Heinz, Matthias; Normann, Hans-Theo; Rau, Holger A.

Working Paper

How competitiveness may cause a gender wage gap: Experimental evidence

DICE Discussion Paper, No. 213

Provided in Cooperation with:
Düsseldorf Institute for Competition Economics (DICE)

Suggested Citation: Heinz, Matthias; Normann, Hans-Theo; Rau, Holger A. (2016) : How competitiveness may cause a gender wage gap: Experimental evidence, DICE Discussion Paper, No. 213, ISBN 978-3-86304-212-7, Düsseldorf Institute for Competition Economics (DICE), Düsseldorf

This Version is available at:
http://hdl.handle.net/10419/128594

Terms of use:
Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.
How Competitiveness May Cause a Gender Wage Gap: Experimental Evidence

Matthias Heinz, Hans-Theo Normann, Holger A. Rau

February 2016
DICE DISCUSSION PAPER

Published by
düsseldorf university press (dup) on behalf of
Heinrich-Heine-Universität Düsseldorf, Faculty of Economics,
Düsseldorf Institute for Competition Economics (DICE), Universitätsstraße 1,
40225 Düsseldorf, Germany
www.dice.hhu.de

Editor:
Prof. Dr. Hans-Theo Normann
Düsseldorf Institute for Competition Economics (DICE)
Phone: +49(0) 211-81-15125, e-mail: normann@dice.hhu.de

DICE DISCUSSION PAPER

All rights reserved. Düsseldorf, Germany, 2016

ISSN 2190-9938 (online) – ISBN 978-3-86304-212-7

The working papers published in the Series constitute work in progress circulated to stimulate discussion and critical comments. Views expressed represent exclusively the authors’ own opinions and do not necessarily reflect those of the editor.
How Competitiveness May Cause a Gender Wage Gap: Experimental Evidence*

Matthias Heinz†, Hans-Theo Normann‡ and Holger A. Rau§

1University of Cologne
2Duesseldorf Institute for Competition Economics (DICE)
3University of Göttingen

February 2016

Abstract

We show that choices in competitive behavior may entail a gender wage gap. In our experiments, employees first choose a remuneration scheme (competitive tournament vs. piece rate) and then conduct a real-effort task. Employers know the pie size the employee has generated, the remuneration scheme chosen, and the employee’s gender. Employers then decide how the pie will be split, as in a dictator game. Whereas employers do not discriminate by gender when tournaments are chosen, they take substantially and significantly more from female employees who choose piece-rate remuneration. A discriminatory wage gap occurs which cannot be attributed to employees’ performance.

JEL Classification: C91, J16, M52

Keywords: dictator game, discrimination, gender wage gap, laboratory experiment, real-effort task.

*We are grateful to Uri Gneezy, Muriel Niederle, Nikos Nikiforakis, Beatrice Pagel, Gerhard Riener, Eric Schniter and Matthias Sutter for helpful comments. We are indebted to audiences at ESA Zurich 2013, Bavarian Micro Day Nuremberg 2013, Royal Economic Society Manchester 2014, 7th M-BEES 2014 in Maastricht, NERD Workshop 2014 Nuremberg and seminar audiences at Chapman University and the University of Duesseldorf for helpful comments. Last but not least we owe a debt of gratitude to Lise Vesterlund for sharing the z-Tree code with us.

†Email: heinz@wiso.uni-koeln.de, telephone: +49 221 470 7263.
‡Email: normann@dice.hhu.de, telephone: +49 211 8135297.
§Corresponding author, email: holger.rau@uni-goettingen.de, telephone: +49 551 39 2281.
1 Introduction

Experimental research has repeatedly documented that women and men differ in their inclination to compete. Niederle and Vesterlund (2007) show that women tend to shy away from competition: they choose a competitive environment less often than men, even though they perform the underlying task equally well (see also Niederle et al., 2012). Dohmen and Falk (2011) find in a remuneration-choice experiment that women prefer fixed wages to a higher extent than men.\footnote{Gneezy et al. (2009) show in a field setting that women’s competitiveness strongly depends on the nature of the societies investigated [matrilineal vs. patriarchal]. See also Hoffman et al. (2011).} In a field experiment with children, Gneezy and Rustichini (2004) find that competition enhances the performance of males only (which is not the case in Niederle and Vesterlund, 2007). Altogether, the evidence suggests that women tend to avoid competition whereas men embrace it (see Croson and Gneezy, 2009, for a survey of this literature).

In light of these results, it has been suggested that differences in competitiveness may entail pay gaps (Sutter and Rützler, 2010; Dohmen and Falk, 2011; Flory et al., 2014). If men choose competitive (better paid) environments more often, a wage gap emerges even when women and men perform equally well. Shying away from competitive environments may lead to women missing opportunities for better paid jobs.\footnote{The issue that women sometimes perform less well in competitive environments can be related to the remuneration choices. Gneezy et al. (2003) show that women may be less effective than men in competitive environments even if they perform equally well in non-competitive settings.} Thus, the differences in the attitude toward competition may cause a gender wage gap.

At first sight, it may appear that such a wage gap due to differences in the attitude toward competition would not be discriminatory as it would be consistent with the \textit{equal pay for equal work} principle. Generally, various explanations have been proposed for the gender wage gap (Blau and Kahn, 1992; 1994; Fryer and Levitt, 2010): performance-based explanations refer to differences in, for example, education, skills, and occupational choices, but the wage gap may also result from discrimination when equally qualified and equally performing men and women are paid differently. It seems that a gender pay gap due to differences in the inclination to compete would fall into the group of non-discriminatory explanations as it is performance based.

In this paper we argue that gender differences in the attitude toward competition may bring along an additional discriminatory effect. In the aforementioned experiments, the payments for the different remuneration schemes are exogenously fixed and do not allow for discrimination by gender. Within the remuneration schemes, there is equal pay for equal work by design. In
settings relevant in the field, this may not always be the case as remuneration is often flexible (bonuses, negotiations, award of perks, etc.) and is determined by superiors with the discretion to decide. If so, the remuneration decisions itself and the interaction with gender may play a role. How do employers reward performance given the choice for either competitive or non-competitive remuneration by females and males? The novelty of our experiments is that we allow for an *adjusted* wage gap to emerge (whereas in existing experiments the adjusted wage gap is inevitably zero). When superiors use their discretion to decide on payments, discrimination beyond what can be justified based on performance may result.

The goal of our research can also be put this way. The previous experimental literature mostly focused on the supply side (Bohnet et al., 2012, is an exception, see below). These papers analyze whether men and women differ in their labor supply decisions by way of their attitudes toward competitive environments. We mainly look at the demand side—the employers—and at the interaction between the supply and the demand side. We investigate how employers react to tournament vs. piece-rate choices by women and men.

To investigate this research question, we use a laboratory experiment with more than 600 participants. As in previous experiments, employees’ choices and performances determine the pie size: they first choose a remuneration scheme (competitive tournament vs. piece rate) and then they conduct a real-effort task. This part of the experiment is similar to Niederle and Vesterlund (2007). We add to this a dictator-game stage (for example, Hoffman et al., 1996; Eckel and Grossman, 1998) where the pie is generated by the employees (recipients) and where employers (dictators) decide how much to take from the pie. When deciding, employers know the pie size, the remuneration choice, and the employee’s gender. While this design is perhaps extreme in that employers have 100 percent discretion, it is suitable for identifying discriminatory effects which previous experimental research could not detect. We also conduct a control treatment where the remuneration scheme is determined by a random computer move.

Our findings are as follows. When the employees decide to work under piece rate, employers

---

4 On gender effects regarding negotiations and wage expectations, see Bowles et al. (2008), Rigdon (2012), and Babcock andLaschever (2003) which are discussed below.

5 Other dictator games include, for instance, Cherry et al. (2002), and Cappelen et al. (2007). Erkal et al. (2011) analyze a multi-player, real-effort dictator game. Brock et al. (2013) and Cappelen et al. (2013) investigate dictator decisions involving risks.

6 Dictator game experiments with earned money (real-effort experiments) were suggested by Ruffle (1998) where the performance of recipients in a general-knowledge contest determined the total surplus. In Cherry et al. (2002) the real-effort task was done by the dictator. Oxoby and Spraggon (2008) analyze both dictator- and recipient-earned wealth, also using GMAT questions. Further, see Konow (2000) and Heinz et al. (2012) for real-effort dictator games.
take substantially and significantly more from female workers. By contrast, employers do not discriminate based on gender when tournaments are chosen. Neither do they discriminate against women when the payment scheme is imposed by a random computer move. We thus observe a discriminatory wage gap which cannot be attributed to performance.

