.

Our experimental results offer new evidence identifying the detrimental effects that cheating behaviour at the workplace may have on the degree of reciprocity between managers and workers. First, we replicate existing evidence showing that, in the absence of monitoring (cheating is possible), subjects cheat on their actual performance by self-reporting higher numbers of correctly solved pairs of matrices in a real-effort task. This cheating behaviour influences how managers decide to set their wages in a subsequent gift-exchange game. Specifically, managers offer higher wages for workers who cheat and interestingly, workers expect such behaviour by managers. It seems that managers do not fully take into account that workers could have been cheating. Given the absence of cheating possibilities in the performance after labour market matching, these higher wages are not matched by workers' performance in the task, rendering the wage-effort relationship insignificant. In contrast, in the presence of monitoring (cheating is not possible), we find a positive relationship between wages offered and effort chosen by the workers. Our findings provide implications for adopting ex ante measures that eliminate employees' opportunities to cheat on their previous performance. They also show that the positive wage-effort relationship might depend on the availability of such ex ante measures, e.g., screening devices to detect cheating on productivity on the individual level.

The rest of the paper is organized as follows: Section 2 introduces the details of our experimental design. In Section 3, we provide our experimental results with regard to cheating in part 1 and the choices in part 2. Finally, Section 4 concludes the paper.

2. Experimental Design and Hypotheses

We implement a between-subjects design consisting of three treatments. Each treatment has two parts: part 1 has subjects perform a real-effort task, and part 2 has subjects participate in a one-shot gift exchange game, i.e. a labour market that determines wages of workers followed by the same real-effort task as in part 1 to determine worker effort. Our experimental design varies along two dimensions: i) subjects are or are not provided with opportunities to mis-report their performance in part 1 (cheating is possible or not); and ii) subjects are or are not aware of the content of part 2 when they perform the real-effort task during part 1 and report their productivity. In all treatments, subjects knew in advance that the experiment will consist of two parts.

Prior to the beginning of part 1, subjects were asked to fill in a brief questionnaire eliciting basic demographic information. In particular, we asked questions about subjects' age, gender and fields of study.[1] Once everyone in the laboratory had completed this brief questionnaire, we proceeded with the main experiment which consisted of the two parts. Below, we describe in detail each of the parts.

---

[1] Subjects were asked to complete this questionnaire without receiving any information about the content of part 1 and part 2 of the experiment.

2.1 Part 1: the matrix task

During Part 1, subjects were asked to perform the so-called “matrix task” introduced by Mazar et al. (2008). Specifically, subjects were given a sheet of paper with a series of 10 different pairs of matrices containing nine non-integer numbers each and were asked to find in each of the pairs two numbers that add up to exactly 10 (see the Appendix for the 10 pairs of matrices that subjects were provided in part 1). Subjects were told that they had three minutes to solve as many of the pairs as possible and that they get paid based on how many correct answers they provide (self-report), depending on the treatment. For each correctly (reported as) solved pair of matrices, subjects received £0.50; the piece rate was clearly stated in the instructions.

After subjects performed the matrix task, we varied across treatments whether their performance was monitored or not. In the former condition (“non-cheating condition”), the experimenter verified the number of correctly solved pairs of matrices; while in the latter condition (“cheating condition”), subjects were asked to self-report the number of matrices they had solved correctly without the experimenter verifying the correct pairs of matrices.[2] To ensure that cheating opportunities were possible, at the beginning of the experiment subjects were told that they can destroy their work sheets through a paper shredder at the back of the room. This was also the case in the condition where cheating is not possible, as we wanted to follow exactly the same procedures in both conditions. This means that after the number of solved pair of matrices had been verified by the experimenter in the “non-cheating conditions”, the sheets of papers were also shredded. At the end of part 1, subjects were informed about their (self-reported) performance in the matrix task.

By observing behaviour in part 1, we obtain a measure of subjects’ performance under a piece rate scheme. In the remainder of the analysis, we therefore refer to “Part 1 performance” as the number of correct answers or self-reported correct answers in this task. Any difference observed between the cheating and the non-cheating conditions is attributed to the fact that subjects can lie and mis-report their performance in the former condition but not in the latter one.

In addition, we elicited incentivised beliefs about what subjects’ think the average performance of others (excluding themselves) in the matrix task is (always only within their treatments). Subjects were compensated with £1 if their guess was exactly correct, £0.50 if their guess deviated by one pair of matrices (in either direction), £0.20 if their guess deviated by two pairs of matrices (in either direction) and £0 if their guess deviated by more than two pairs of matrices (in either direction) from the actual average performance of others. Subjects were not provided with feedback regarding the accuracy of guesses about others’ behaviour before the end of the experiment.

Following part 1, subjects participated in part 2 which we outline below.

2.2 Part 2: gift exchange game

During part 2, a one-shot gift-exchange game was implemented in which the task to be fulfilled as a worker is identical to the real effort task in part 1 of the experiment, but cheating was not

---

[2] A screenshot for reporting the number of correctly solved pairs of matrices in the real-effort task can be found in Appendix A (Figure A.1).

possible in any of the treatments. We adopted our setting by following the paper of Englmaier et al. (2014). Subjects were randomly allocated to be either a manager (firm) or a worker at the beginning of part 2. The role assignment was independent from subjects' performance in part 1, and this information was clearly stated in the instructions.[3] We employ the strategy method in order to elicit managers' wage decisions. In particular, managers have to submit a binding wage offer (ranging from £0 up to £8, inclusive, using up to two decimals) for each of the workers in the session such that we obtain the full wage profile that managers are submitting for all workers.[4] The information that the managers had when setting their wage offers for the workers was their performance from part 1 of the experiment (which could be either verified (monitored) or self-reported (non-monitored), depending on the treatment) and their gender (as elicited from the questionnaire given to subjects before the beginning of part 1). After all wage offers had been submitted, every worker was matched randomly to a single manager, i.e. every firm hires only one worker.

Subsequently, workers perform the matrix task as in part 1 for a period of three minutes. Workers learn only the wage offer that their matched managers have determined for them before they start working for their manager. Thus, there is no possibility for workers to be influenced by offers that the manager has submitted to other workers or by offers that other managers have submitted to them. The interaction is one-shot to rule out any effects of repetition over time and to focus in the most parsimonious way on the effects of part 1 performance on wage setting decisions by managers.

Workers' performance then determines the earnings of the manager in part 2, π_M, according to the following equation:

$$\pi_M = 2X - W$$

where X is equal to the number of correctly solved pairs of matrices by the randomly matched worker and W is the wage.

Workers are paid their predetermined fixed wage, and we assume for the sake of succinctness that they have non-monetary costs of effort from solving the matrix task (in any case, costs of effort cannot be observed in our real-effort environment). Hence, workers earn:

$$\pi_W = W$$

It is possible for managers to make losses if the wage exceeds the revenues generated by their worker (i.e., $W > 2X$). Subjects were informed that losses had to be paid from earnings in part 1 of the experiment. Given the nature of the task which requires no particular training for performing it, there should be little learning possibilities from doing the task a second time. In part 2 of the experiment, subjects' performance was monitored and verified by the experimenter in all treatments. This was clearly stated and announced in the instructions that subjects received for part 2. After the real-effort task in part 2 was completed, subjects received feedback about the number of correct answers and the payoff to the manager and the worker.

---

[3] Note that the matrix task in part 1 was performed by both the managers and the workers, enabling both parties to become familiar with the nature of the task.

[4] A screenshot for the wage setting decisions of the managers in the experiment is provided in Appendix A (Figure A.2).

Both manager and workers learn only the details from their interaction and not from the interaction between any other manager-worker pair.

To gain a better understanding of the managers' and workers' behavioural patterns, we elicited incentivised beliefs before giving any feedback. In particular, we asked managers to indicate what they thought that the performance of their matched worker in part 2 was, and we asked workers to indicate the expected wage that their matched manager would offer to them. Belief elicitation was done before managers (workers) were informed of the workers' performances (managers' wages).[5]

2.3 Treatments and procedures

Our main treatment variation consists in the pieces of information elicited in part 1 of the experiment that were made available to managers when submitting their wage profiles. In our two main treatments, information from part 1 performance is provided to managers which set their wage offers, knowing whether part 1 performance had been monitored or not. This allows us to compare whether and, if so, how cheating opportunities when reporting part 1 performance affects wage setting by managers. In both of these treatments, the details of the content of part 2 was not announced until after part 1 was completed. We refer to the two resulting treatments in the remainder of the paper as follows: i) the "Monitored – Part 2 is not announced" treatment, whereby part 1 performance is verified by the experimenter (cheating is not possible) and subjects are not aware of the content of part 2; ii) the "Non-monitored – Part 2 is not announced" treatment, whereby part 1 performance is not verified by the experimenter (cheating is possible), and subjects are not aware of the details of the content of part 2. We also included a treatment where part 1 performance is not monitored but the details of the content of part 2 is announced to subjects from the beginning of the experiment. We refer to this treatment as the "Non-monitored – Part 2 is announced" treatment.[6] Table 1 provides an overview of our experimental treatments, presenting the number of observations we collected and the number of sessions we conducted in each treatment.[7]

---

[5] To incentivise belief elicitation, we employed the same scoring rule as in part 1. Specifically, if allocated the role of manager (worker), subjects received £1 if their guess was exactly correct, £0.50 if their guess deviated by one matrix (in either direction), £0.20 if their guess deviated by two matrices (in either direction) and £0 if their guess deviated by more than two matrices (in either direction) from the actual average performance of their matched worker in part 2 (from the actual wage offered by their matched manager).

[6] The inclusion of a "Monitored – Part 2 is announced" treatment seemed not necessary for answering our main research questions. Had it turned out that the ex ante announcement of Part 1 would have been relevant for behaviour in part 1, a fourth treatment ("Non-monitored – Part 2 is announced") could have had revealed potential motivation effects from the announcement. Since announcing part 2 had no effect on choices, we did not implement the treatment.

