



HAL
open science

Gender differences in tournament and flat-wage schemes: An experimental study

David Masclet, Emmanuel Peterle, Sophie Larribeau

► **To cite this version:**

David Masclet, Emmanuel Peterle, Sophie Larribeau. Gender differences in tournament and flat-wage schemes: An experimental study. *Journal of Economic Psychology*, 2015, 47, pp.103-115. 10.1016/j.joep.2015.01.003 . halshs-01105414

HAL Id: halshs-01105414

<https://shs.hal.science/halshs-01105414>

Submitted on 20 Jan 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Gender Differences in Tournament and Flat-wage schemes: An Experimental Study

David Masclet^a, Emmanuel Peterle^b and Sophie Larribeau^c

March 2014

Abstract

We present a new experiment that explores gender differences in both performance and compensation choices. While most of the previous studies have focused on tournament *vs.* piece-rate schemes, the originality of our study consists in examining the gender gap in the context of a flat wage scheme. Our data indicate that females exert a significantly higher effort than men in fixed payment schemes. We find however no gender difference in performance under the tournament scheme, due to a combination of two effects. On the one hand, men more significantly increase their effort when switching from a flat wage to a tournament scheme. On the other hand, when switching from the flat wage to a tournament scheme, women have less margin to increase performance since their effort was already relatively high with a flat wage. We also find that females are more likely than males to choose a flat-wage scheme than a tournament. This gap however narrows dramatically when feedback on previous experience is provided.

JEL codes: C91, J16, J31, M52.

Keywords: Experiment, Gender Differences, Tournament Scheme, Flat-wage Scheme.

We would like to thank Elven Priour for programming for the experiment. Financial support from the *Agence Nationale de la Recherche* (ANR-08-JCJC-0105-01, "CONFLICT" project) is gratefully acknowledged. We also thank David Dickinson, Enrique Fatas, Catherine Eckel, Ragan Petrie, Lise Vesterlund, Marie-Claire Villeval, seminar participants at the 2012 French Experimental Economics Association's Annual Meeting in Montpellier for their helpful comments. We are also grateful to participants at the 2012 ESA Meeting held in New York.

^a Department of Economics, University of Rennes 1, France; and CIRANO, Montreal Canada,
e-mail: david.mascllet@univ-rennes1.fr

^b Department of Economics, University of Rennes 1, France, e-mail: emmanuel.peterle@univ-rennes1.fr

^c Department of Economics, University of Rennes 1, France, e-mail: sophie.larribeau@univ-rennes1.fr

Gender Differences in Tournament and Flat-wage schemes: An Experimental Study

Abstract

We present a new experiment that explores gender differences in both performance and compensation choices. While most of the previous studies have focused on tournament vs. piece-rate schemes, the originality of our study consists in examining the gender gap in the context of a flat wage scheme. Our data indicate that females exert a significantly higher effort than men in fixed payment schemes. We find however no gender difference in performance under the tournament scheme, due to a combination of two effects. On the one hand, men more significantly increase their effort when switching from a flat wage to a tournament scheme. On the other hand, when switching from the flat wage to a tournament scheme, women have less margin to increase performance since their effort was already relatively high with a flat wage. We also find that females are more likely than males to choose a flat-wage scheme than a tournament. This gap however narrows dramatically when feedback on previous experience is provided.

JEL codes: C91, J16, J31, M52.

Keywords: Experiment, Gender Differences, Tournament Scheme, Flat-wage Scheme.

1. Introduction

Throughout the world, in a range of contexts involving businesses and governments, relatively few women hold top positions. Using data from a large panel of U.S. firms, Bertrand and Hallock (2001) reported that women held only 2.5% of the highest-paid executive posts. According to the Corporate Gender Gap Report (2010), published by the World Economic Forum, women occupied 14% of the seats on American Boards of Directors (Zahidi and Ibarra, 2010). Traditional explanations for the labor market gender gap focus on either differences in abilities (see, for instance, Gneezy *et al.*, 2003) or discrimination in the workplace (see Altonji and Blank, 1999 for an overview).