2 Experimental Design

Our design combines the remuneration-choice experiments of Niederle and Vesterlund (2007) with a real-effort dictator game (Ruffle, 1998; Konow, 2000; Cherry et al., 2002). The first two stages are similar to “Task 3—Choice of Compensation Scheme” of Niederle and Vesterlund (2007) where subjects first choose a remuneration scheme and afterwards take part in a real-effort task. We add a third stage where a dictator game is played: dictators (employers) decide how to split the pie, and the recipients (employees) determine the pie size in the mathematical real-effort task.

The game is played only once, and the timing is as follows:

1. employees (recipients) choose between tournament and piece-rate remuneration,
2. employees conduct a mathematical real-effort task which determines the pie size,
3. employers (dictators) decide how much to take from the pie, after learning the following information: the employee’s gender, the size of the pie generated, the employee’s remuneration choice, and (if tournament was chosen) whether the employee was a winner of the tournament.

We now explain these stages in detail.

The remuneration choice

The participants acting as employees were told upfront that they would be partitioned into groups of six people and that all groups would consist of three female and three male participants. This was the only time that gender was mentioned in the experimental instructions.

Employees could choose between two remuneration schemes: tournament and piece rate. This part is similar to Niederle and Vesterlund’s (2007) “Task 3—Choice of Compensation Scheme.”

---

7 This stage is somewhat similar to Bosman and van Winden’s (2002) power-to-take game where two players each earn an income in an individual real-effort task. One player (the authority) decides how much she will take away from the responder. In a second stage, the responder can punish the authority by destroying her own income. The transfer from the responder to the take authority is based on the responder’s remaining income after the second stage.

8 One difference to Niederle and Vesterlund’s (2007) design is that their participants do the real-effort task under
If tournament is chosen, payments depend on own performance and on the performance of other participants. Employees who choose tournament are compared to the other participants in their group, regardless of whether the other participants chose piece rate or tournament. When a subject achieves a score which is among the two best results in the group of six, she earns 90 “Talers” for each correct answer. Those participants who opted for tournament but did not manage to achieve one of the two best scores only receive 30 Talers for each correct answer.

For employees who choose piece rate, the payment rate for each correct answer is constant. They are given 50 Talers for each question that is answered correctly. Note that, for a one-third chance of winning the tournament, the expected payoff is the same under both payment schemes. We applied an exchange rate where 1.5 Taler corresponded to 1 cent.

The real-effort task

The real-effort task we use is the same as in Niederle and Vesterlund (2007). Employees have to add up five two-digit numbers. They are allowed to use scratch paper but pocket calculators are prohibited. The numbers are randomly drawn by the computer. Employees have 15 minutes in total for the real-effort task and can proceed at their own pace.

The real-effort task is as presented in Table 1. Subjects have to fill the correct answer into the blank box. After submitting an answer, subjects are immediately informed whether it was correct. On the screen, there is an overview of the number of correct and wrong answers.

| 58 | 83 | 76 | 13 | 85 |

Table 1: Example of a problem in the real-effort task

The employer decision

Employers have complete discretion as to how to divide the pie generated by the employees (Ruffle, 1998). In the instructions, the employer decision reads as follows: “You decide on the allocation of the money between you and the assigned other participant.”

both the non-competitive piece-rate condition and the competitive tournament environment before deciding which payment scheme to take. In our setup, the remuneration choice is upfront; our subjects knew the task to be performed, but had no experience as to how well they would do. Another difference to Niederle-Vesterlund (2007) is that our subjects had 15 minutes rather than 5 to solve mathematical problems. Finally, and perhaps most importantly, there is no dictator-game stage in Niederle and Vesterlund (2007). Our subjects knew that the final payoffs would be determined by the employer.
Before deciding, employers are informed of the size of the pie, the employee’s remuneration choice and the employee’s gender. Employers are also informed of whether the employee is a winner of the tournament, provided tournament is chosen. Table 2 summarizes the feedback given to employers. Employees have to fill in this form themselves, but the entries are checked by the experimenters.

<table>
<thead>
<tr>
<th>Participant ID</th>
<th># 01</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td>female</td>
</tr>
<tr>
<td>Chosen remuneration scheme</td>
<td>tournament</td>
</tr>
<tr>
<td>Among the two best performers (if tournament)</td>
<td>yes</td>
</tr>
<tr>
<td>Earned sum of money</td>
<td>17.50€</td>
</tr>
</tbody>
</table>

Table 2: Example of the feedback given before employers decide on the final allocation

**Treatments**

We have three treatments, called *Choice*, *Feedback*, and *Random*. The main treatment, *Choice*, is as in the sections above. It includes the three aforementioned stages where, in stage 1, the remuneration scheme is chosen by the employees.

In *Feedback* we varied the information feedback. In addition to the items listed in Table 2, we gave information on the relative performance of the employee. For each possible outcome of the remuneration choice (piece rate, tournament winner, tournament loser), we told the employers the corresponding average earnings in *Choice*.9

In *Random* there is no choice in stage 1; instead, the remuneration scheme is given by a random computer draw. To ensure comparability of treatments, we explained both payment schemes to employers and employees before they were told the relevant scheme that had been allocated to them. Participants were told that a random draw would select whether an employee would work under piece-rate or tournament conditions. This random draw was designed such that it roughly matched the frequency of tournament choices we observed in the *Choice* and *Feedback* treatments (52%).

---

9 We told the employers that “a similar experiment has already been conducted some time ago. The earned sum of money of participants A who had chosen piece rate was 9.60 euros.” In cases where the employee chose tournament and won [lost] we wrote “tournament and who were [not] among the two best performers” instead of “piece rate,” and the amounts were 26.20 and [5.80] euros, respectively. Employees did not know that the employers would be informed of the average earnings when they chose the remuneration scheme and worked on the real-effort task.
which we rounded down to 50 percent, so precisely half of the employees worked under tournament conditions. The groups of six participants were randomly composed, so they potentially contained both piece-rate and tournament employees.

3 Insights from previous studies

A first finding in the literature is that employers reward good employee performance. This has been shown in dictator games where, as in our experiment, the recipients earn the surplus in real-effort tasks (Ruffle, 1998; Oxoby and Spraggon, 2008; Heinz et al., 2012). Depending on their performance, recipients either generate a small or a large pie. The surplus generated is hence a clear-cut measure of relative performance which the dictators may take into account. Typically, dictators give more to well-performing recipients; low-performing recipients receive less on average and are frequently given zero.\(^\text{10}\)

A second group of findings concerns the accountability (Könow, 2000) of decisions. Accountability means that people will be called to account only for discretionary decisions, but not for exogenous factors which they cannot reasonably influence.\(^\text{11}\) Könow (2000) analyzes real-effort dictator games and shows that dictators’ decision making will be affected only when the recipient is accountable for certain variables.

What do these insights imply for our setup? As far as performance is concerned, employers’ decisions should be particularly affected by it in Feedback. Note that we give feedback on relative performance only in Feedback, whereas performance is difficult to assess in Choice and Random. In all our treatments, employees’ absolute performance is observable for the employers only in terms of the money earned. From this, employers can also conclude the number of correctly solved questions. In Choice and Random, however, it is difficult to assess how good a single employee’s performance is. Due to this ambiguity, we expect that employers will not base their decisions on performance only but also on the information we provide. For example, employers may believe that employees work harder under tournament conditions and may thus reward employees better in that case.\(^\text{12}\) In Feedback, we provide an unambiguous feedback on relative recipient performance, so we expect employers to reward good performance, as in previous studies. Regarding the accountability

\(^\text{10}\) In Heinz et al. (2012) this result is driven by female dictators being more generous than male dictators.

\(^\text{11}\) Cappelen et al. (2007) suggest liberal egalitarianism as an equivalent concept: that people should only be held responsible for their choices (for example, effort). This is in contrast to libertarian notions which also hold people responsible for their ability and talent on the one hand, and to notions of strict distributional egalitarianism on the other.