[7] In each treatment, subjects received the instructions reproduced in Appendix B.

Table 1. Overview of experimental treatments

Treatments	Number of observations	Number of observations per session (S)	
Monitored – Part 2 is not announced	100	S1: 14 S2: 18 S3: 18	S4: 18 S5: 14 S6:18
Non-monitored – Part 2 is not announced	108	S1: 18 S2: 20 S3: 18	S4: 14 S5: 20 S6: 18
Non-monitored – Part 2 is announced	104	S1: 16 S2: 18 S3: 16	S4: 18 S5: 18 S6: 18

Procedures: The experiment was conducted in the BEEL laboratory at the University of Birmingham in March 2019 with the help of z-tree (Fischbacher, 2007). A total number of 312 subjects from various academic backgrounds were recruited via ORSEE (Greiner, 2015). Subjects remained anonymous throughout the experiment, and cash payments were made privately at the end of the experiment. On average, an experimental session lasted for about 45 minutes. Subjects earned an average of £7.55 (including a show-up fee of £2.50).

2.4 Hypotheses

This study is partly exploratory, and we want to be clear about this. Expected effects depend on assumptions regarding the inclination to cheat and the sophistication of managers to fully take potential cheating into account. If decision makers are rational and selfish, the prediction is straightforward for the one-shot game: Since effort provision is costly and effort cannot be enforced, workers will shirk in part 2, and managers will – correctly expecting shirking – choose the lowest permissible wage, regardless of the productivity signal they get from performance in part 1. In part 1, when cheating is possible, rational and selfish decision makers will report the highest possible performance. Gender should not play any role.

We know from many previous experiments that there is a positive relationship between wages and effort that relies on the trust of managers in the reciprocity of workers. Our working hypothesis is that the ex ante cheating possibility creates a distorted productivity signal. If the signal is ignored by the managers, there is no a priori reason to assume that the positive relationship is attenuated. It is also possible, at least in principle, that the willingness to provide more effort for higher wages is independent of whether workers are inclined to cheat on productivity. Then, the productivity signal and potential cheating in part 1 would have no effect on the subsequent labour market.

Let us assume for the sake of the argument that workers cheat about their performance X when cheating is possible, that managers do not fully take cheating into account when setting wages W and thus expect workers to provide more output ($b(X)$), and that workers cannot put in enough effort in order to provide the expected output. Subscripts C and NC refer to non-monitored (cheating) conditions and the monitored (non-cheating) condition, respectively. Hence in part 2:

$$W_C > W_{NC}$$

$$b(X)_C > b(X)_{NC}$$

$$X_C \approx X_{NC}$$

As a consequence, the wage-output (effort) relationship should become weaker in the cheating conditions than in the non-cheating condition.

While our design does not allow studying the correlation between cheating inclination and reciprocal tendency on the individual level – we do not observe cheating on the individual level – it helps us disentangle some source of labour market outcomes with ex ante cheating possibilities.

We can disentangle at least two potential mechanisms that might contribute to a weakening of the wage-effort relationship: belief-based explanations and preference-based explanations. Belief-based explanations require two ingredients. First, performance in part 1 must have an impact on performance beliefs of managers about workers. Second, managers do not or not fully take cheating on performance in part 1 into account when making wage offers. Preference-based explanations are less well-defined. Managers could have preferences over being honest or over other characteristics that are independent of their beliefs about workers' performance.

There is evidence that females tend to lie and cheat less than males (for a survey of results, see Rosenbaum et al., 2014), but the tendency is not very pronounced. However, we know very little on the perception of gender when it comes to interpret an honest or dishonest signal. Since we thought that there could be a stronger stereotype than the actual difference in honesty between females and males, we hypothesized that revealing gender information could lead to different wage offers to female and male workers in the treatments where cheating was possible in part 1.

3. Experimental results

In the description of the results that follows, we begin by examining subjects' behaviour in part 1 of the experiment in the three treatments. We then investigate behaviour in the gift exchange game in part 2 by analysing managers' wage offers and workers (monitored) performance in parts 1 and 2. Finally, we analyse how the relationship between wages and effort provision in part 2 is influenced when cheating opportunities are (not) possible.

3.1 Performance in part 1

Figure 1 shows the distribution of number of pairs of matrices (reported as) solved in part 1 across treatments. We observe that, when monitoring is possible in part 1, subjects' correct number of pairs of matrices are mainly concentrated on the left-hand side part of the distribution with the vast majority (79% of subjects) correctly solving between (and including) 0 and 3 pairs of matrices. Under monitoring, there is no subject who solved more than 7 pairs of matrices correctly. This is in stark contrast to the distributions of the other two treatments where cheating is possible. For these two treatments, we observe that approximately one third of subjects report more than 7 pairs (30.56% of subjects in the "Non-monitored – Part 2 is not announced" treatment and 27.88% in the "Non-monitored – Part 2 is announced" treatment), and in contrast, very few subjects report between 0 and 3 pairs of matrices (23.15% of subjects in the "Non-monitored – Part 2 is not announced" treatment and 19.23% in the "Non-monitored – Part 2 is announced" treatment). A Kruskal-Wallis test[8] indicates statistically significant differences in how the pairs of matrices (reported as) solved correctly in part 1 are distributed in the treatments (p<0.001).

Figure 1. Distribution of number of pairs of matrices (reported as) solved in part 1

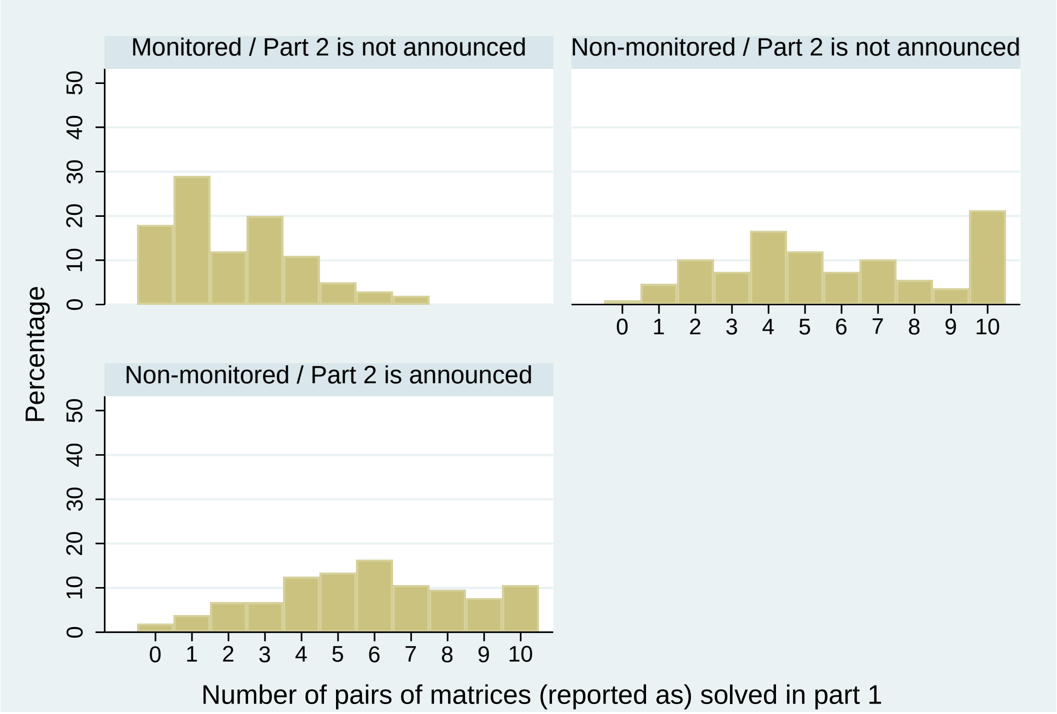


Table 1 presents the average number of pairs of matrices (reported as) solved in part 1 as well as subjects' beliefs about others' average performance in part 1 in the three treatments. Starting with the actual performance, we find that, on average, subjects provide 2.14 correct answers in the "Monitored – Part 2 is not announced" treatment, whereas the average number of pairs of matrices reported as solved is 5.81 and 5.79 in the "Non-monitored – Part 2 is not announced" treatment and the "Non-monitored – Part 2 is announced" treatment, respectively. A Mann-Whitney test shows significant differences between the "Monitored – Part 2 is not announced" and the other two treatments where cheating is possible (2.14 vs. 5.81, p<0.001; 2.14 vs. 5.79, p<0.001). However, the average reported number of correctly solved pairs of matrices is not significantly different between the two treatments with cheating being possible, regardless of whether part 2 is announced or not (5.81 vs. 5.79, p=0.908).[9]

---

[8] All statistical tests in the following are two-sided tests.