Recent developments in experimental economics have explored whether gender differences in individual or social preferences might also explain some of this gender gap (see Croson and Gneezy, 2009 for an overview). Special attention has been devoted to the role played by gender differences in attitudes toward competition through games in which participants are asked to choose between a tournament and a piece-rate scheme (Gneezy *et al.*, 2003; Gneezy and Rustichini, 2004; Vandegrift and Brown, 2005; Niederle and Vesterlund, 2007; Datta Gupta *et al.*, 2013). Most of these studies report that women shy away from competition while men are more eager to choose the tournament scheme¹. Other studies have investigated whether males and females perform differently in competitive environments. The evidence regarding gender differences in performance within competitive environments however is less clear-cut. Though some studies observe that males tend to outperform females under competitive schemes (e.g. Gneezy *et al.*, 2003; Gneezy and Rustichini, 2004; Shurchkov, 2012), others conclude no significant gender difference (Niederle and Vesterlund, 2007; Günther *et al.*, 2009). A possible reason to explain these opposite findings is that the nature of the task may matter, males being more motivated to outperform when the tasks are "stereotypical-male tasks (Günther *et al.*, 2009).

¹ Going one step further, some studies have investigated the determinants of gender differences in tournament entry, by focusing on the roles played by: risk aversion and self-confidence (e.g. Niederle and Vesterlund, 2007; Datta Gupta *et al.*, 2013), other-regarding preferences (e.g. Teyssier, 2008; Bartling *et al.*, 2009; Balafoutas *et al.*, 2012), personality traits (Müller and Schwierien, 2012), or pure distaste for a competitive environment (Niederle and Vesterlund, 2007). Another branch of the literature has sought to separate the respective roles of nature and nurture in order to explain these gender differences in attitude toward competition. This issue has led some economists to extend experiments in the field with younger participants (e.g. Dreber *et al.*, 2014; see also Niederle and Vesterlund, 2010 for a relevant discussion).

In this paper, we attempt to contribute to the existing literature by experimentally investigating gender differences in both performance and remuneration choice between a flat-wage scheme and a tournament. More specifically, our contribution to the literature is threefold.

A first contribution of our study is to investigate gender differences in performance under a fixed-payment setting. While most of the existing literature has focused on gender gap in tournaments or piece rates, less is known regarding gender differences under flat wages. To the best of our knowledge, no previous study has investigated gender differences in performance in the context of an experiment with a fixed-payment scheme. This relatively minor interest in flat wage is surprising since such schemes are prevalent across a number of firms and institutions (e.g. Bartling and von Siemens, 2010; Franceschelli *et al.*, 2010).²

Second, we revisit previous studies by investigating gender differences in performance under a tournament scheme. Indeed, as noted above, previous results concerning the existence of a gender gap in performance under a tournament scheme are mixed. It is thus important to replicate previous findings and in particular to test whether observed differences are robust to the nature of the task. This is done here by investigating whether there exist a gender gap in performance under a competitive environment in the context of a new original task that consists in decoding sets of numbers into letters from a grid of letters (Charness *et al.*, 2013).³

A third contribution of our paper is to explore gender differences in flat-wage entry with a tournament as an alternative. This sharply contrasts with the existing literature which has almost exclusively focused on piece-rate vs. tournament schemes. One notable exception is Dohmen and Falk (2010), who also investigated flat-wage entry with a tournament scheme as an alternative. Our paper differs however from theirs in several ways. First, while Dohmen and Falk (2010) concentrated on the issue of productivity sorting, here we abstract from this issue by imposing a remuneration scheme in some of our treatments. This allows us to examine gender differences in greater detail not only in terms of remuneration choice but also

² Several factors may explain the persistence of flat-wage schemes within firms, in particular in the public sector; these include: the role of egalitarian concerns (Ding *et al.*, 2001; Bartling and von Siemens, 2010), job security and intrinsic motivation for public service that may compensate for the quite small and relatively flat extrinsic incentives the sector offers (Georgellis *et al.*, 2011).

³ To some extent, the decoding task may appear more neutral than other tasks used in previous studies, in particular multiplying one-digit numbers by two-digit numbers. We acknowledge however that determining whether a particular task is more or less neutral is difficult, to some extent somewhat arbitrary and is still open to debate.

in terms of performance. Second, our design allows us to test whether feedback on performance affects gender gap with respect to flat-wage entry. This is done by asking players to choose between the two remuneration schemes, both before and after gaining experience with them.

Our experiment consists of three treatments. In the first treatment, called *exogenous tournament*, each participant is matched with a counterpart to form a pair and then asked to perform a decoding task in a tournament setting where the winner receives a monetary prize. The second treatment (*exogenous flat wage*) is similar to the first treatment, except for the fact that participants in each pair are paid the same fixed wage, irrespective of their performance. To study gender differences in selecting one's preferred compensation scheme, we also conduct a third treatment (*choice*), in which participants are asked to choose between a flat wage and a tournament, with the knowledge that their preferred scheme will be the one implemented. The comparison between these various treatments allows us to investigate the gender difference existing in both performance and choice of remuneration scheme.