\(^\text{12}\) In these treatments, we realistically capture situations in the field: If performance and success were always fully obvious to employers, performance evaluation and fair remuneration would perhaps be less of a problem in the field.
argument, participants are fully responsible for their decisions in Choice and Feedback, whereas they are not accountable for it in Random.

Putting these arguments together, we state our main hypothesis as follows. In Choice, employers will reward employees who choose tournament more generously than those who choose piece-rate remuneration. This will not occur in the other treatments. In Feedback, employers have direct evidence on the relative performance and, in Random, subjects cannot be held accountable for the remuneration choices.

The existing literature also suggests that employers may discriminate based on employee gender. One hypothesis is that an employer’s assessment of a performance is affected by a gender stereotype. It is well documented that a significant fraction of the population shares a stereotypical belief that men are superior at mathematical tasks.\textsuperscript{13} Bohnet et al. (2012) use Niederle and Vesterlund’s (2007) mathematical task and let employers decide whether to hire a job candidate. The information they give to employers is the candidate’s gender and his or her past performance in the task. For this mathematical task, Bohnet et al. (2012) find that employers are more likely to employ a male employee, suggesting that they believe the male candidate will perform better in a (payoff-relevant) second round. Importantly, female employees are discriminated only if they are evaluated separately; the discriminatory effect vanishes when the performance of one male and one female candidate are simultaneously presented to employers.\textsuperscript{14} This is consistent with our design: the separate evaluation could be seen as not having information on relative performance. In our experiment, biased employers might take the information “male employee” as an indicator of good performance and pay them better, accordingly.

Another hypothesis is based on findings that male employees expect and demand more in negotiations.\textsuperscript{15, 16} Bowles et al. (2008) report results from negotiations where men penalize female job candidates more than male candidates for initiating negotiations. Women are therefore less inclined to initiate negotiations with male evaluators. In a setting without information on employees’ gender, Rigdon (2012) finds that women request lower wages than men and, in turn, female negotiators receive less. She shows that this effect vanishes when female negotiators are given information in advance on the average requests of men. Babcock et al. (2003) report similar results. In our setting, when employers expect male employees to demand more, they might be inclined to give more to men in order to meet this expectation.

\textsuperscript{13} See, for instance, Fennema and Sherman (1977, 1978), Eccles et al. (1990), Swim (1994), Spencer et al. (1999).

\textsuperscript{14} Bohnet et al. (2012) also show that the stereotype effect does not occur when they employ a non-mathematical task.

\textsuperscript{15} We are grateful to Muriel Niederle for suggesting this hypothesis to us.

\textsuperscript{16} Eckel and Grossman (2001) also report this finding in the ultimatum game, that is, they find male second-movers to be more demanding than female responders.
4 Procedures

Before entering the laboratory, participants were randomly assigned to two groups of equal size and divided into two separate rooms (A and B). Upon arrival, a random draw for each individual participant determined the A vs. B allocation. Subjects in room A were the employees. Subjects in room B were the employers. We ensured that half of the participants in room A were female.

The employees in room A were informed that they had been partitioned into cohorts of six people and that all groups consisted of three female and three male participants. (The instructions said: “In your group there are exactly three women and three men.”) In the meantime, the employers in room B had to wait approximately 20 minutes. They knew the nature of the task the employees had to conduct, and they also knew about the remuneration choice. Employers were not allowed to talk during the waiting time. After the employees had completed the real-effort task, we informed them about how much money they had generated. Subjects who had chosen the tournament were also informed as to whether or not they had won the tournament.

Next, employees were given a form and asked to fill in their participation number, gender, the chosen payment scheme and, in the case of tournament remuneration, whether they were one of the two best participants in their group. Finally, they had to fill in the amount of money they had generated. The experimenters checked that all forms had been filled in correctly. Each sheet was collected and put in an envelope. The envelopes were brought to the employers who each received one envelope. The allocation was random and the randomization was publicly visible. Employers had to decide in a dictator-like fashion on the allocation of money. Finally, we informed the participants in room A of the allocation decision of their employer. Subjects in rooms A and B were paid privately. All participants read the same instructions. They were fully informed about all details of the procedures of the experiment at the beginning. We applied neutrally framed treatments without using the words employer, employee, worker. Instead, the instructions said “participant A/B.”

The experiment was conducted in the FLEX laboratory at the Goethe-University Frankfurt. The subject pool consisted of 632 students from various fields. We conducted four sessions of each treatment, each session with 48 (one session: 36) subjects. We had 192 subjects participating in Choice, 192 in Random, 180 in Feedback. Three employee-employer pairs were discarded from the data set.\textsuperscript{17} After dropping these six subjects from the data, we were left with 558 participants.

\textsuperscript{17} One participant in the Choice treatment did not give the information about gender (this employee wrote “irrelevant” in the form) and we decided not to insist on the correction of the entry. (We know that this subject was female, however, we still cannot include the data because the employer did not know the employee’s gender when deciding.) One male employee in the Choice treatment appears to be an outlier: he solved 108 problems, 106 of them correctly. This is more than three times the average number of correct scores and still far above the second-highest
whose decisions we include in the analysis: there were 94 employee-employer pairs in Choice (47 female employees), 95 in Random (47 female employees), and 90 in Feedback (45 female employees). The remaining 68 subjects were employed in three additional sessions which were designed to elicit employee expectations. See section 6 below.

One session typically lasted about 75 minutes, subjects earned on average 17.26 euro including a 5 euro show-up fee. Subjects were recruited with ORSEE (Greiner, 2004). The experiment was conducted using z-Tree (Fischbacher, 2007) with an adapted version of the code used in Niederle and Vesterlund (2007).

5 Results

Throughout this section we employ the following variables. The pie size is the amount of money employees generate by solving the problems. The employee share measures the employees’ income relative to the pie generated. The employee share is a clear-cut measure for discrimination as it analyzes dictator giving based on the pie generated, that is, employees’ performance. The relative measure also enables us to compare employer behavior between low- and high-performing employees. To get a better idea on the final earnings of employees we also focus on the absolute amount of money employers give to the employees. We denote this by employee income. This is the employers’ one and only decision. We also discuss employees’ performance which measures the number of correctly solved problems. We report two-sided $p$ values throughout.

We will present the results in a sequence that reverses the order of decision making. We start with the employers’ decisions and move backwards to employees’ performance and remuneration decisions.

5.1 Employee share

Table 3 summarizes the employer decisions. Our research question is how the choice of tournament vs. piece-rate remuneration affects the share of the pie the employers give away. The main variable of interest is thus employee share. It is suitable to identify discrimination since it measures the employees’ income relative to performance.

scorer (71 questions, 63 correct). Furthermore, one male employer of the Random treatment indicated after the experiment that he had misunderstood the value given and requested permission to change his giving rate. Our results do not change qualitatively when we include these observations in the analysis.
Table 3: Descriptive statistics on employers’ decisions in our treatments. Employers decide on employee income, taken as given the pie size and the remuneration scheme. The employee share is the ratio of employee income and pie size. For each individual employee we calculated the employee share separately, and took the average of the shares. This average is reported in the table. Standard deviations in parentheses.