[9] When we look separately at the distribution of (self-reported) performance for those subjects allocated the role of managers and the role of workers in part 2, we obtain similar results. Remember that nobody knew his or her role at this stage of the experiment, yet. In particular, we find that, on average, managers' average performance in part 1 is: 1.96 in the "Monitored – Part 2 is not announced" treatment, 5.98 in the "Non-monitored – Part 2 is not announced" treatment, and 5.90 in the "Non-monitored – Part 2 is announced" treatment. The corresponding

Table 2. Average number of pairs of matrices (reported as) solved and average beliefs about others' average performance in part 1

Treatment	Average number of pairs of matrices (reported as) solved in part 1	Average beliefs about others' average performance in part 1
Monitored – Part 2 is not announced [n=100]	2.14 (1.76)	3.38 (1.42)
Non-monitored – Part 2 is not announced [n=108]	5.81 (2.95)	5.70 (2.11)
Non-monitored – Part 2 is announced [n=104]	5.79 (2.62)	5.45 (2.12)

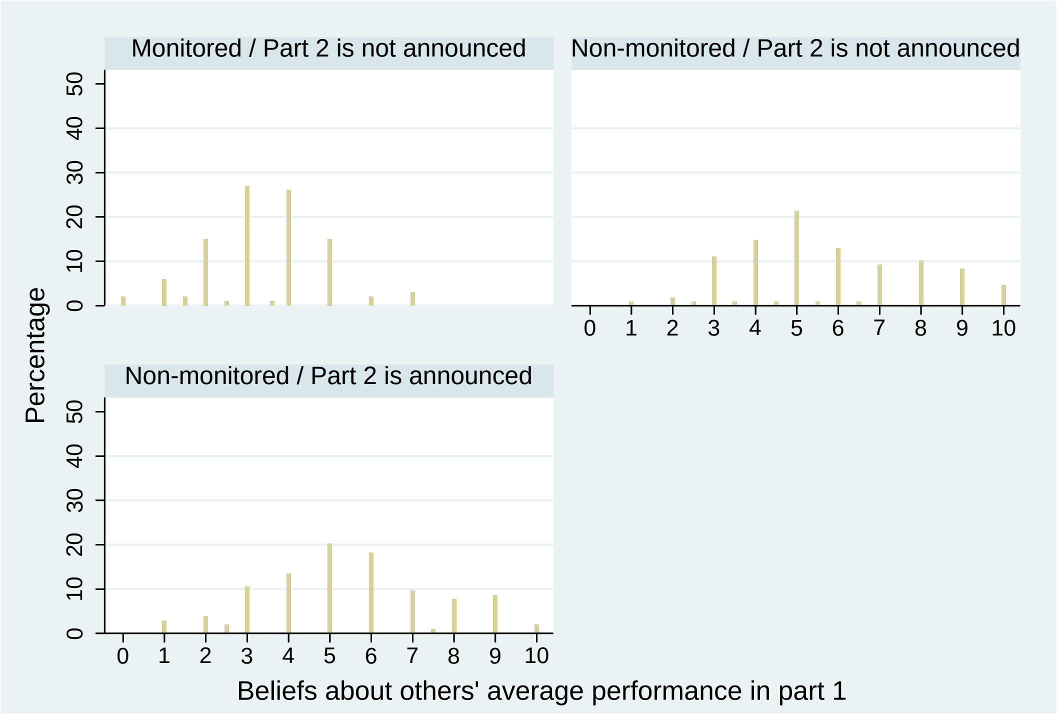
Note: Standard deviations are reported in parentheses. Number of observations are reported in square brackets.

A similar pattern of behaviour is obtained when we look at subjects' beliefs about others' performance in part 1. In sum, we find that subjects expect others to also report higher levels of reported performance when cheating is possible. Figure 3 shows the distribution of beliefs about others' performance in the three treatments. A Kruskal-Wallis test again suggests significant differences across treatments (p<0.001). Beliefs and behaviour are strongly correlated on the aggregate level. Correlation coefficients are positive and highly significant (p<0.01) for all three treatments. Interestingly the correlation coefficients are even higher in the non-monitored treatments ("Non-monitored – Part 2 is not announced": 0.56; "Non-monitored – Part 2 is announced": 0.65) than in the monitored treatment (0.32).

---

average performance of workers in part 1 in each treatment is: 2.32 in the "Monitored – Part 2 is not announced" treatment, 5.63 in the "Non-monitored – Part 2 is not announced" treatment, and 5.67 in the "Non-monitored – Part 2 is announced" treatment. By comparing differences in performance across roles (managers vs. workers) for a given treatment, we do not find any statistically significant difference (for any pairwise comparison, p>0.223). This implies that managers and workers performed/reported similarly in our three treatments. By examining treatment differences across treatments, for a given role, we find that there are significant differences when we compare treatments with and without monitoring (any pairwise comparison yields p<0.001 for both managers and workers), but not when we compare the two treatments with monitoring (p=0.869 for managers and 0.707 for workers). This implies that both managers and workers report significantly higher numbers of correctly solved matrices in the treatments without monitoring compared to the verified performance when monitoring is possible. Since assignment of roles was random, none of these results are surprising.

Figure 2. Distribution of beliefs about others' average performance in part 1



By looking at average beliefs as reported in Table 1, we find that, on average, subjects expect others to provide 3.38 correct answers in the "Monitored – Part 2 is not announced" treatment, whereas, the average expected number of pairs of matrices reported as solved by others is 5.70 and 5.45 in the "Non-monitored – Part 2 is not announced" treatment and the "Non-monitored – Part 2 is announced" treatment, respectively. A Mann-Whitney test shows significant differences between the "Monitored – Part 2 is not announced" and the other two treatments where cheating is possible (3.38 vs. 5.70, p<0.001; 3.38 vs. 5.45, p<0.001). However, the average expected reported number of correctly solved pairs of matrices by others is not significantly different between the two treatments where cheating is possible, regardless of whether part 2 is announced (5.70 vs. 5.45, p=0.526).

Taken together, our analysis from part 1 provides strong evidence that, in the absence of monitoring, subjects report significantly higher performance levels than in the monitoring treatment. In addition, subjects expect that others' average (reported) performance will be higher in the treatments when there is no monitoring. Our first result is summarised as follows.

Result 1 *Without monitoring, (self-reported) performance is significantly higher than when monitoring is available. Without monitoring, subjects also expect that others report significantly higher levels of performance.*

3.2 Gift exchange game behaviour in part 2

Our analysis starts by first examining managers' behaviour in terms of wage setting in part 2 across treatments. We then turn to workers' performance in part 2. Finally, we look at how the wage-effort relationship varies for the three treatments.

3.2.1. Managers' behaviour

Figure 3 shows the distribution of wage offers in the three treatments. We observe that, in the "Monitored – Part 2 is not announced" treatment, most managers make low wage offers to their workers: 80% of managers offer less than £3. By contrast, only 10% offer more than £4, with no manager offering an amount greater than £5. In the treatments where cheating is possible, we observe that a significant minority of managers make offers above £3 (31.40% in the "Non-monitored – Part 2 is not announced" treatment and 44.23% in the "Non-monitored – Part 2 is announced" treatment). There are also a few managers who make offers above £7. Still, the majority of them offer less than £3, but importantly, this percentage drops to 68.50% in the "Non-monitored – Part 2 is not announced" treatment and 55.77% in the "Non-monitored – Part 2 is announced" treatment. A Kruskal-Wallis test indicates significant differences in the wage offers across treatments (p=0.007).

Figure 3. Distribution of managers' wage offers in part 2

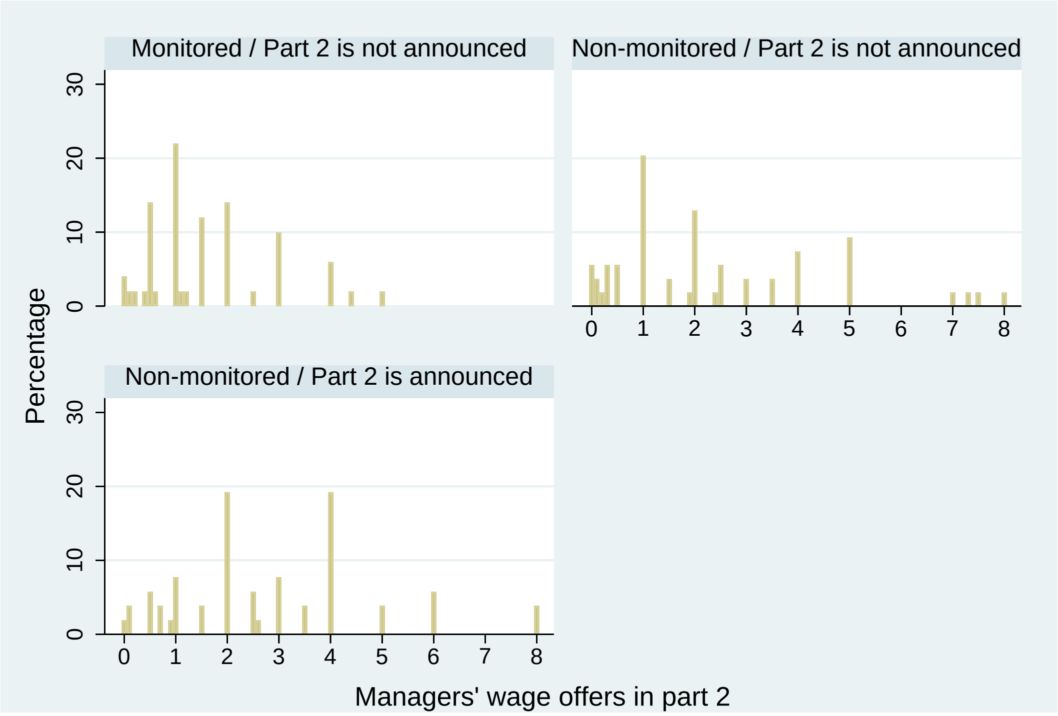


Table 3 reports the average wage offers made by managers in part 2 as well as the corresponding workers' beliefs about managers' wages across treatments. Starting with managers' wage decisions, and as observed in Figure 2, we find that, on average, managers offer wages of £1.60 in the "Monitored – Part 2 is not announced" treatment, whereas the average wage offers are £2.34 and £2.77 in the "Non-monitored – Part 2 is not announced" treatment and the "Non-monitored – Part 2 is announced" treatment, respectively. A Mann-Whitney test shows significant differences between the "Monitored – Part 2 is not announced" and the "Non-monitored – Part 2 is announced" treatments (1.60 vs. 2.77, p=0.001). For the other two pairwise comparisons, we do not find statistically significant differences at conventional levels (1.60 vs. 2.34, p=0.175 and 2.34 vs. 2.77, p=0.122). By pooling the two treatments where monitoring is not present (i.e. “Non-monitored – Part 2 is not announced” and “Non-monitored – Part 2 is announced”), we find that, in the two treatments without monitoring in part 1, average wage offers are significantly higher than in the monitoring treatment (1.60 vs. 2.55, p=0.008).