To anticipate our findings, we observe that: *i*) Women exert significantly more effort than men under a flat-wage scheme; *ii*) Although men raise their level of effort when switching from a flat wage to a tournament to a far greater extent, we find no gender difference in performance under the tournament scheme; *iii*) Females are less likely to enter into competition; and *iv*) The gender gap is significantly reduced with performance feedback.

The remainder of our paper is organized as follows. Section 2 describes the experimental design and procedures. Section 3 then discusses the behavioral hypotheses proposed for evaluation. Results from the different treatments are reported in Section 4. Finally section 5 discusses our main findings and section 6 concludes the paper.

2. Experimental design

2.1 The game

Our experiment consists of three treatments that differ both in their compensation scheme (*flat wage vs. tournament*) and in the way this scheme is to be implemented (either *exogenously* or *endogenously*). The *exogenous tournament* treatment involves a simple real-effort task. At the beginning of the game, each participant is matched with a counterpart and given three minutes to perform a decoding task (Charness *et al.*, 2013). Participants are asked to decode sets of numbers into letters from a grid of letters displayed on their computer screen

(see the screenshot in Fig. A1 of online Appendix A). We chose a simple task that does not require any specific skills and is sufficiently fastidious to induce disutility for participants. Participants are continuously informed of their current number of correct answers. If a submitted answer is wrong, the participant has to decode the same letter until the correct answer is provided. Participants can solve as many problems as they like and are allowed to read any personal documents they have brought into the lab in their bags as long as they do not communicate with other participants. Once the three minutes have elapsed, the computer displays a summary screen of the two counterparts' respective performance. The player who decodes more tasks correctly is awarded 180 ECU⁴, while the counterpart receives 20 ECU. Ties are broken by a random draw. The *exogenous flat wage* treatment is similar to the previous treatment, except that each participant is paid exactly the same wage of 100 ECU, irrespective of her/his performance.

The third treatment, called *choice*, comprises two stages. During stage 1, participants are asked to choose between a flat wage and a tournament scheme, being informed that their preferred scheme will be the one implemented. The implemented schemes follow exactly the same rules as the treatments described above. In stage 2, participants perform the same task as in the previous games.

2.2 Experimental procedure

The experiment consists of 9 sessions with between 18 and 24 participants. These sessions are conducted via the ORSEE platform (Greiner, 2004) at the University of Rennes I (France). The experiment is computerized using the Ztree software (Fischbacher, 2007). A total of 202 participants (95 females and 107 males) are recruited from undergraduate classes in business, science, law and economics. None of the subjects has previously participated in a similar experiment. We gather information regarding their academic field of study and level of education⁵. On average, the sessions last about 60 minutes, which include the initial instruction period and payment of participants. Participants earn an average of €17.11 (S.D. = €4.74) with a show-up fee of 4 Euros.

Our design is a within-treatment design since participants successively play all treatments. More specifically, they first play the *choice* treatment, followed by the *exogenous flat wage*

⁴ The experimental currency, called ECU, was converted into Euros at a predetermined conversion rate of 40 ECU = 1 Euro.

⁵ The level of education is defined as the number of years spent studying after high school.

treatment, *exogenous tournament* treatment and lastly *choice* treatment.⁶ The *choice* treatment is implemented twice in order to check whether having previously experienced a flat wage and tournament influences participants' decisions during the second choice game.

Instructions for each treatment are only given to participants at the beginning of each game (see instructions in online Appendix B). Participants receive feedback about their payoff at the end of both exogenous treatments (second and third games), but no feedback is provided at the end of the choice treatment. In some sessions (sessions 7-9), measures of self-confidence relative to the issue of an exogenous tournament are elicited as follows: after performing the task under the exogenous tournament scheme, subjects are asked to predict how many participants in the room outperformed them during the tournament. This result provides an accurate measure of subjects' confidence regarding their relative abilities. If the answer is correct, participants receive 100 ECU; this payoff decreases to 50 ECU if their answer is within +/-1 of the correct number of participants, and to 0 ECU otherwise.