<table>
<thead>
<tr>
<th></th>
<th>piece rate</th>
<th></th>
<th>tournament</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>men</td>
<td>women</td>
<td>avg.</td>
<td>men</td>
</tr>
<tr>
<td><strong>Choice (n = 94)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>employee share</td>
<td>0.36</td>
<td>0.18</td>
<td>0.26</td>
<td>0.31</td>
</tr>
<tr>
<td></td>
<td>(0.22)</td>
<td>(0.21)</td>
<td>(0.23)</td>
<td>(0.23)</td>
</tr>
<tr>
<td>employee income (euro)</td>
<td>3.09</td>
<td>2.03</td>
<td>2.53</td>
<td>5.71</td>
</tr>
<tr>
<td></td>
<td>(2.31)</td>
<td>(2.50)</td>
<td>(2.45)</td>
<td>(6.70)</td>
</tr>
<tr>
<td>pie size (euro)</td>
<td>8.90</td>
<td>10.16</td>
<td>9.57</td>
<td>16.77</td>
</tr>
<tr>
<td></td>
<td>(2.90)</td>
<td>(2.83)</td>
<td>(2.90)</td>
<td>(12.58)</td>
</tr>
<tr>
<td><strong>Feedback (n = 90)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>employee share</td>
<td>0.22</td>
<td>0.32</td>
<td>0.29</td>
<td>0.35</td>
</tr>
<tr>
<td></td>
<td>(0.22)</td>
<td>(0.22)</td>
<td>(0.22)</td>
<td>(0.21)</td>
</tr>
<tr>
<td>employee income (euro)</td>
<td>2.65</td>
<td>3.47</td>
<td>3.27</td>
<td>6.04</td>
</tr>
<tr>
<td></td>
<td>(3.19)</td>
<td>(3.06)</td>
<td>(3.07)</td>
<td>(5.47)</td>
</tr>
<tr>
<td>pie size (euro)</td>
<td>10.52</td>
<td>10.54</td>
<td>10.54</td>
<td>16.38</td>
</tr>
<tr>
<td></td>
<td>(4.30)</td>
<td>(3.90)</td>
<td>(3.94)</td>
<td>(10.44)</td>
</tr>
<tr>
<td><strong>Random (n = 95)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>employee share</td>
<td>0.24</td>
<td>0.32</td>
<td>0.28</td>
<td>0.33</td>
</tr>
<tr>
<td></td>
<td>(0.25)</td>
<td>(0.22)</td>
<td>(0.24)</td>
<td>(0.19)</td>
</tr>
<tr>
<td>employee income (euro)</td>
<td>2.80</td>
<td>3.21</td>
<td>3.00</td>
<td>5.68</td>
</tr>
<tr>
<td></td>
<td>(3.22)</td>
<td>(2.38)</td>
<td>(2.81)</td>
<td>(5.78)</td>
</tr>
<tr>
<td>pie size (euro)</td>
<td>10.67</td>
<td>10.01</td>
<td>10.35</td>
<td>17.27</td>
</tr>
<tr>
<td></td>
<td>(3.34)</td>
<td>(2.91)</td>
<td>(3.12)</td>
<td>(12.56)</td>
</tr>
</tbody>
</table>

A conspicuous finding in Table 3 occurs in Choice when piece rate is chosen. Employers pay a much smaller fraction to female employees (18%) than to men (36%). Indeed, the employee share for women is only half of the share men receive. This difference is significant (Mann-Whitney test, \( p = 0.005 \)). Statistical significance prevails after we Bonferroni-correct the \( p \) level to \( 0.1/6 = 0.017 \) because of multiple comparisons.\(^{18}\) The employee share gap between men and women in Choice and under piece rate is also in great contrast to the other five gender differences in Table 3 which are all insignificant. Specifically, with the tournament payment scheme, no gender differences whatsoever occur: employers give 31 percent of the pie to male employees and 30 percent to female employees who opted for the tournament remuneration. We highlight this finding in Figure 1 which shows the employee share for women and men in Choice.

Splitting the tournament data further into winners and losers does not lead to unambiguous

\(^{18}\)With three treatments and two remuneration schemes, we have \( n = 6 \) male-female comparisons of employee share in Table 3.
insights, see Figure 1. Male winners of the tournament have an average employee share larger than losers, whereas for women it is the other way round. (See also Section 5.2 below.) These differences are not significant (Mann-Whitney tests, $p = 0.382$ (winners); $p = 0.554$ (losers)).

Looking at the cases where employers give nothing is also helpful. Figure 2 shows the relative frequency of employee share $= 0$ and employee share $> 0$ for the Choice treatment.\footnote{See Figures 4 and 5 in Appendix A for presentations of the Feedback and Random treatments.} Nothing is given most often when female employees choose piece rate (48%) whereas male tournament winners never receive a payment of zero. A Fisher exact test ($6 \times 2, p = 0.026$) suggests that the differences in Figure 2 are significant. Getting into pairwise male-female comparisons, we find
significant differences for piece rate \( (p = 0.021) \) but not for tournament losers \( (p = 0.640) \) or winners \( (p = 0.189) \).

Does employees’ performance matter for the employee share? Figure 3 presents scatter plots showing the relation between employee performance and employee share in our three treatments. In Choice (left panel) it can be seen that there exists no significant correlation between employees’ performance and the employee share. At the same time, the scatter plots show that employees obtain a higher employee share when performing better in Feedback and Random. This is further evidence that employers concentrate on factors other than performance when deciding on employees’ remuneration in the Choice treatment.

![Scatter plots showing the relation between employee performance and employee share](image)

Figure 3: The relation between employees’ performance and the employee share they obtained.

We now provide comprehensive regression analyses of employee share. The goal of the regressions is to give a full understanding whether the findings are driven by discrimination or by other factors such as performance. We use tobit regressions because the variables are left censored and, indeed, have many observations with a zero employee share.
<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>choice</td>
<td>-0.016</td>
<td>-0.002</td>
<td>0.026</td>
</tr>
<tr>
<td></td>
<td>(0.042)</td>
<td>(0.042)</td>
<td>(0.164)</td>
</tr>
<tr>
<td>random</td>
<td>0.020</td>
<td>0.028</td>
<td>-0.017</td>
</tr>
<tr>
<td></td>
<td>(0.041)</td>
<td>(0.041)</td>
<td>(0.161)</td>
</tr>
<tr>
<td>female employee</td>
<td>-0.041</td>
<td>-0.030</td>
<td>-0.245*</td>
</tr>
<tr>
<td></td>
<td>(0.034)</td>
<td>(0.034)</td>
<td>(0.129)</td>
</tr>
<tr>
<td>female employer</td>
<td>0.094***</td>
<td>0.087**</td>
<td>0.076**</td>
</tr>
<tr>
<td></td>
<td>(0.034)</td>
<td>(0.334)</td>
<td>(0.334)</td>
</tr>
<tr>
<td>piece rate</td>
<td>-0.009</td>
<td>-0.129</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.040)</td>
<td>(0.113)</td>
<td></td>
</tr>
<tr>
<td>tournament winner</td>
<td>0.022</td>
<td>0.006</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.054)</td>
<td>(0.084)</td>
<td></td>
</tr>
<tr>
<td>performance</td>
<td>0.003*</td>
<td>0.004</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.003)</td>
<td></td>
</tr>
<tr>
<td>piece rate × female employee</td>
<td></td>
<td></td>
<td>0.360**</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.166)</td>
</tr>
<tr>
<td>tournament winner × female employee</td>
<td></td>
<td></td>
<td>0.176</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.168)</td>
</tr>
<tr>
<td>[Interactions with choice]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>piece rate × choice</td>
<td>0.246*</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.147)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>tournament winner × choice</td>
<td></td>
<td></td>
<td>0.330*</td>
</tr>
<tr>
<td></td>
<td>(0.158)</td>
<td></td>
<td>(0.170)</td>
</tr>
<tr>
<td>piece rate × female employee × choice</td>
<td>-0.680***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.215)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>tournament winner × female employee × choice</td>
<td>-0.361</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.235)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>choice × performance</td>
<td>-0.003</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[Interactions with random]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>piece rate × random</td>
<td>-0.004</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.147)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>tournament winner × random</td>
<td></td>
<td>-0.101</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.155)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>female employee × random</td>
<td>0.166</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.161)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>piece rate × female employee × random</td>
<td>-0.153</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.208)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>tournament winner × female employee × random</td>
<td>0.115</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.246)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>random × performance</td>
<td>0.002</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>constant</td>
<td>0.227***</td>
<td>0.102</td>
<td>0.118</td>
</tr>
<tr>
<td></td>
<td>(0.040)</td>
<td>(0.072)</td>
<td>(0.109)</td>
</tr>
<tr>
<td>pseudo R²</td>
<td>0.049</td>
<td>0.086</td>
<td>0.197</td>
</tr>
<tr>
<td>observations</td>
<td>279</td>
<td>279</td>
<td>279</td>
</tr>
</tbody>
</table>

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 4: Tobit regression of the *share paid* to the employees by employers.
Consider the three regressions reported in Table 4 in turn, in which \textit{employee share} is the dependent variable. Regression (1) uses exogenous regressors only (that is, dummies for employer gender, employee gender and the \textit{Choice} and \textit{Random} treatment; the omitted treatment, captured by the constant, is \textit{Feedback}). We see that \textit{female employers} give significantly more. This finding also holds in all other regressions and will be discussed in Section 5.3. Neither our treatment variables nor the employee gender turn out to be significant in regression (1). Regression (2) adds the performance-related variables \textit{performance}, \textit{piece rate}, and \textit{tournament winner}. Again, we observe no significant regressors (except \textit{female employers}). Tournament winners do not receive a higher \textit{employee share}.