Turning to workers’ beliefs about managers’ wages, we obtain a similar behavioural pattern in the sense that, in the treatments without monitoring in part 1, workers expect managers to offer higher wages compared to the treatment where monitoring is present. Specifically, on average, workers expect that managers will offer a wage of £2.34 in the “Monitored – Part 2 is not announced” treatment, which is significantly lower than the average expected wage of £3.29 in the “Non-monitored – Part 2 is not announced” treatment (p=0.005) and of £3.08 in the “Non-monitored – Part 2 is announced” treatment (p=0.019). We also find that average expected wages by workers do not differ significantly between the two treatments where there is no monitoring of reported performance in part 1 (£3.29 vs. £3.08, p=0.648).

Table 3. Average wage offers made by managers and workers’ beliefs about managers’ wage offers in part 2

	Average (implemented) wage offers by managers in part 2	Average workers’ beliefs about managers’ wage offers in part 2
Monitored – Part 2 is not announced [n=50]	1.60 (1.22)	2.34 (1.36)
Non-monitored – Part 2 is not announced [n=54]	2.34 (2.09)	3.29 (1.81)
Non-monitored – Part 2 is announced [n=52]	2.77 (1.90)	3.08 (1.72)

Note: Standard deviations are reported in parentheses. Number of observations are reported in square brackets.

Our results are confirmed by performing parametric analyses as presented in Table 4. In particular, we report four regression models where the dependent variable is manager’s wage offers in part 2 (Models 1-3) and worker’s expected wage (Model 4). Since the dependent variable in both regressions is censored from above and below (as the minimum (maximum) wage that a manager can offer to their matched worker is £0 (£8)), we us Tobit models. The independent variables of the models comprise: i) a dummy variable called “Non-monitored” which is equal to 1 for the treatments without monitoring (i.e. both the “Non-monitored – Part 2 is announced” and the “Non-monitored – Part 2 is not announced” treatments), and 0 for the “Monitored – Part 2 is not announced” treatment; ii) a dummy variable called “Part 2 is announced” which is equal to 1 for the treatments where part 2 is announced (i.e. both the “Non-monitored – Part 2 is announced” and the “Monitored – Part 2 is not announced” treatments), and 0 for the “Non-monitored – Part 2 is not announced” treatment; iii) a dummy variable capturing gender effects denoted “Female” that equals 1 for female workers and 0 otherwise; (iv) the number of (self-reported) correct answers in part 1; and (v) in interaction term of (i) and (iv). Our regression results are shown in Table 4.

Table 4. Wage and workers' beliefs about managers' wage offers in part 2 – Regression results

	Dependent variable: Manager's (implemented) wage offer			Dependent variable: Worker's beliefs about manager's wage offer
	Model (1)	Model (2)	Model (3)	Model (4)
Non-Monitored (=1 if part 1 performance is not monitored, 0 otherwise)	0.769** (0.345)	1.414*** (0.447)	2.104*** (0.588)	0.989*** (0.314)
Part 2 is announced (=1 if part 2 is announced, 0 otherwise)	0.452 (0.398)	0.435 (0.387)	0.432 (0.387)	-0.263 (0.343)
Worker's gender (=1 if female, 0 otherwise)	-0.163 (0.300)	-0.243 (0.300)	-0.251 (0.298)	-0.289 (0.269)
(Self-reported) correct answers in part 1	-	-0.159** (0.063)	0.069 (0.098)	-
(Self-reported) correct answers in part 1 X Non-monitored	-	-	-0.268** (0.121)	-
Constant	1.640*** (0.257)	1.992*** (0.292)	1.550*** (0.311)	2.487*** (0.223)
Obs.	156	156	156	156

*Notes: Tobit estimates. Robust standard errors are presented in parentheses. *** denotes significance at the 1-percent level, ** denotes significance at the 5-percent level, and *** at the 1-percent level.*

Three main observations stand out from Table 4. First, the coefficient of the dummy variable "Non-monitored" is significant at least at the 5% level in Models (1) to (3), implying that managers' wage offers are significantly higher in the treatments where performance in part 1 was not monitored. We obtain the same pattern when we look at Model (4) with beliefs as the dependent variable. Workers expect managers to offer significantly higher wages in the treatments without monitoring. Second, we find no significant relationships for "Part 2 is announced" and "Female" in any of the models. This indicates that announcing the content of part 2 or not as well as the worker's gender did not affect manager wage setting. Third, managers seem to be aware of potential mis-reporting of performance in part 1. Particularly, in the non-monitored treatments, the number of self-reported correct answers from part 1 reduces the positive effect of the non-monitoring treatments on wage offers (see Model 3). Obviously, they still underestimate mis-reporting, given the magnitude and the significance of the "Non-monitored" dummy in Model (3). Interestingly the sign of the coefficient is reversed, when we add the number of self-reported correct answers from part 1 to the regression on worker's beliefs (not shown). We have also looked at models that use the squared term of "(self-reported) correct answers in part 1" in order to address potential non-linearities. Not surprisingly this squared term is negative and significant, indicating that too high (self-reported) correct answers may be perceived as suspicious by managers. Thus, they relationship between (self-reported) correct answers and wages is inverse-U-shaped.

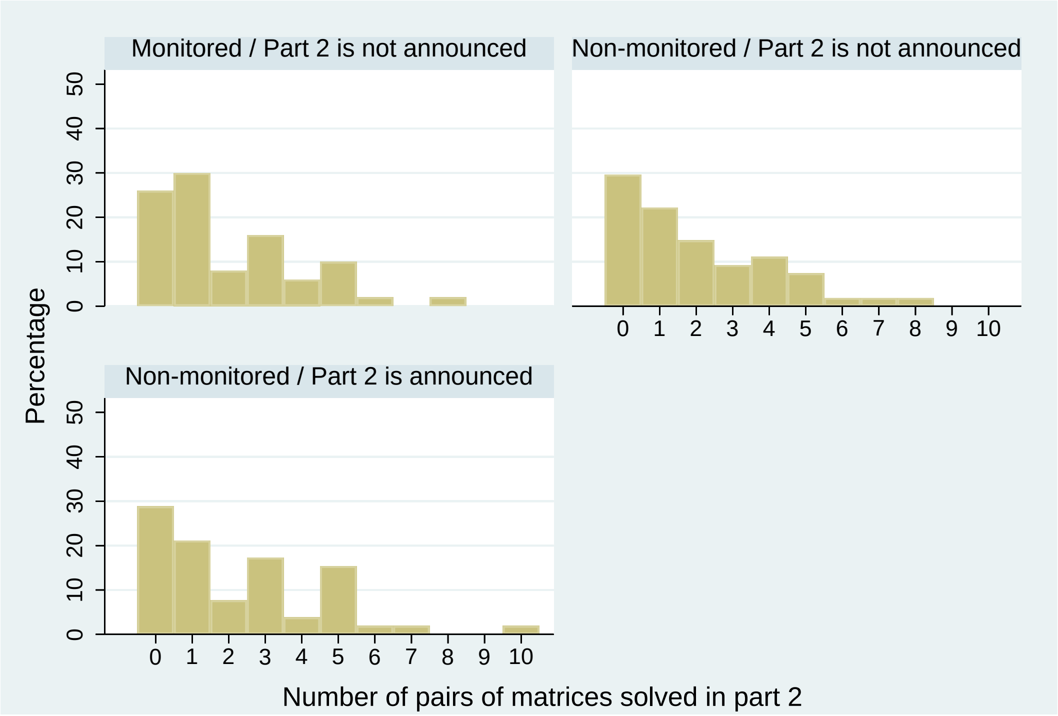
Following our analysis in this section, we summarise our main findings as follows.

Result 2 *Managers offer, on average, significantly higher wages in the absence of performance monitoring in part 1. Workers also expect higher wages, on average, in the treatments without monitoring.*

3.2.2. Workers' behaviour

Our next step is to analyse workers' behaviour (as measured by observing their performance) in part 2, where monitoring is always available. Figure 4 shows the distribution of the number of pairs of matrices solved in part 2 in the three treatments. We observe that the distributions of performance in part 2 are similar across treatments. This is confirmed by performing a Kruskal-Wallis test (p=0.879).

Figure 4. Distribution of number of pairs of matrices solved in part 2



We next look at how average workers' performance in part 2 is shaped. It is apparent that workers solve similar numbers of pairs of matrices regardless of the treatment (using a Mann-Whitney test, any pairwise comparison yields p>0.649). Table 5 also reports managers' expectations about the workers' performance in part 2.[10]

---

[10] The distribution of managers' expectations about workers' performance in part 2 can be found in Appendix C (Figure C.3).

Table 5. Average number of pairs of matrices solved and average managers' beliefs about workers' performance in part 2

	Average number of correct answers for workers in part 2	Average manager's beliefs about workers' correct answers in part 2
Monitored – Part 2 is not announced [n=50]	1.96 (1.94)	2.84 (1.63)
Non-monitored – Part 2 is not announced [n=54]	2 (2.03)	4.48 (2.13)
Non-monitored – Part 2 is announced [n=52]	2.25 (2.27)	4.38 (2.36)

Note: Standard deviations are reported in parentheses. Number of observations are reported in square brackets.

Our findings indicate that managers' expectations do not match closely workers' performance in part 2. We find that managers expect workers to perform significantly better in part 2 in the treatments without monitoring in part 1. When we compare manager's average beliefs about workers' performance between the "Monitored – Part 2 is not announced" treatment and either of the "Non-monitored – Part 2 is not announced" or the "Non-monitored – Part 2 is announced" treatments, we find statistically significant differences (the two Mann-Whitney tests, $p<0.01$). In contrast, we do not find any statistically significant differences in managers' average expectations when monitoring in part 1 is not possible (4.48 vs. 4.38, $p=0.794$). Taken together with Result 2, which shows that managers offer higher wages in the non-monitoring treatments, our analysis here indicates that this result is likely to be associated with the higher expectations about workers' performance in part 2 that managers have in the treatments without monitoring.