2.3 Preliminary measures

Before running the experiment, subjects participate in a preliminary stage to collect measures of their individual and social preferences. We elicit a measure of risk aversion based on Holt and Laury's (2002) lottery choice experiment. Our risk aversion indicator corresponds to the normalized number of "safe choices". Our inequality aversion measures are based on the procedure provided in Blanco *et al.* (2010). Precisely measures of disadvantageous inequality aversion correspond to the respondent's decision in an ultimatum game. The advantageous inequality aversion indicator is based on a modified dictator game. More detailed information on these games and the computational methods used is available in online Appendix A. All these preliminary measures are incentivized⁷. To avoid any wealth effect, participants are however not informed of the issue of these games until the end of the experiment.

3. Behavioral hypotheses

Let's begin by considering our conjecture regarding the level of effort exerted in each scheme. Regarding performance under a flat wage scheme, several previous studies have

⁶ We deliberately chose not to reverse the order of the games so as to measure the effect of switching from a flat-wage scheme to a tournament.

⁷ More specifically, in each preliminary game, one row of the payoff matrix is randomly selected and played to determine the payoffs for the selected option.

shown that people exert positive effort in such setting despite the absence of monetary incentive (see for instance Gächter and Thöni, 2009; Charness *et al.*, 2013; Kuhnen and Tymula, 2012). This surprising finding is often interpreted in the literature in terms of intrinsic motivation to work (Deci and Ryan, 1987; Gneezy and List, 2006; Cohn *et al.*, 2009; Gächter and Thöni, 2009; Charness *et al.*, 2013; Kuhnen and Tymula, 2012). Furthermore other studies have produced evidence that females may be more intrinsically motivated than males in several contexts, such as education and occupations (see, for instance, Vallerand and Bissonnette, 1992 and Lanfranchi and Narcy, 2013).⁸ Other authors have found that in absence of monetary compensation, females are significantly more likely to donate blood (52%) than males (30%) (see Mellström and Johannesson, 2008).⁹ Based on these evidences, we conjecture that females should outperform males in a fixed-payment setting by virtue of being more intrinsically motivated than males.

Regarding performance under a tournament scheme, previous evidence on the existence of a gender gap in performance in a competitive environment is not clear-cut. Some studies have found a significant gap in performance within competitive environments (e.g. Gneezy *et al.*, 2003; Gneezy and Rustichini, 2004; Shurchkov, 2012), while others have not observed any differences (Niederle and Vesterlund, 2007; Günther *et al.*, 2009)¹⁰. Some authors explain this discrepancy by the fact that males may be more motivated to perform better when the tasks are "stereotypical-male tasks", like those requiring spatial or math skills (Günther *et al.*, 2009). It would therefore seem that the nature of the task may indeed matter. In this current study we replicate previous studies using a different task, i.e. a decoding task. Based on this existing literature, we conjecture that if a gender gap in performance under a tournament, it should be relatively weak. Our conjectures regarding gender differences in performance under the tournament and the flat wage schemes are summarized in C1.

⁸ Using a database of 1,042 first-term, junior college students enrolled in a compulsory course, Vallerand and Bissonnette (1992) observe that women appear to be more intrinsically motivated to study than men. According to Lanfranchi and Narcy (2013), the higher representation of women in sectors characterized by lower extrinsic incentives, such as the public and nonprofit sectors, could be partly explained by the fact that men's and women's tastes differ in terms of intrinsic and extrinsic job characteristics. In the words of the authors, "the weaker pecuniary motivation of women in their job choices contributes slightly to the feminization of the public and nonprofit sectors."

⁹ Mellström and Johannesson (2008) also tested whether monetary compensation for donating blood might crowd out the supply of blood donors. The authors observed that for men the percentage of blood donors remained unchanged with and without monetary compensation while it dropped significantly for women when a payment was introduced.

¹⁰ Niederle and Vesterlund (2007) find no gender differences in performance under a tournament scheme. Günther *et al.* (2009) obtain results similar to those by Gneezy *et al.* for a "male" task but observe no gender differences for a neutral task.

C1. a) Men exert slightly greater effort than women under the tournament scheme. b) Women outperform men under the flat-wage scheme.

Our second conjecture relates to the decision to enter either the flat-wage or the tournament scheme. Previous studies have found that males are more likely to choose a tournament than a piece-rate scheme (Gneezy *et al.*, 2003; Gneezy and Rustichini, 2004; Vandegrift and Brown, 2005; Niederle and Vesterlund, 2007; Datta Gupta *et al.*, 2013). However the evidence regarding the choice between a flat wage and a tournament is relatively sparse. One notable exception is Dohmen and Falk (2010). The authors report on average gender differences as 53.6 (46.4) percent of all women prefer the flat wage scheme (tournament), compared to 30.2 (69.8) percent of all men. However this gender gap almost totally disappears after controlling for performance, risk attitudes, confidence and social preferences. As is clear from this study, the results are not clear cut, which calls for further investigation. In particular it is important to investigate which factors, if any, may matter in explaining a potential gender gap.