Finally, regression (3) adds all relevant interactions. Those with \textit{Choice} can be found in the middle of the table and those with \textit{Random} at the bottom. As in Table 3, the low employee share given to women in \textit{Choice} under piece-rate remuneration stands out in (3): \textit{piece rate} \times \textit{female employee} \times \textit{choice} is economically and statistically significant. This strong effect is only partially compensated by women’s significantly higher \textit{employee share} with piece-rate remuneration (\textit{piece rate} \times \textit{female employee}) as the sum of both effects is negative (Wald test, \(p = 0.004\)). Also significant (at the 10 percent level) is \textit{piece rate} \times \textit{choice}, but this effect applies to both sexes and thus does not change the gender gap. To sum up, the regression analysis confirms the descriptive analysis above: a significant gender gap in terms of \textit{employee share} occurs only in \textit{Choice} under piece-rate conditions. The gender gap is purely discriminatory and does not depend on employees’ performance. No other significant effects occur.

\section*{5.2 Employee income}

We now turn to the absolute income employees receive (\textit{employee income}). This is the actual decision the employers make. Corresponding to the low \textit{employee income} in Table 3 for women who choose piece rate in \textit{Choice}, women also get a 34 percent lower \textit{employee income} in that case (2.03 vs. 3.09 euros). This difference fails to be significant when we Bonferroni-correct for multiple (six) comparisons (Mann-Whitney, test \(p = 0.077\)). The other five Mann-Whitney tests comparing female-male \textit{employee income} are all \(p > 0.231\).
<table>
<thead>
<tr>
<th></th>
<th>employee income</th>
<th>normalized</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>choice</td>
<td>-1.171</td>
<td>-0.015</td>
</tr>
<tr>
<td></td>
<td>(0.835)</td>
<td>(0.038)</td>
</tr>
<tr>
<td>random</td>
<td>-0.478</td>
<td>0.006</td>
</tr>
<tr>
<td></td>
<td>(0.828)</td>
<td>(0.037)</td>
</tr>
<tr>
<td>female employer</td>
<td>-1.421**</td>
<td>-0.072</td>
</tr>
<tr>
<td></td>
<td>(0.674)</td>
<td>(0.048)</td>
</tr>
<tr>
<td>female employer</td>
<td>1.888***</td>
<td>0.028**</td>
</tr>
<tr>
<td></td>
<td>(0.683)</td>
<td>(0.042)</td>
</tr>
<tr>
<td>piece rate</td>
<td>1.294**</td>
<td>0.004</td>
</tr>
<tr>
<td></td>
<td>(0.571)</td>
<td>(0.042)</td>
</tr>
<tr>
<td>tournament winner</td>
<td>5.561***</td>
<td>0.165***</td>
</tr>
<tr>
<td></td>
<td>(0.766)</td>
<td>(0.033)</td>
</tr>
<tr>
<td>performance</td>
<td>0.171***</td>
<td>0.190***</td>
</tr>
<tr>
<td></td>
<td>(0.025)</td>
<td>(0.040)</td>
</tr>
<tr>
<td>piece rate × female employer</td>
<td>2.648</td>
<td>0.111*</td>
</tr>
<tr>
<td></td>
<td>(2.418)</td>
<td>(0.061)</td>
</tr>
<tr>
<td>tournament winner × female employer</td>
<td>-0.112</td>
<td>0.037</td>
</tr>
<tr>
<td></td>
<td>(2.418)</td>
<td>(0.061)</td>
</tr>
<tr>
<td>[Interactions with choice]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>piece rate × choice</td>
<td>2.080</td>
<td>0.075</td>
</tr>
<tr>
<td></td>
<td>(2.129)</td>
<td>(0.054)</td>
</tr>
<tr>
<td>tournament winner × choice</td>
<td>-1.375</td>
<td>-0.025</td>
</tr>
<tr>
<td></td>
<td>(2.213)</td>
<td>(0.051)</td>
</tr>
<tr>
<td>female employer × choice</td>
<td>0.711</td>
<td>0.057</td>
</tr>
<tr>
<td></td>
<td>(2.357)</td>
<td>(0.060)</td>
</tr>
<tr>
<td>piece rate × female employer × choice</td>
<td>-5.709*</td>
<td>-0.213***</td>
</tr>
<tr>
<td></td>
<td>(3.112)</td>
<td>(0.079)</td>
</tr>
<tr>
<td>tournament winner × female employer × choice</td>
<td>-3.627</td>
<td>-0.121</td>
</tr>
<tr>
<td></td>
<td>(3.367)</td>
<td>(0.085)</td>
</tr>
<tr>
<td>choice × performance</td>
<td>-0.064</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.062)</td>
<td></td>
</tr>
<tr>
<td>[Interactions with random]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>piece rate × random</td>
<td>0.000</td>
<td>0.005</td>
</tr>
<tr>
<td></td>
<td>(2.125)</td>
<td>(0.054)</td>
</tr>
<tr>
<td>tournament winner × random</td>
<td>-1.375</td>
<td>-0.025</td>
</tr>
<tr>
<td></td>
<td>(2.213)</td>
<td>(0.051)</td>
</tr>
<tr>
<td>female employer × random</td>
<td>0.711</td>
<td>0.057</td>
</tr>
<tr>
<td></td>
<td>(2.357)</td>
<td>(0.060)</td>
</tr>
<tr>
<td>piece rate × female employer × random</td>
<td>-0.581</td>
<td>-0.055</td>
</tr>
<tr>
<td></td>
<td>(3.015)</td>
<td>(0.077)</td>
</tr>
<tr>
<td>tournament winner × female employer × random</td>
<td>5.914*</td>
<td>0.093</td>
</tr>
<tr>
<td></td>
<td>(3.522)</td>
<td>(0.090)</td>
</tr>
<tr>
<td>random × performance</td>
<td>0.112</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.061)</td>
<td></td>
</tr>
<tr>
<td>constant</td>
<td>3.453***</td>
<td>-5.114***</td>
</tr>
<tr>
<td></td>
<td>(0.808)</td>
<td>(0.025)</td>
</tr>
<tr>
<td>pseudo R²</td>
<td>0.010</td>
<td>0.129</td>
</tr>
<tr>
<td></td>
<td>0.011</td>
<td>-0.720</td>
</tr>
<tr>
<td>observations</td>
<td>279</td>
<td>279</td>
</tr>
<tr>
<td></td>
<td>279</td>
<td>279</td>
</tr>
</tbody>
</table>

Standard errors in parentheses

*** p < 0.01, ** p < 0.05, * p < 0.1

Table 5: Tobit regression of the euros paid to the employees by employers. Regression (4) analyzes the euro payment for each correctly solved problem.
Table 5 reports regressions on employee income with the same regressors as in Table 4. In regression (1), female employees obtain an income that is approximately 1.4 euros lower than men. But this pay gap disappears when we control for performance in regressions (2) and (3). Female employers give more income to employees in all regressions. In (2) and (3), tournament winner and performance are highly significant. The variable performance strongly correlates with pie size and thus has a major impact on employee income. Regression (3) includes the interactions. Regression (4) uses a different dependent variable: we normalize the employee income by performance such that in (4) we analyze the payment for each correctly solved problem. This enables us to control whether male and female employees’ performance is acknowledged the same way. All regressors in (4) are the same as in regression (3) but we exclude performance.

In models (3) and (4) we confirm that female employees receive a significantly lower income than men when choosing the piece-rate remuneration. Whereas the effect becomes insignificant at conventional significance levels in model (3) (Wald test, $p = 0.111$), the effect remains significant in model (4) (Wald test, $p = 0.016$). Our findings can be summarized as follows:

**Result 1.** Comparing payments to women and men, we find that women earn less both in relative terms (employee share) and absolute terms (employee income) for piece-rate choices in treatment Choice. This effect is significant in a multiple regression analysis for employee share and is insignificant ($p = 0.111$) for employee income. The gender differences in the other treatments and remuneration schemes are insignificant.