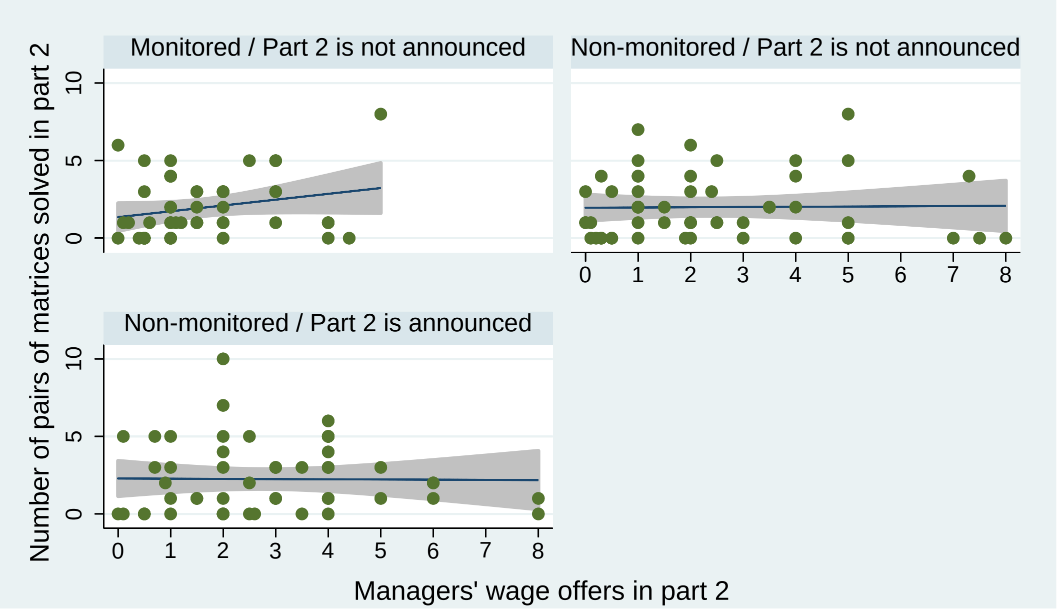
Our main findings form this section can be summarised as follows.

Result 3 *Workers' performance in part 2 is not statistically different in our three treatments. In contrast, managers expect workers to perform, on average, significantly better in part 2 in the treatments where there was no monitoring in part 1.*

3.2.3 The wage-effort relationship

In this section, we try to gain a better understanding of the relationship between managers' wages and workers' performance in part 2. Previous experimental studies, many of which rely on abstract instead of real-effort provision, have shown that typically there is a positive relationship between wages offered by managers and effort provided by the workers. In the following analysis, we examine whether the wage-effort relationship is affected when cheating is possible. In Figure 5, we plot managers' implemented wage as a function of workers' performance in part 2 for each of the three treatments separately. Note that the managers first set up their wage offers to workers based on the number of pairs of matrices (reported as) solved in part 1, depending on the treatment. After observing the wage offered by their matched manager, the workers decide how much effort to put in by performing the matrix task as in part 1, but with full monitoring in all treatments.

Figure 5. The wage-effort relationship across treatments



Note: The shaded areas indicate the 95% confidence interval around the linear regression line.

It can be seen that, when performance monitoring in part 1 is possible, there is a positive relationship between wage offers chosen by the managers and the performance of workers in part 2. This observation is confirmed by performing a Pearson correlation test that finds that the wage-effort relationship is positive and weakly significant ($\rho=0.236$, $p<0.10$) for the "Monitored – Part 2 is not announced" treatment. In the absence of monitoring in part 1, our previous analysis shows that managers offer (high) wages that are not matched by the performance of the workers. In turn, as shown in Figure 5, the wage-effort relationships in both the "Non-monitored – Part 2 is not announced" treatment and the "Non-monitored – Part 2 is announced" treatment are virtually flat. A Pearson correlation test indicates that the slope of the wage-effort relationship is not significantly different from zero in either of the two treatments ("Non-monitored – Part 2 is not announced" treatment: $\rho=0.015$, $p=0.913$; "Non-monitored – Part 2 is announced" treatment: $\rho=-0.011$, $p=0.941$).

Result 4 *The presence of cheating opportunities breaks the positive wage-effort relationship. In contrast, when monitoring is possible, wages are positively associated with workers' effort in part 2.*

3.2.4 Additional analyses

As supplementary analysis, we also look at managers' wage profiles, conditional on the workers' performance as measured by the number of pairs of matrices (reported as) solved in part 1. Figure 6 shows the relationship between these two variables of interest when the full wage profile of managers is plotted. It suggests that in the treatments where monitoring of performance is possible, there is a positive correlation between wage setting and part 1 performance ("Monitored – Part 2 is not announced" treatment: Pearson correlation test, $\rho=0.544$, $p<0.001$). For the treatments where monitoring is not possible, Figure 6 suggests that there is a weaker association between wage offers and part 1 performance. By performing a

Pearson correlation test, we nonetheless find significant correlations in both treatments ("Non-monitored – Part 2 is not announced" treatment: Pearson correlation test, ρ=0.235, p<0.001; "Non-monitored – Part 2 is announced" treatment: Pearson correlation test, ρ=0.251, p<0.001). Notably, we observe that, in the treatments without monitoring, the magnitude of the correlation coefficients is less than half of the counterpart in the treatment with monitoring. To check whether these differences are significant, we perform a test of equality of pairwise correlation coefficients. We find that there are statistically significant differences in the two pairwise comparisons between the correlation coefficients of the monitored treatments vs. the coefficient of the non-monitored treatment (ρ=0.544 vs. ρ=0.235, p<0.001 and ρ=0.544 vs. ρ=0.251, p<0.001), but not between the two treatments where monitoring is not possible (ρ=0.235 vs. ρ=0.251, p=0.800). This provides evidence that the relationship between managers' wages in part 2 and workers' part 1 performance becomes weaker in the presence of cheating opportunities.

Figure 6. Managers' wage profiles as a function of workers' performance in part 1 across treatments

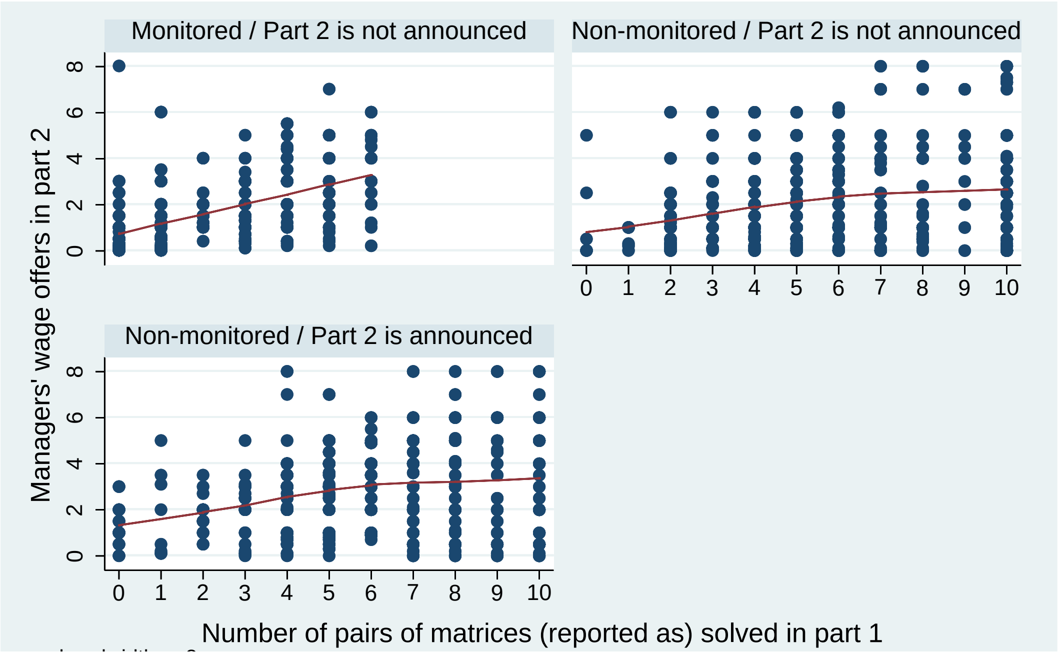


Figure 6 also indicates that the positive and significant relationship between wage profiles and workers' performance in part 1 does not seem to hold across the whole distribution of their performance. To test for this possibility, we calculate the median performance in each treatment separately and then estimate the correlation coefficients between managers' wage offers and part 1 performance when the latter falls below the median and above the median (reported) performance. Starting with the "Monitored – Part 2 is not announced" treatment, we find that the median part 1 performance is equal to 2 pairs of matrices. The correlation coefficient between managers' wage offers and workers' below (above) the median performance is ρ=0.321 (ρ=0.338). In both cases (i.e. for the below and above the median performance), we find that the correlation coefficients are highly statistically significant (p<0.001). We next turn to the “Non-monitored – Part 2 is not announced” treatment with a median reported performance in part 1 of 5 pairs of matrices. In this treatment, the corresponding correlation coefficient between managers’ wage offers and workers’ below (above) the median performance is ρ=0.241 (ρ=0.038). We observe that managers’ wage levels are not significantly correlated with high self-reported performance (p=0.585) as opposed to low self-reported performance (p<0.001). A similar pattern is obtained when we consider the “Non-monitored – Part 2 is announced” treatment. Here, the median reported performance is 6 pairs of matrices, and we find that below (above) the median performance, there is (not) a significant correlation between wage offers and corresponding performance (specifically, for below the median performance: ρ=0.365, p<0.001 and for above the median performance: ρ=0.041, p=0.612). Taken together, our analysis indicates that managers offer higher wages, the higher the workers’ performance is from part 1 in all three treatments. Importantly, however, this relationship holds only for low (below the median) (self-reported) performance levels across all treatments. When we look at high (above the median) (self-reported) performance levels, we find that the relationship between wage setting and performance is flat and the corresponding correlation coefficients are very small in magnitude. This suggests that, in the presence of cheating opportunities, managers do not associate positively their wage offers with the workers’ reported performance above the median but rather the relationship becomes flat (still it remains at (too) high levels).