There are at least three main dimensions that distinguish the two compensation schemes. First, flat-wage schemes generate certain payments, whereas tournament payoffs are uncertain. The second dimension that distinguishes the two compensation schemes is that flat-wage schemes provide equal payoffs for all participants, which is obviously not the case in a competitive scheme. Third, tournaments may induce a situational pressure to perform compared to a fixed payment (Niederle and Vesterlund, 2010). We conjecture that these differences between the two remuneration schemes may lead females to prefer the flat-wage scheme to a tournament. Indeed, previous evidence has shown that females are more risk-averse and less overconfident than males (see Eckel and Grossman, 2008 for an overview of the literature, see also Niederle and Vesterlund, 2007; Kamas and Preston, 2012 and Datta Gupta *et al.*, 2013). Furthermore, although consensus is evident in the literature, several papers have shown that females are more inequality-averse than males (Fehr *et al.*, 2006; Kamas and Preston, 2012). Some studies have also shown that more egalitarian subjects are less likely to choose to compete than non-egalitarian subjects (see Teyssier, 2008; Bartling *et al.*, 2009; Balafoutas *et al.*, 2012). Finally some studies have argued that females may be more likely than males to dislike performing under pressure related to competitive environments (see for instance Niederle and Vesterlund, 2007). Based on all these findings, we conjecture that women may be more likely than men to choose a flat wage, as summarized in C2.

C2. Women are more likely to choose the flat-wage scheme than men.

Our third conjecture concerns the role of feedback in remuneration choices. We conjecture that the gender gap may be sensitive to previous experience and should thus be narrower in the second choice game compared to the first, as experience would reduce overconfidence by revealing to individuals their true abilities. Some studies have indeed shed light on the effects of experience on overconfidence. In particular, empirical evidence has indicated that overconfidence decreases with experience (e.g. Locke and Mann, 2001; Christoffersen and Sarkissian, 2002). This assumption is consistent with the findings of Wozniak *et al.* (2010) and Cotton *et al.* (2013), who conclude that feedback and repetition of competition decreases the gender gap in tournament entry. Consequently, we conjecture that:

C3. The gender gap is narrower in the second choice treatment.

4. Results

Let's first examine potential gender differences in performance under both flat-wage and tournament schemes. Afterwards, we investigate the choice of compensation scheme and the impact of feedback on tournament entry.

4.1 Gender gap in performance

4.1.1 Performance under a flat-wage scheme

Figure 1 displays the performance distribution in each exogenous compensation scheme separately for men and women. Consistent with previous findings, we observe a positive effort in the flat-wage scheme, despite the absence of a penalty for shirking. More surprising is the fact that women perform a significantly greater number of tasks (mean=40.20, SD=9.58) than men (mean=33.86, SD=15.78) under the exogenous flat-wage scheme. This difference is statistically significant ($z=2.688$, $p=0.0072$)¹¹. We also note that the proportion of participants decoding a very low number of tasks is higher among men. On average, 17.14% of men decoded 10 or fewer letters, while such is the case for only 2.11% of women.

These findings are consistent with our conjecture C1b. They are summarized in Result 1.

Result 1. *Women exert significantly greater effort than men under a flat-wage scheme.*

¹¹ Women also significantly outperform men with a flat wage when this compensation scheme is chosen. Such is the case in both the first choice game (although borderline significant: $z = 1.558$, $p = 0.1192$) and the second choice game ($z = 2.751$, $p = 0.0059$). We have no explanation for the fact that this difference is smaller under the flat-wage scheme of the first choice game.

Support for Result 1. The first panel of Table 1 shows estimates of the determinants of performance under the flat-wage scheme. Model (1) includes both flat-wage treatments while models (2) and (3) report similar findings for the exogenous and endogenous treatments separately.¹² In model (1), we appeal to panel data methods with individual random effects since each participant is observed several times. The independent variables include several demographics, such as gender and level of education. We also control for ability. Since our decoding task required no particular skill or qualification, except perhaps moderate computer familiarity¹³, our measure of ability is based on speed in using a computer.¹⁴ Moreover, we control for both risk aversion and inequality aversion. The measure of risk aversion is based on lottery choices (Holt and Laury, 2002). Inequality aversion indicators are built using decisions in the dictator game and respondents' decisions in the ultimatum game, as described in online appendix A. We control for inconsistent behavior as well, by including dummies for individuals making multiple switches during preliminary games.¹⁵ The "choice game" variable is a dummy that takes a "1" value when participants performed under an endogenously chosen flat wage (the choice game) and "0" when the flat wage was implemented exogenously. Finally, a trend variable is included to control for participants' learning or tiredness.