Our findings on tournament winners and losers appear unexpected at first sight. Table 4 suggests that winners of the tournament do not receive a larger employee share than losers. But should the superior performance of the winners not be much better rewarded by employers? Looking at the absolute employee income solves the puzzle and, indeed, tournament winner is highly significant in regressions (2) and (3) of Table 5. Tournament winners generate a much larger pie (because they perform better and obtain the large payment of 90 Taler for each correctly solved problem): on average, winners generate a pie size of 26.5 euros. By contrast, the average pie size for tournament losers is 5.5 euros. Now consider the overall average employee share of 30 percent. This would amount to an employee income of roughly 8 euros for a tournament winner (one could consider this a reasonable sum for a 75-minute experiment) but only mere 1.65 euros (by most standards a petty amount) for a tournament loser. So, seen from the perspective of the absolute payments, winners and losers are treated rather differently but the employee share would

---

$^{20}$An unexpected finding is that tournament winner $\times$ female employee $\times$ random is significant (at the 30 percent level) in Model (3). However, the result should be treated with caution, as we only have four observations where women won the tournament in Random.
nevertheless be the same.

5.3 Employer gender

Tables 4 and 5 strongly suggest that female employers give more. Looking at simple non-conditional averages, we find that female employers give 32.7 percent of the pie whereas male employers give 25.9 percent. In terms of employee income, women give 1.26 euros more than men. Both differences are significant (Mann-Whitney test, $p = 0.012$) and confirm the findings in Heinz et al. (2012). We summarize:

**Result 2.** *Female employers give significantly more both in absolute terms (employee income) and relative terms (employee share).*

More pertinent to our research is the question to what extent the employer gender matters with respect to employee gender and remuneration scheme. The problem is that if we condition employer behavior on employee gender and payment scheme, we would need to split our data set into eight different subcases in each treatment. So we only have a few observations in each category and the findings are somewhat speculative in nature. Nevertheless, we report on the most conspicuous results.

Consider the cases where piece-rate remuneration is chosen in *Choice*. Both male and female employers give 14–15 percentage points more to male employees in this case, although women do so at a higher level (female employers allocate 43 percent to male workers and 29 percent to women; men give 27 percent to male employees and 12 percent to female workers). Worth pointing out is the rather low share of 12 percent that male employers pay to female employees. Instances with a zero giving rate are also telling: when men pay women under piece-rate, 11 out of 18 employers gave zero and one more employer gave next to zero. By contrast, male employers pay more (27 percent) to men. It turns out that only three out of 11 male employees receive zero from the male employers. It appears that male employers especially penalize the piece-rate choices of female employees. When men remunerate tournament choices, there is a sharp increase in the giving rate to women (35 percent).

5.4 Employee decisions

The employees make two decisions in *Choice* and *Feedback* (the remuneration choice and their effort decision which affects performance) whereas, in *Random*, they decide on effort only. The treatments *Choice* and *Feedback* are identical from the employees’ perspective, so we do not distinguish between the two in this section.
Table 6: Average number of correctly solved problems. Standard deviation in parentheses.

Table 6 presents the employee’s performance, that is, the average number of correctly solved problems. We focus on two main findings, both of which have been given attention in the previous literature. First, we find no significant gender differences in performance for both remuneration schemes ($t$ tests, all possible combinations in the treatments with and without choice: $p = 0.145$ or higher). This is in line with Niederle and Vesterlund (2007). Second, we find that employees have a substantial higher performance under tournament conditions compared to piece rate. This difference is significant, taking all treatments into account ($t$-test, $p < 0.0001$) and also when considering the treatments with and without choice separately ($t$-tests, all $p < 0.021$).

Table 7 presents an OLS regression analysis of employee performance. In regression (1) we include dummies for the treatments, piece rate, and the gender of the employee as independent variables. Regression (2) adds all interactions of female employee and the treatments and the compensation scheme. Except for piece rate, we observe no significant regressor. In particular, we do not find any significant performance differences between treatments or any gender effects. The negative sign of piece rate indicates that the performance of the employees is significantly higher under the tournament compared to piece-rate compensation, irrespective of the employee’s gender or the treatment. We summarize:

**Result 3.** (i) No significant gender differences can be found regarding employee performance under piece-rate and tournament remuneration in our treatments. (ii) Employees perform significantly better under tournament than under piece-rate conditions in our treatments.
<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>choice</td>
<td>-2.215</td>
<td>-1.159</td>
</tr>
<tr>
<td></td>
<td>(1.600)</td>
<td>(2.331)</td>
</tr>
<tr>
<td>random</td>
<td>-0.893</td>
<td>1.729</td>
</tr>
<tr>
<td></td>
<td>(1.591)</td>
<td>(2.315)</td>
</tr>
<tr>
<td>piece rate</td>
<td>-5.710***</td>
<td>-8.105***</td>
</tr>
<tr>
<td></td>
<td>(1.316)</td>
<td>(1.941)</td>
</tr>
<tr>
<td>female employee</td>
<td>-1.059</td>
<td>-0.062</td>
</tr>
<tr>
<td></td>
<td>(1.307)</td>
<td>(3.098)</td>
</tr>
<tr>
<td>female employee \times choice</td>
<td>-2.961</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(4.257)</td>
<td></td>
</tr>
<tr>
<td>female employee \times random</td>
<td>-5.533</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(4.132)</td>
<td></td>
</tr>
<tr>
<td>female employee \times piece</td>
<td>2.456</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(3.846)</td>
<td></td>
</tr>
<tr>
<td>female employee \times random \times piece</td>
<td>2.026</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(4.577)</td>
<td></td>
</tr>
<tr>
<td>female employee \times choice \times piece</td>
<td>2.781</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(4.601)</td>
<td></td>
</tr>
<tr>
<td>constant</td>
<td>37.810***</td>
<td>37.532***</td>
</tr>
<tr>
<td></td>
<td>(1.380)</td>
<td>(1.656)</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.082</td>
<td>0.097</td>
</tr>
<tr>
<td>observations</td>
<td>279</td>
<td>279</td>
</tr>
</tbody>
</table>

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 7: OLS regression on subjects' performance.

We conclude this section with a discussion of the remuneration decisions. Around half of the employees in these two treatments chose tournament (52.2 percent). Men did so more often than women: 64 percent of male participants opted for tournament as opposed to 40 percent of female participants. This difference is statistically significant ($2\times2$ Fisher’s exact test, $p = 0.002$).
Although we observe similar gender differences, the effect is somewhat less pronounced than in the previous literature. Niederle and Vesterlund (2007) find that 35 percent of their female participants but 73 percent of the males choose tournament. Several major differences of our experimental design may account for this. In addition to the fact that participants had no experience with the payment schemes in our case, it could be that some men are disinclined to compete in our setting where employers have full discretion over the final allocation of the money. Our findings can be summarized as follows:

**Result 4.** Female employees choose tournament less often than men.

### 6 Discussion

Our insights from previous studies (Section 3) suggest the following hypothesis. Employers will reward tournament choices more generously because they can expect employees to perform better in the tournaments. The reward will, however, materialize in *Choice* only because employers have direct evidence on relative performance in *Feedback* and, in *Random*, subjects cannot be held accountable for the remuneration choices. We further expect employer decisions to be affected by a gender stereotype which should be particularly forceful in the absence of performance information.

Supporting this hypothesis, we find that the payment scheme affects employer behavior in *Choice* but not in *Random* and *Feedback*. Given a lack of information on relative performance, it appears that employers see tournament choices as a sign of good performance. For piece rate choices, they are affected by a gender stereotype. Stereotypical thinking may induce employers to believe that male employees perform better in the math task: given tournament was not chosen, the feedback “male employee” may be perceived as good news by the employers and so men accordingly receive higher rewards. It appears that, given the lack of information on relative performance in *Choice*, there is more “wiggle room” to discriminate against women. In the end, men choosing piece rate receive an *employee share* even slightly higher than employees (male or female) who choose tournament. In *Feedback* this effect indeed disappears. One possible interpretation could be that providing employers with information on performance eliminates the misjudgment of female performance. Indeed, the strongest correlation of performance and employee share materializes in *Feedback*. The discriminatory effect also disappears in *Random* where no remuneration choice is made.

To get a better idea of subjects’ (stereotype-affected) beliefs about employers’ performance, we ran a “bystander” treatment. An additional 45 subjects who did not take part in the main experiments were asked to read the instructions of the *Choice* treatment and to report their expectations.
of employees’ performances under tournament and piece rate. Each subject was asked three times to guess the average performance of either female or male workers for the following three cases: piece rate, tournament winners, and tournament losers. The guesses were incentivized using a quadratic scoring rule. The data confirm the gender stereotype: male and female bystanders expect lower performances for female workers throughout.