We then look at determinants of managers’ wages again. In Table 6, the three models from Table 4 are re-evaluated, using an additional independent variable: the expectation (belief) of the manager regarding the performance of the worker in part 2. Remember that we elicited beliefs, using monetary incentives. Not surprisingly, the belief has a significant and robust effect on managers’ wages. Since, it was elicited after setting the wage offer, it is not clear whether beliefs are fully independent. However, first of all, the results from Table 6 show that wages are consistent with beliefs about performance and behaviour follows a clear pattern. Second, knowing that performance in part 1 influences these beliefs, managers fail to take potential cheating into account. The coefficient seems to be very robust in the three models, both in terms of significance and in terms of magnitude.

Table 6. Wages in part 2 – Regression results

	Dependent variable: Manager's (implemented) wage offer		
	Model (1)	Model (2)	Model (3)
Non-Monitored (=1 if part 1 performance is not monitored, 0 otherwise)	0.027 (0.317)	0.549 (0.393)	1.279** (0.507)
Part 2 is announced (=1 if part 2 is announced, 0 otherwise)	0.508 (0.352)	0.496 (0.344)	0.495 (0.344)
Worker's gender (=1 if female, 0 otherwise)	-0.163 (0.260)	-0.225 (0.258)	-0.234 (0.255)
(Self-reported) correct answers in part 1	-	-0.123** (0.054)	0.119 (0.088)
(Self-reported) correct answers in part 1 X Non-monitored	-	-	-0.286*** (0.106)
Mangers' beliefs on part 2 performance	0.442*** (0.081)	0.427*** (0.079)	0.430*** (0.078)
Constant	0.385 (0.284)	0.702** (0.297)	0.224 (0.349)
Obs.	156	156	156

*Notes: Tobit estimates. Robust standard errors are presented in parentheses. *** denotes significance at the 1-percent level, ** denotes significance at the 5-percent level, and *** at the 1-percent level.*

We have also looked at gender effects, in addition to Table 4 that does not reveal a significant gender effect regarding wages to males and females. Interestingly, we do not observe any significant gender effects for our main variables of interest. Even when we consider the full wage profiles assigned by the firms to each worker, we find that neither gender nor gender matching between managers and workers plays a significant role in determining wages. Since we have enough statistical power, we think that it is relevant to mention these null results explicitly.

4. Discussion and conclusion

We have implemented an experiment that allows studying cheating at the application stage of a labour market that is characterized by incomplete contracts. Our results from a laboratory experiment show that managers do not take the cheating option fully into account (even though they have experienced the opportunity themselves), when making wage offers. In other words, managers clearly underestimate cheating. Wages are much higher in the non-monitored treatments than in the monitored treatment. Managers thus expect also higher levels of output from their workers in the non-monitored treatments. As a consequence, the positive wage-effort relationship is destroyed when the selection decision is based on productivity data that might be distorted by cheating. Gender does not seem to play a significant role (nor does the interaction of gender and cheating).

It seems that our main results can be explained by managers that do not take cheating into account when setting wages. They expect much higher levels of performance by workers, based on self-reported part 1 performance than can be realized by workers in part 2. Hence, we conclude that our treatment differences are belief-based and not preference-driven.

We think that the experiment provides relevant findings. If cheating at the application stage is possible, the standard result of a positive wage-effort relationship does not hold anymore. It suggests that labour markets in which cheating about past performance is easier should suffer more from the lack of monitoring than labour markets in which cheating is more difficult, for instance as a consequence of signalling through the use of credible devices such as diplomas or exam results.

Obviously, we use a very stylized setting, because we thought that the literature should start off with a parsimonious design. Numerous relevant extensions or changes of the basic setup are conceivable and potentially interesting. One important route for further research is repeated interaction on the labour market. Managers could learn quickly that workers cheated and decrease wages. However, on many labour markets such reactions would be impossible or come with other costs. It would also be very interesting to study the effects of better signaling options for honest workers. Finally, one could look at a second part in which cheating is also possible, and firms can invest in technologies that detect cheating. However, such a setup would be much more complicated and does not make much sense in a bilateral market on which always on worker is matched with a manager. Obviously, our setting implements a specific technology that implies a specific efficiency gain from trade and potential ceiling effects for effort provision. Hence, setups with different parameters could address the generalizability of our results.

References

Abeler J, Nosenzo, D, Raymond C (2019) Preferences for truth-telling. Econometrica 87(4), 1115–1153.

Akerlof, GA (1982) Labor contracts as partial gift exchange. Quart. J. Econ. 97, 543–569.

Becker, GS (1968) Crime and punishment: an economic approach. J. Polit. Econ. 76, 169–217.

Bellemare, C, Shearer, B (2009) Gift giving and worker productivity: Evidence from a firm-level experiment. Games Econom. Behav, 67(1), 233–244.

Brown, M, Falk, A, Fehr, E (2004) Relational contracts and the nature of market interactions. Econometrica 72, 747–780.

Dur, R (2009) Gift exchange in the workplace: Money or attention? J. Europ. Econom. Assoc. 7(2-3), 550–560.

Englmaier, F, Strasser, S, Winter, J (2014) Worker characteristics and wage differentials: Evidence from a gift-exchange experiment. J. Econom. Behav. and Organ. 97(1): 185–203.

Falk, A (2007) Gift exchange in the field. Econometrica 75, 1501–1511.

Falk, A, Szech, N (2013). Morals and markets. Science. 340: 707–711.

Fehr, E, Falk, A (1999) Wage rigidities in a competitive incomplete contract market. J. Polit. Econ. 107, 106–134.

Fehr, E, Kirchsteiger, G, Riedl, A (1993) Does fairness prevent market clearing? An experimental investigation. Quart. J. Econ. 108, 437–460.

Fehr, E, Kirchler, E, Weichbold, A, Gächter, S (1998) When social norms overpower competition: Gift exchange in experimental labor markets. J. Labor Econ. 16, 324–351.

Fischbacher U (2007) z-Tree: Zurich toolbox for ready-made economic experiments. Experiment. Econom. 10(2):171–178.

Fischbacher U, Föllmi-Heusi F (2013) Lies in disguise—An experimental study on cheating. J. Eur. Econom. Assoc. 11(3):525–547.

Gneezy U (2005) Deception: The role of consequences. Amer. Econom. Rev. 95(1):384–394.

Gneezy, U, List, JA (2006) Putting behavioral economics to work: Testing for gift exchange in labor markets using field experiments. Econometrica 74, 1365–1384.

Greiner B (2015) Subject pool recruitment procedures: Organizing experiments with ORSEE. J. Econom. Sci. Assoc. 1(1):114–125.

Kirchler, M, Huber, J, Stefan, M, Sutter, M (2016) Market design and moral behaviour. Man. Sci. 62(9): 2615–2625.

Klimm F (2019) Suspicious success – Cheating, inequality acceptance, and political preferences. Europ. Econom. Rev. 117: 36-55.

Kocher M, Schudy S, Spantig L (2018) I Lie? We Lie! Why? Experimental evidence on a dishonesty shift in groups. Man. Sci. 64(9): 3995–4008.

Kube, S, Maréchal, MA, Puppe, C (2012) The currency of reciprocity: Gift exchange in the workplace. American Economic Review 102(4), 1644–1662.

Mazar N, Amir O, Ariely D (2008) The dishonesty of honest people: A theory of self-concept maintenance. J. Marketing Res. 45(6): 633–644.

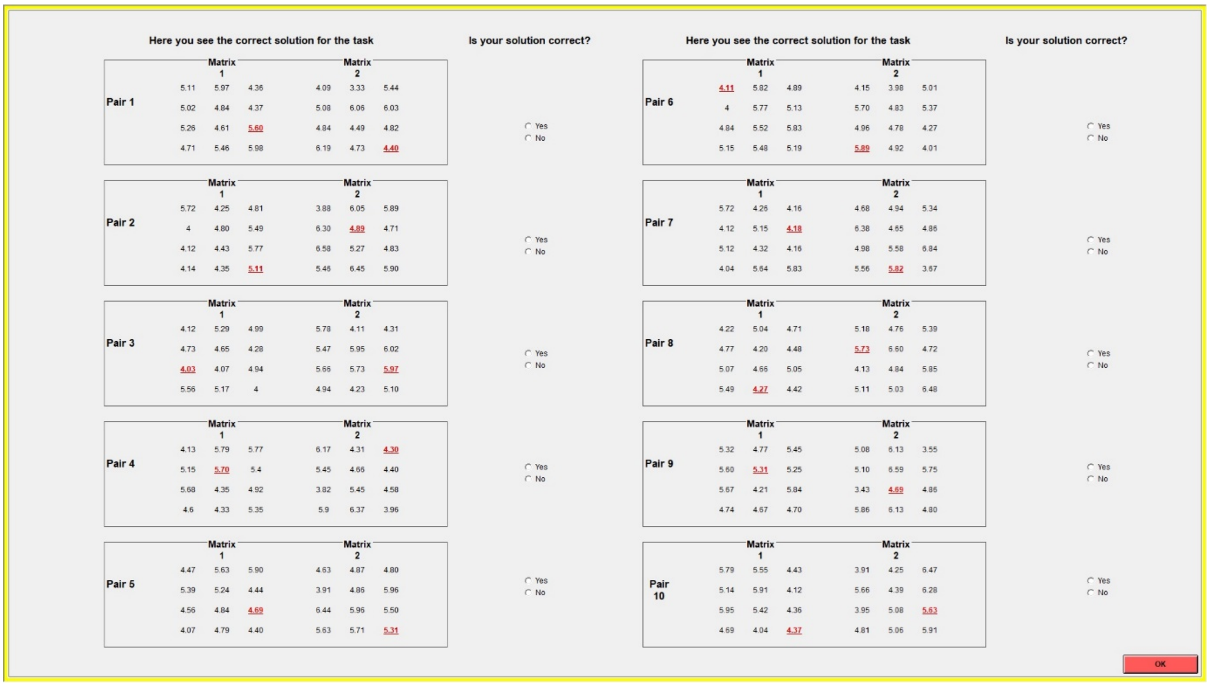
Rosenbaum SM, Billinger S, Stieglitz N (2014) Let’s be honest: A review of experimental evidence of honesty and truth-telling. J. Econom. Psych. 45(4):181–196.