[Table 1 and Figure 1 about here]

The coefficient associated with the female variable is positive and highly significant in all three models of the left panel of Table 1. These findings indicate that women perform a significantly greater number of tasks than men in a flat-wage scheme. The "choice game" variable captures a negative and highly significant coefficient in column (1), thus suggesting the presence of a sorting effect, whereby those who chose a flat wage in the choice treatment were less likely to perform well. The "risk aversion" coefficient is positive and significant in models (1) and (2), which is surprising since participants were informed in the instructions

¹² We controlled for selection bias in the endogenous treatments using Heckman procedures for both remuneration schemes. In both cases, no selection bias appeared given that the inverse Mills ratios were not significant. These regressions, while not reported herein, are available upon request.

¹³ It may also require a working memory capacity able to significantly accelerate performance. We thank an anonymous reviewer for this remark.

¹⁴ This indicator of ability measures the time elapsed to complete the preliminary tasks; it corresponds to the inverse of the individual average time spent fulfilling these previous preliminary tasks. We acknowledge that this measure is imperfect. Alternatively, a similar decoding task performed with no payment or under a piece-rate scheme could have been used, but this might have led to possible learning effects and endogeneity problems. Note that this indicator is positively correlated with performance in both the flat-wage and tournament schemes.

¹⁵ Regressions on consistent players only (155 observations) lead to the same conclusions and are available upon request.

that no penalty would be imposed in case of shirking. One possible explanation for this finding is that highly risk-averse people may have decided to irrationally protect themselves against any eventuality by exerting a non-zero effort. Note however that we observed no gender differences in risk aversion ($z = -0.755$, $p\text{-value} = 0.4503$).

4.1.2 Tournament performance

Let's now consider the effort put forth in the tournament compensation scheme. Figure 1 shows a higher average performance under the tournament scheme (mean = 46.00, SD = 8.27) than in a flat wage (mean = 36.84, SD = 13.57). A Wilcoxon signed-rank test indicates that this difference is highly significant ($z = 10.816$, $p = 0.0000$)¹⁶. Introducing a tournament scheme also reduces performance variability and the frequency of minimal effort. Figure 1 also shows that if both men and women are positively influenced by a competitive environment, men react more strongly than women to the introduction of competition. On average, men increase their effort by 11.89 units from the exogenous flat wage to the exogenous tournament, whereas women increase their performance by only 6.09 units. A Wilcoxon signed-rank test confirms that this difference in extra effort is significant ($z = -2.916$, $p = 0.0035$). Our data also indicate no difference in performance between women (mean = 46.29, SD = 7.24) and men (mean = 45.75, SD = 9.12) in the exogenous tournament ($z = 0.449$, $p = 0.6534$)¹⁷. This outcome does not support our conjecture C1a yet is consistent with a number of previous studies that also found no difference between males and females (e.g. Niederle and Vesterlund, 2007). It seems that this finding results from two effects: *i*) the fact that men more significantly increase their effort when switching from the flat-wage to the tournament scheme given their increased sensitivity to competitive incentives compared to women (e.g. Gneezy *et al.*, 2003); and *ii*) the fact that women have less margin to increase their effort when switching from flat wage to tournament since women's efforts had already been higher. These findings are summarized in Result 2.

¹⁶ Similar findings are obtained when considering the choice games. On average, participants who chose the tournament during the first choice game performed 42.39 tasks (SD = 6.97), while those opting for the flat-wage scheme exerted less effort (mean = 33.68, SD = 10.62). This difference is statistically significant ($z = 6.358$, $p = 0.0000$). Similar findings were obtained with the second choice game (mean = 50.08, SD = 7.14 for tournament, mean = 33.14, SD = 16.15 for flat wage, with $z = 8.734$, $p = 0.0000$).

¹⁷ No difference has been found either between men and women when choosing the tournament scheme in both the first and second choice games ($z = 0.330$, $p = 0.7414$ with $z = -0.132$, $p = 0.8951$).