While this gender bias regarding performance is intriguing, we concede that it cannot fully explain our results since the bias is present in all cases (piece rate, tournament winners, tournament losers). Our data, however, exhibit such a bias only in the case of piece-rate remuneration in Choice but not for Random.

A different explanation is based on employees’ expectations of the money they would receive. Women who choose the piece-rate payment scheme may be seen as expecting a lower giving rate. Put differently, employers might think that women choosing tournament will demand a higher share of the pie; as do men for either payment scheme (Rigdon, 2012). To elicit employees’ expectations, we asked a further 23 bystanders - who, as above, did not participate in the actual experiment - to read the instructions of the Choice treatment and report to us what they thought the employees’ expectations were. Specifically, we asked these subjects whether, given a choice of either piece rate or tournament, women or men expect more (or whether they expect just as much). People were paid an 8-euro flatrate for participation. Given a choice for piece rate, we find indeed that 11 of the 23 subjects suggest that men expect a higher giving rate; only four subjects assumed that women would be more demanding, and the rest thought both genders would be equally demanding. If employers are willing to meet these anticipated expectations, this would support the pattern we observe in the data; specifically, men obtain a higher share for piece rate in Choice. Having said that, the expectations data also suggest that men expect more when tournament is chosen (again, 11 subjects thought that men would expect a larger share of the pie). So men should also receive more when tournament is chosen - which is not the case. So this explanation cannot fully explain the results either.

21 We elicited these expectations after running our main treatments, so we did not ask the actual dictators themselves. However, using neutral bystanders for this task may even be advantageous since the actual dictators might exaggerate their estimates to justify their allocation decisions, perhaps mistakenly thinking the estimates would be communicated to the employees.

22 As an aside, we note that even if the “biased expectations” explanation had bite it would be entirely at odds with actual performances. The gender stereotype maintains that men perform better at the mathematical task, so ultimately it is the (allegedly) better performance that is rewarded, not male gender per se. In our data, men who choose piece-rate remuneration actually perform significantly worse than men under the tournament condition. And women outperform men (insignificantly) when working under the piece-rate condition. Hence, the feedback ‘male employee chooses piece rate’ should be perceived as bad news by employers, not as good news.

23 To avoid a demand effect, we also asked whether they think people younger/older than 21 would expect more.
We conclude as follows. The gender stereotype is consistent with our results only with the limitation that employers hold a mistaken belief about employee performance and misperceive the performance of men who choose piece rate. We find partial indication of that in our first bystander treatment. Bystanders also (wrongly) believe that female workers choosing tournament perform worse but, in the main experiment, women who choose tournament are not discriminated against. From our second bystander treatment, it also appears that giving rates are not exclusively expectation-based, either. Multiple and possibly contradictory forces may be at work here and further research is needed to come to a fully comprehensive understanding of the discriminatory effect we observe.

7 Conclusion

Our research is motivated by a simple but relevant question: if women and men differ in their inclination to compete, how do employers reward performances given the choice for either competitive or non-competitive remuneration? In our experiment, employees choose a payment scheme as in Niederle and Vesterlund (2007) but then employers, knowing the scheme chosen and the employee’s gender, decide with full discretion on how much to give to the employee (like in a dictator game).

Our findings suggest that gender differences in the attitude to compete (Niederle and Vesterlund, 2007; Gneezy and Rustichini, 2004; Gneezy et al., 2009; Dohmen and Falk, 2011; Niederle et al., 2012; Buser et al., 2014) may cause a discriminatory gender wage gap. Women who decide to work under the non-competitive piece-rate remuneration scheme receive significantly less than men who make the same choice. By contrast, female workers who choose the competitive tournament are not paid less. Gender differences in payments also vanish when the payment schemes are randomly determined by the computer or when the employers are provided with the information on employees’ relative performance. Women and men receive in all conditions a similar share of the pie, except when female employees chose the piece rate in our main treatment. Thus, it seems women are less rewarded for entering the tournament. Instead, women appear to be punished for choosing the non-competitive piece rate.

The wage gap we observe is discriminatory since it is adjusted for performance and payment scheme. In most countries such discrimination is illegal and there is equal pay for equal work by law. However, the adjusted wage gap in the field is, as a matter of fact, not zero in many countries (OECD, 2012). The implications of our experiments for the field have to be taken with a grain of salt: the remuneration scheme is often not a choice and information on relative performance may be available. Also, our experiments are one-shot whereas decision makers in the field may gain experience. Nevertheless, our experiment suggests one possible channel for why adjusted wage
gaps may occur. Whenever employers or superiors in general have discretion over wages, bonuses, promotions, perks, etc., negative discrimination of women may take effect.

Next to our main research question, we confirm several results along the way. First, women give more to employees than men in all our treatments. Heinz et al. (2012) found the same result in a dictator-game experiment with real effort. The authors find that female employers (37%) send significantly more than male employers (25%). Secondly, we find that tournament incentives work in that they lead to better performance but we find no gender differences in performance. Third, as in Niederle and Vesterlund (2007), women choose tournament less often.

References


Appendix A

Figure 4: Distributions of giving rates conditional on remuneration choices and gender in Feedback.
Appendix B

Translated instructions in English (Treatment: Choice)

—not intended for publication—

—note: differences in the Random Treatment are presented in square brackets—

Information about the Experiment

General Information:

Please read these instructions carefully. For taking part in the experiment, you will get a single flat-rate amount of 5€, which is not dependent on your decisions in the experiment. During the experiment you can earn an additional amount in “Taler”; how much depends on your decisions and those of the other participants. This sum will be converted into Euros at the end of the experiment, added to the 5€, and paid out anonymously, always provided you completed all parts of the experiment. The Taler-to-euro conversion rate is 1.5 Taler = 1 cent.

The experiment is also anonymous, which means you will not learn who you are interacting with. Please turn off your mobile phone and do note that you may not speak to other participants during the complete experiment. If you have any questions, please raise your hand and we will come to you.
At the beginning of the experiment, you received a participant ID. It consists of a letter and a number. All participants with an “A” are type A participants. All participants with a “B” are type B participants. The number is your participant number, which is required at the end in order to pay you.

Note that the participants of both types, A and B, have been given the same instruction sheet.

Experimental Procedure:

In the context of the experiment each participant of type A is assigned to a participant of type B. Each participant B must split a sum of money between him/herself and the assigned participant A. The amount of money that participant B can allocate depends on the results that participant A achieved in a test run previously.

Instructions for Participants of type A

In a moment, the participants of type B will receive an envelope. In this they will find a piece of paper with an amount of money written on it, together with further information (more about this below). The amount of money depends on the result that you achieved in the preceding test (more about this below). Your allocated participant of type B must divide this between you and him/her. However, you get the lump-sum payment of 5€, regardless of the decision of participant B. Please keep in mind that the sum of money that was assigned to you by participant B will be paid together with the lump-sum payment at the end of the experiment.

The test: How does it work?

- The test is performed on the computer and is to calculate the sum of five randomly selected two-digit numbers.
- In total, you have 15 minutes for the test.
- You may not use a calculator.
- You will get scratch paper and pencils, which you are allowed to use to do the calculations.
- Using the computer keyboard, type your answer directly into the box “Sum” and click the button “Enter” with the mouse to submit it.

Please refer to the following Screen Shot that shows the structure of the input screen, as well as an example of a possible task. The computer will tell you at once if your answer was correct. Here too, your answers are kept confidential.
The test: Determining the sum of money

- During the experiment, the participants of type A are split into groups of six.
- The 6 groups have exactly 3 females and 3 males.
- The test lasts a total of 15 minutes.
- The amount of money is determined in two ways called piece rate and tournament.

Before the test begins you chose the payment type you would like. To do this, click on the initial computer screen on either “Piece rate” or “Tournament”. Then press “OK” and the experiment will begin. The selected method of payment applies for the entire test. The payment types are as follows:

[Random treatment: before the test begins the payment type is randomly set by the computer. This payment type applies for the entire test.]