Sutter M (2009) Deception through telling the truth?! Experimental evidence from individuals and teams. Econom. J. 119(534):47–60.

Weisel O, Shalvi S (2015) The collaborative roots of corruption. Proc. National Acad. Sci. USA 112(34):10651–10656.

Appendix A – Screenshots (not intended for publication)

Figure A.1. Screenshot for reporting the number of correctly solved pairs of matrices in the real effort task in part 1



Here you see the correct solution for the task — **Is your solution correct?**

	Matrix 1			Matrix 2			
Pair 1	5.11	5.97	4.36	4.09	3.33	5.44	○ Yes ○ No
	5.02	4.84	4.37	5.08	6.06	6.03	
	5.26	4.61	**5.60**	4.84	4.49	4.82	
	4.71	5.46	5.98	6.19	4.73	**4.40**	

	Matrix 1			Matrix 2			
Pair 2	5.72	4.25	4.81	3.88	6.05	5.89	○ Yes ○ No
	4	4.80	5.49	6.30	**4.89**	4.71	
	4.12	4.43	5.77	6.58	5.27	4.83	
	4.14	4.35	**5.11**	5.46	6.45	5.90	

	Matrix 1			Matrix 2			
Pair 3	4.12	5.29	4.99	5.78	4.11	4.31	○ Yes ○ No
	4.73	4.65	4.28	5.47	5.95	6.02	
	4.03	4.07	4.94	5.66	5.73	**5.97**	
	5.56	5.17	4	4.94	4.23	5.10	

	Matrix 1			Matrix 2			
Pair 4	4.13	5.79	5.77	6.17	4.31	**4.30**	○ Yes ○ No
	5.15	**5.70**	5.4	5.45	4.66	4.40	
	5.68	4.35	4.92	3.82	5.45	4.58	
	4.6	4.33	5.35	5.9	6.37	3.96	

	Matrix 1			Matrix 2			
Pair 5	4.47	5.63	5.90	4.63	4.87	4.80	○ Yes ○ No
	5.39	5.24	4.44	3.91	4.86	5.96	
	4.56	4.84	**4.69**	6.44	5.96	5.50	
	4.07	4.79	4.40	5.63	5.71	**5.31**	

Here you see the correct solution for the task — **Is your solution correct?**

	Matrix 1			Matrix 2			
Pair 6	**4.11**	5.82	4.89	4.15	3.98	5.01	○ Yes ○ No
	4	5.77	5.13	5.70	4.83	5.37	
	4.84	5.52	5.83	4.96	4.78	4.27	
	5.15	5.48	5.19	**5.89**	4.92	4.01	

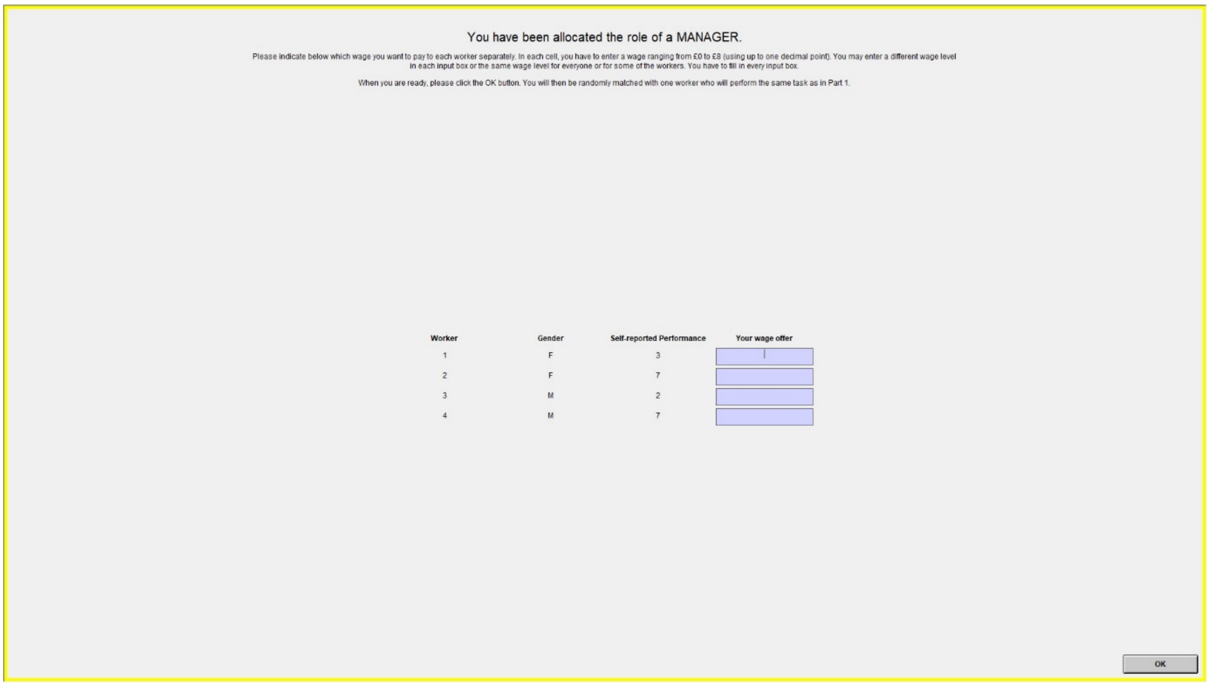
	Matrix 1			Matrix 2			
Pair 7	5.72	4.26	4.16	4.68	4.94	5.34	○ Yes ○ No
	4.12	5.15	**4.18**	6.38	4.65	4.86	
	5.12	4.32	4.16	4.98	5.58	6.84	
	4.04	5.64	5.83	5.56	**5.82**	3.67	

	Matrix 1			Matrix 2			
Pair 8	4.22	5.04	4.71	5.18	4.76	5.39	○ Yes ○ No
	4.77	4.20	4.48	**5.73**	6.60	4.72	
	5.07	4.66	5.05	4.13	4.84	5.85	
	5.49	**4.27**	4.42	5.11	5.03	6.48	

	Matrix 1			Matrix 2			
Pair 9	5.32	4.77	5.45	5.08	6.13	3.55	○ Yes ○ No
	5.60	**5.31**	5.25	5.10	6.59	5.75	
	5.67	4.21	5.84	3.43	**4.69**	4.86	
	4.74	4.67	4.70	5.86	6.13	4.80	

	Matrix 1			Matrix 2			
Pair 10	5.79	5.55	4.43	3.91	4.25	6.47	○ Yes ○ No
	5.14	5.91	4.12	5.66	4.39	6.28	
	5.95	5.42	4.36	3.95	5.08	**5.63**	
	4.69	4.04	**4.37**	4.81	5.06	5.91	

Figure A.2. Screenshot for the wage setting decisions of the managers in part 2



You have been allocated the role of a MANAGER.

Please indicate below which wage you want to pay to each worker separately. In each cell, you have to enter a wage ranging from £0 to £8 (using up to one decimal point). You may enter a different wage level in each input box or the same wage level for everyone or for some of the workers. You have to fill in every input box.

When you are ready, please click the OK button. You will then be randomly matched with one worker who will perform the same task as in Part 1.

Worker	Gender	Self-reported Performance	Your wage offer
1	F	3	
2	F	7	
3	M	2	
4	M	7	

Appendix B – Experimental Instructions

[Note: These are the written instructions as presented to subjects facing the "Monitored – Part 2 is not announced" treatment. Amendments to the other two treatments are given in square brackets.]

Instructions

Welcome! You are about to take part in a decision-making experiment. This experiment is run by the "Birmingham Experimental Economics Laboratory" and has been financed by various research foundations. Just for showing up you have already earned £2.50. You can earn additional money depending on the decisions made by you and other participants. It is therefore very important that you read these instructions with care.

It is important that you remain silent and do not look at other people's work. If you have any questions, or need assistance of any kind, please raise your hand and an experimenter will come to you. If you talk, laugh, exclaim out loud, etc., you will be asked to leave and you will not be paid. We expect and appreciate your following of these rules.

We will first jointly go over the instructions. After we have read the instructions, you will have time to ask clarifying questions. We would like to stress that any choices you make in this experiment are entirely anonymous. Please do not touch the computer or its mouse until you are instructed to do so. Thank you.

The amount that you earn during the experiment will privately be paid to you in cash at the end of the experiment.

The experiment consists of two parts. The total amount you will earn from the experiment will be the sum of the earnings you make in the two parts of the experiment as specified in the instructions below. At the beginning of the experiment, we would like you to answer a brief questionnaire.