Result 2: *a) Both men and women significantly increase their effort when switching from a flat wage to a tournament. b) The extra effort for men is greater. c) No gender difference is observed under the tournament scheme.*

Support for Result 2. The right panel of Table 1 reports estimates of the determinants of performance under the tournament scheme. Model (4) includes all treatments while models (5) and (6) respectively consider the exogenous and endogenous treatments in isolation. The female coefficient is not significant in any of the models, thus confirming our previous findings. The "ability" variable captures a positive and significant coefficient in models (4) and (6), indicating that the most capable individuals are also more likely to exert effort under the tournament scheme. Note however that no significant difference in ability is observed between men and women ($z=-0.242$, $p=0.8085$). Moreover, the round variable captures a positive and significant coefficient, which may be interpreted as a learning effect.

4.2 Gender differences in the choice of remuneration scheme

This subsection focuses on remuneration choices. Figure 2 displays the frequency of choosing each compensation scheme during the two choice treatments. Consistent with our conjecture C2, Figure 2 indicates that despite producing identical performances in the tournament scheme, males are more (less) likely than females to choose tournaments (flat wage schemes). On average, 61.68% of men choose the tournament scheme vs. 33.68% of women in the first choice game ($\chi^2=15.793$, $p=0.0000$). These findings closely resemble those obtained by Dohmen and Falk (2010) (i.e. 62.3% vs. 37.3%). We also observe that those who won the exogenously imposed tournament are more likely to choose the tournament scheme during the second choice game as well (63.04% for women and 70.91% for men).

Figure 3 displays the proportion of women and men who enter the tournament conditional on their performance quartile in the exogenous tournament game for choice games 1 and 2, respectively. This figure suggests that performance in the exogenous treatment is a good predictor of a participant's tournament entry decision, which implies productivity sorting (Lazear, 2000; Dohmen and Falk, 2010). Interestingly, Figure 3 also indicates that the gender gap remains positive for almost all quartiles except the fourth quartile.¹⁸

¹⁸ Proportion tests indicate significant gender differences for all quartiles except the fourth (Q1: $z=2.1097$, $p=0.0349$; Q2: $z=3.4155$, $p=0.0006$; Q3: $z=1.6772$, $p=0.0935$; Q4: $z=-1.1685$, $p=0.2426$). This finding suggests that top performing women do not compete significantly less than top performing men in our experimental setting, which contrasts with Niederle and Vesterlund (2007). One possible reason may pertain to design differences, especially the fact that our experiment introduces flat wage as an alternative to the tournament compensation

Both Figures 2 and 3 show that the gender gap is strongly reduced by repetition; it is narrower in the second choice game than in the first and no longer significant ($\chi^2 = 2.345$, $p = 0.1260$), which supports our conjecture C3. Figure 3 points out that this reduction in gender gap is mainly due to a reduction in entry tournament of the lowest performing males, while for all quartiles females remain relatively stable in their choices. Our findings are summarized in Result 3.

[Figures 2, 3 and Table 2 about here]

Result 3. *a) Women are more likely than men to choose the flat-wage scheme. b) The gender gap is significantly reduced with experience. c) Although productivity is a good predictor of tournament entry, it fails to explain the gender gap.*

Support for Result 3. Table 2 consists of two panels. The left panel displays the results of several estimates of the probability of choosing the tournament scheme during the first choice game. The right panel reports similar findings for the second choice game. Most independent variables reported in Table 2 are similar to those of Table 1. In the spirit of Dreber *et al.* (2014), Niederle and Vesterlund (2007) and Dohmen and Falk (2010), we include an indicator of performance in the exogenous tournament to verify whether the decision to enter the tournament may be conditional on performance.¹⁹ Models (4) and (5) replicate models (2) and (3) for Sessions 7-9 only when measures of self-confidence (*expected rank*) were elicited after the exogenous tournament game.²⁰ Models (6) to (9) show similar estimates for the second choice game. These regressions contain the variable "won the previous tournament", which takes a value of "1" if the player had won the previous exogenous tournament game and "0" otherwise. The "female" variable captures a negative and highly significant coefficient in column (1), thus indicating that women are less likely than men to choose the tournament in the first choice game and therefore, conversely, more likely to opt for the flat wage scheme. After controlling for performance, risk aversion and social preferences, the gender gap remains significant (columns 2 to 5), suggesting that neither productivity nor risk aversion can

choice. It may be argued that women are more inclined to shy away from competition when the alternative scheme is a piece rate, since remuneration would still be directly affected by performance. This may be less prevalent when the alternative is a flat wage, in that selecting a fixed remuneration may be very costly for high-performing individuals.