Payment type: Piece rate

The sum of money achieved is based on the number of correctly answered items as calculated at the end of the 15 minutes. Each correctly answered task earns 50 Talers. Note: Only the correct answers count. Wrong answers do not reduce the sum of money. Thus for Piece rate:

**Piece rate:** The remuneration is as follows

\[
\text{earned money} = \text{number of correctly solved questions} \times 50 \text{ Talers}
\]

Payment type: Tournament

The sum of money achieved is based on the number of correctly answered items as calculated at the end of the 15 minutes. In contrast to payment type “Piece-rate,” the payment here depends on the relative performance compared to the other five participants of the group. There are two possible cases:
Case 1: Occurs when you belong to the two participants in your group that have correctly answered the most tasks. In this case you are a winner of the tournament and get 90 Talers for each correctly answered task.

Case 2: Occurs when you don’t belong to the two participants in your group that have correctly answered the most tasks. In this case, you are not a winner of the tournament and you will get 30 Talers for every correctly answered task.

Thus, for Tournament:

**Tournament:** If you are a winner of the tournament (case 1), the remuneration is as follows

| Earned money: number of correctly solved questions x 90 Talers |

**Tournament:** If you are not a winner of the tournament (case 2), the remuneration is as follows

| Earned money: number of correctly solved questions x 30 Talers |

Note: Only the correct answers count. Wrong answers do not reduce the sum of money. If several participants correctly answered the same number of tasks, a random draw will be used, if needed.

After completion of the test, we will distribute two questionnaires, which you will please fill out.

Examples:

- “Piece rate” selected by you [Random treatment: the computer set “Piece-rate” for you] and you answered 30 questions correctly.
  The amount of money generated is then:
  Amount = 30 x 50 Taler = 1500 Taler
  In Euro = 1500 Taler /150 = 10 Euro

- “Tournament” selected by you [Random treatment: the computer set “Tournament” for you], you answer 40 questions correctly and are therefore one of the two best participants of your group. You are paid according to case 1.
  The generated amount of money is then:
  Amount = 40 x 90 Taler = 3600 Taler
  In Euro = 3600 Taler / 150 = 24 Euro
• “Tournament” selected by you [Random treatment: the computer set “Tournament” for you], you answer 30 questions correctly and therefore do not belong to the two best participants of your group. You are paid according to case 2. The generated amount of money is then:

\[ \text{Amount} = 30 \times 30 \text{ Taler} = 900 \text{ Taler} \]

In Euro: \( \frac{900}{150} = 6 \text{ Euro} \)

**Instructions for participants of type B**

For explanations of this test, please refer to the sections “The test: how does it work” and “test: determining the sum of money” on pages 2 + 3 of these instructions.

• Please ensure that your assigned participant of type A has completed all the necessary fields.

  Then enter your participant-ID in the field provided.

• Decide on the allocation of the total amount between you and your assigned participant of type A:

  Enter your chosen distribution on the sheet of paper provided for this.

Please keep in mind that you can choose all splits between 0 Taler and the total amount of money achieved in Taler (please enter only whole numbers, no decimal points). The following are a few examples of possible splits:

1. Example: Amount of money: 3,000 Taler

   I keep 3,000 Taler.
   I leave participant A 0 Taler

2. Example: Amount of money: 1,500 Taler

   I keep 750 Taler.
   I leave participant A 750 Taler

3. Example: Amount of money: 1,500 Taler

   I keep 100 Taler.
   I leave participant A 1,400 Taler

Please note that these amounts will be paid at the end of the experiment together with the lump-sum payment (5€).
Translated instructions in English (additional treatment 1: Employees’ performance)

—not intended for publication—

Note: In the performance treatment we first handed out the full instructions of Choice. We then distributed three additional questions where we asked the participants about employees’ performance. Below, we list these three questions. Note, there were two versions of treatments. In each version we either ask all three questions for male or for female employees. That is, why we present the questions in the form: “female/male”.

[. . .]

Below are examples of possible remuneration schemes employees could have opted for. Your task is now to assess average performance of the employees under the different conditions.

What do you think:

1. What is the average number of correctly solved questions by female/male participants who opted for piece rate?

2. What is the average number of correctly solved questions by female/male participants who opted for tournament and won it?

3. What is the average number of correctly solved questions by female/male participants who opted for tournament and lost it?
Translated instructions in English (additional treatment 2: *Expectations*)

—not intended for publication—

**Note:** In the *Expectations* treatment we first handed out the full instructions of *Choice*. We then distributed additional questions where we asked the participants about employees’ expectations. Below, we list one of these additional questions.

[...]

Below are examples of possible amounts that a participant of type A obtained in the previous test. Your task is now to assess the expectations of that participant.

What do you think:

*What did the participant of type A expect to get from the participant of type B?*

**Note:** that we are not asking which split you would have chosen personally or what split you would consider to be fair. It is simply about what you think the participants of type A expected. Reminder: Participants of type A had to solve the math problems, participants of type B determined the distribution of the money earned. Please read all four cases before you decide.

Assume the following: A participant of type A earned 15 Euros as a piece rate.

What do you think:

*Who expected more?*

- A male participant of type A, who chose the piece rate O
- A female participant of type A, who chose the piece rate O
- Both would expect the same amount O

[Note: We exactly asked the same question and gave the same three possible answers for the case where 15 Euros were earned under tournament. We further asked about the expectations of type A participants which are older / younger than 21 years, given a choice for piece rate and tournament.]
PREVIOUS DISCUSSION PAPERS


205  Dauth, Wolfgang, Findeisen, Sebastian and Suedekum, Jens, Adjusting to Globalization – Evidence from Worker-Establishment Matches in Germany, January 2016.

204  Banerjee, Debosree, Ibañez, Marcela, Riener, Gerhard and Wollni, Meike, Volunteering to Take on Power: Experimental Evidence from Matrilineal and Patriarchal Societies in India, November 2015.


201  Demeulemeester, Sarah and Hottenrott, Hanna, R&D Subsidies and Firms’ Cost of Debt, November 2015.


193 Defever, Fabrice, Fischer, Christian and Suedekum, Jens, Relational Contracts and Supplier Turnover in the Global Economy, August 2015.


191 Ciani, Andrea and Bartoli, Francesca, Export Quality Upgrading under Credit Constraints, July 2015.


184 Amess, Kevin, Stiebale, Joel and Wright, Mike, The Impact of Private Equity on Firms’ Innovation Activity, April 2015. Forthcoming in: European Economic Review.


181 Baumann, Florian and Friehe, Tim, Proof beyond a Reasonable Doubt: Laboratory Evidence, March 2015.


178 Buchwald, Achim and Hottenrott, Hanna, Women on the Board and Executive Duration – Evidence for European Listed Firms, February 2015.


174 Buchwald, Achim, Competition, Outside Directors and Executive Turnover: Implications for Corporate Governance in the EU, February 2015.

173 Buchwald, Achim and Thorwarth, Susanne, Outside Directors on the Board, Competition and Innovation, February 2015.


171 Haucap, Justus, Heimeshoff, Ulrich and Siekmann, Manuel, Price Dispersion and Station Heterogeneity on German Retail Gasoline Markets, January 2015.


169 Nowak, Verena, Organizational Decisions in Multistage Production Processes, December 2014.


164 Caprice, Stéphane, von Schlippenbach, Vanessa and Wey, Christian, Supplier Fixed Costs and Retail Market Monopolization, October 2014.
Klein, Gordon J. and Wendel, Julia, The Impact of Local Loop and Retail Unbundling Revisited, October 2014.


Behrens, Kristian, Mion, Giordano, Murata, Yasusada and Suedekum, Jens, Spatial Frictions, September 2014.


Baumann, Florian and Friese, Tim, On Discovery, Restricting Lawyers, and the Settlement Rate, August 2014.


Kholodilin, Konstantin A., Thomas, Tobias and Ulbricht, Dirk, Do Media Data Help to Predict German Industrial Production?, July 2014.


144 Jeitschko, Thomas D., Jung, Yeonjei and Kim, Jaesoo, Bundling and Joint Marketing by Rival Firms, May 2014.


142 Dauth, Wolfgang and Suedekum, Jens, Globalization and Local Profiles of Economic Growth and Industrial Change, April 2014.

141 Nowak, Verena, Schwarz, Christian and Suedekum, Jens, Asymmetric Spiders: Supplier Heterogeneity and the Organization of Firms, April 2014.

140 Hasnas, Irina, A Note on Consumer Flexibility, Data Quality and Collusion, April 2014.

139 Baye, Irina and Hasnas, Irina, Consumer Flexibility, Data Quality and Location Choice, April 2014.


Older discussion papers can be found online at:
http://ideas.repec.org/s/zbw/dicedp.html