Once everybody has completed this questionnaire, we will detail the instructions for Part 1. You will receive new instructions for Part 2 once everyone in the room has completed Part 1. [*Part 2 is announced treatment*: Once everybody has completed this questionnaire, we will detail the instructions for Part 1 and Part 2.]

Detailed Information about Part 1

Your task

You have three minutes to find, in each pair of the following 10 pairs of matrices, which two numbers add up to exactly 10 (one number from Matrix 1 plus one number from Matrix 2). The 10 pairs of matrices can be found in the white envelope on your desk. Please do not open this envelope until you are instructed to do so.

Please circle the numbers that add up to 10 in each pair of matrices clearly.

Here is an example:

Matrix 1

4.12	5.29	4.99
4.73	4.65	4.28
4.03	4.07	4.94
5.56	5.17	4

Matrix 2

5.78	4.11	4.31
5.47	5.95	6.02
5.66	5.73	5.97
4.94	4.23	5.10

The correct solution for this pair is:

Matrix 1

4.12	5.29	4.99
4.73	4.65	4.28
[✓ 4.03]	4.07	4.94
5.56	5.17	4

Matrix 2

5.78	4.11	4.31
5.47	5.95	6.02
5.66	5.73	[✓ 5.97]
4.94	4.23	5.10

Earnings from Part 1

For each pair of correctly solved matrices, you earn £0.50. You have to perform this task sequentially by starting from the first pair of matrices (Pair 1) moving next to the second one etc. in that order until Pair 10. Starting from Pair 1, you are not allowed to move to the next pair of matrices if you have not solved correctly the previous pair of matrices. After the 3 minutes are finished, the experimenter will compare your solutions to the correct solutions. You will be paid according to the number of correct solutions you provide. At the end of the experiment, you will have to destroy the TASK sheet that you have received at the start of Part 1 by using the paper shredder in the back of the room.

[*Not monitored treatments:* After the 3 minutes are finished, you will be asked to compare your solutions to the correct solutions. You will be paid according to the number of correct solutions you self-report. At the end of the experiment, you will have to destroy the TASK sheet that you have received at the start of Part 1 by using the paper shredder in the back of the room. This ensures that we cannot trace back your solutions.]

On the screen you will be informed about your earnings from Part 1 that will be paid to you after the end of Part 2. The TASK sheets are identical for all participants.

If you have any questions, please raise your hand and one of the experimenters will come to your desk to answer your question.

Detailed Information about Part 2

At the beginning of Part 2, you will be assigned a role: either as manager or as worker. Role assignment is random and completely independent of your performance in Part 1.

Your task as a worker:
You have three minutes to find, in each pair of the following 10 pairs of matrices, which two numbers add up to exactly 10 (one number from Matrix 1 plus one number from Matrix 2). The task is exactly the same as in Part 1. This time you work for a manager who earns more money the more matrices you solve correctly, and you will be paid a fixed wage for your work. In this part, the experimenters will verify the number of correctly solved pairs of matrices.

Your task as a manager:
Before the workers can start working, you have to offer the worker a wage in return for them performing the same task as in Part 1. The wage is a fixed amount between (and including) 0 and 8 Pounds (up to one decimal point in increments of £0.10) that you pay to the worker regardless of the performance of the worker. When you set the wage, you do not know yet which worker you will be paired with. Hence, you have to indicate for each worker in the room, how much you want to pay to them.

How you set your wage decisions as a manager:
Managers see a list of workers, indicating their verified performance from Part 1 and their gender. For each worker you have to indicate a wage that you want to pay. Managers may enter a different wage level in each input box or the same wage level for everyone or for some of the workers. Managers have to fill in every input box, i.e. they are required to make a wage offer to every worker. After this, one manager is randomly matched with one worker. The manager will then be informed which wage has been implemented and the worker will learn the wage. At this point, workers will start to work by performing for three minutes the same task as in Part 1 (the pairs of matrices are identical for all workers, but different from Part 1).

Earnings from Part 2

- Workers: Workers earn a fixed wage between (and including) 0 and 8 Pounds indicated by their manager before they start working.

Earnings of a worker = wage

- Managers: Managers earn 2 multiplied by the number of correctly solved pairs of matrices by their worker minus the wage they offered to that worker.

Earnings of a manager = 2 × (Number of correctly solved pairs of matrices by their worker)
minus wage

Managers might make losses in case the offered wage is higher than 2 × (Number of correctly solved pairs of matrices by their worker). In such a case, managers have to pay any losses in Part 2 from their earnings in Part 1.

At the end of Part 2, managers and workers are informed about outcomes (correctly solved pairs of matrices) and your total earnings from today's experiment.

If you have any questions, please raise your hand and one of the experimenters will come to your desk to answer your question.

Appendix C – Additional analysis

Figure C.1. Distribution of number of pairs of matrices (reported as) solved in part 1 (if allocated the role of managers in part 2)

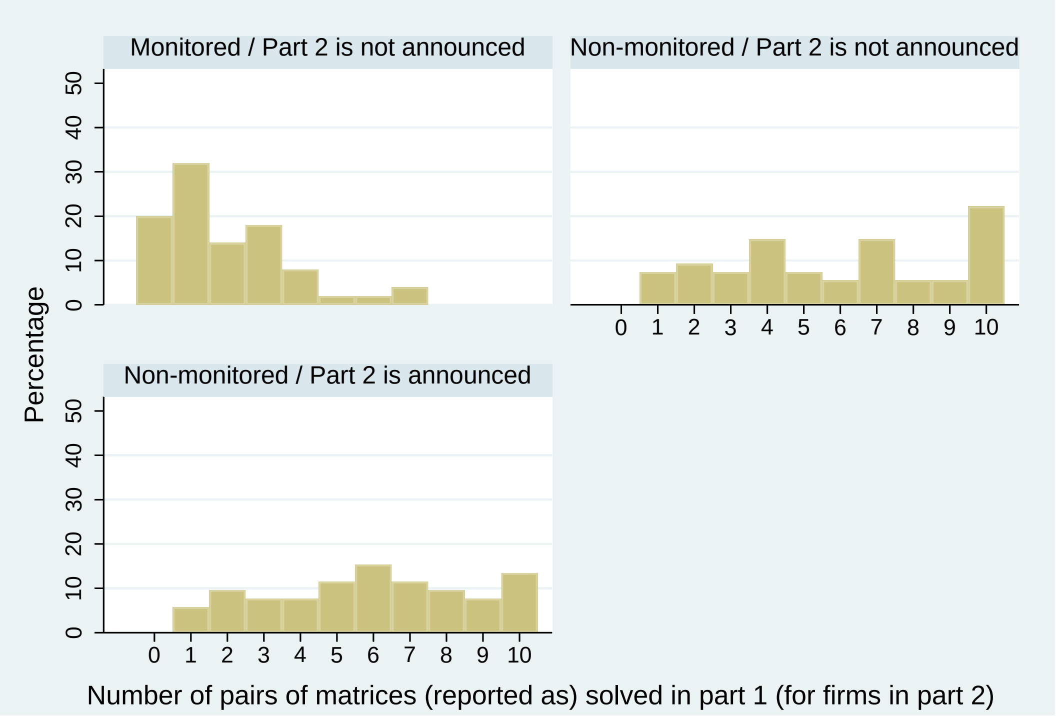


Figure C.2. Distribution of number of pairs of matrices (reported as) solved in part 1 (if allocated the role of worker in part 2)

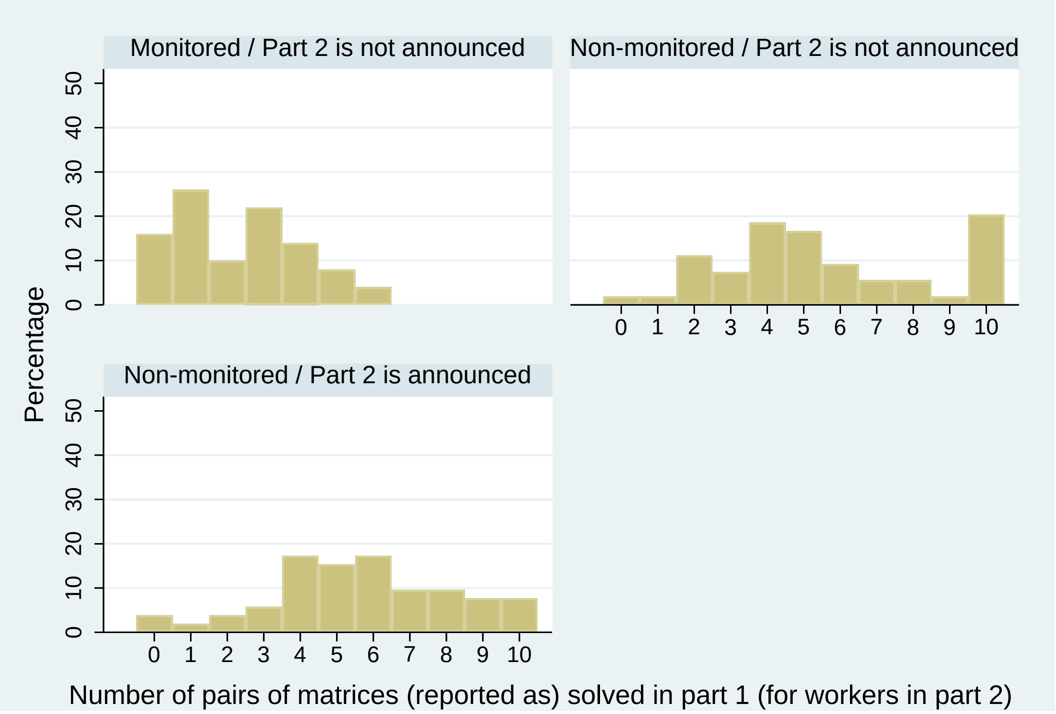


Figure C.3. Distribution of managers' beliefs about workers' performance in part 2