¹⁹ Like Niederle and Vesterlund (2007), our measure of exogenous performance refers to both past and future performance. Given possible multi-colinearity, we excluded our previous "ability" variable from these estimates.

²⁰ The "expected rank" variable corresponds to the predicted number of participants in the room who had performed better than himself/herself plus 1. For instance, if a participant believed that 10 participants had performed better than himself/herself, then the expected rank would equal 11.

explain the gender gap in tournament entry.²¹ Consistent with Dohmen and Falk (2010), Table 2 indicates that performance in the exogenous tournament is a good predictor of a participant's tournament entry decision.²² The "risk aversion" coefficient is negative and significant in column (3), meaning that the more risk-averse participants are less willing to choose the tournament scheme.

The "expected rank" variable in column (5) captures a negative and significant coefficient, thereby indicating that less confident participants are less likely to choose the tournament. Though we observe no gender difference in risk aversion ($z = -0.755$, $p = 0.4503$), males and females differ significantly in their perception of risk (overconfidence). Men are in fact more likely to expect a positive outcome despite the absence of gender difference in tournament success. 51.43% of males underestimated the number of participants who outperformed them, while such was the case for only 39.13% of females ($z = -2.738$, $p = 0.0062$). The coefficients associated with inequality aversion measures in column (3) are not significant, suggesting that difference aversion does not seem to play a significant role in the decision to compete, which contrasts with some previous findings (Andreoni and Vesterlund, 2001; Fehr *et al.*, 2006; Kamas and Preston, 2012). Moreover, we do not find evidence of gender difference in other measures of social preferences.²³ These results are not really surprising in that no clear-cut consensus regarding differences in social preferences can be drawn from the existing literature, hence suggesting a higher variability in women's social preferences than in men's (see Croson and Gneezy, 2009 for a discussion).

Models (6) to (9) report findings on the second choice game. Interestingly, column (6) shows that the gender gap is no longer significant in the second choice, which is consistent

²¹ This is similar to Niederle and Vesterlund (2007) but contrasts with Dohmen and Falk (2010) who found that the gender gap almost disappears after controlling for several individual and social attitudes. As noted by the authors themselves, one reason is that the gender difference may be camouflaged by differences in attributes such as risk aversion. Indeed, the authors found that males and females differed in risk aversion, which is not the case with our sample. Let's remark that debate is still ongoing in the literature on whether a gender difference in risk attitudes actually exists. While most studies observe that women are more averse to risk than men, others have not identified any significant differences, suggesting a high behavioral variability (see Croson and Gneezy, 2009; Eckel and Grossman, 2008 for a discussion).

²² This finding contrasts with Niederle and Vesterlund (2007), who found that performance does not affect tournament entry. A possible reason mentioned above for this difference may be related to differences in the design and, in particular the fact that the alternative option to the tournament is a flat wage and not a piece rate, which may lead the most performing participants to enter the tournament.

²³ On the respondent's side in the ultimatum game, our data indicate that women reject on average 4.57 offers, whereas men reject 5.07 offers. This difference is not significant ($z = -1.109$, $p = 0.2673$). Our data also yield no significant differences in the modified dictator game; more specifically, women choose on average 5.93 egalitarian distributions out of 21 decisions while men choose 5.66 distributions, with this difference also being not significant ($z = 0.247$, $p = 0.8052$).

with both conjecture C3 and previous experimental studies (Wozniak *et al.*, 2010; Cotton *et al.*, 2013). The "won the exogenous tournament" variable captures a positive and significant coefficient in models (6) to (8). The "expected rank" coefficient in column (9) is now only weakly significant, suggesting that feedback on the exogenous tournament issue has lead individuals to review their expectations.

5. Interpretations of our main findings and robustness checks

In this section, we consider a number of possible explanations for our main findings. The finding that females are less likely to choose the tournament does not seem to be explained in our study by differences in performance, nor in risk aversion nor in social preferences.²⁴ Confidence seems to be a better predictor of tournament entry. The fact that experience tends to significantly narrow the gender gap in tournament entry can be interpreted as the overconfidence of low-performing individuals tending to decline over time as experience and feedback reveal individuals' actual relative abilities (Locke and Mann, 2001; Christoffersen and Sarkissian, 2002). We conclude that this reduction in gender gap is mainly due to the lowest-performing men switching from tournament to flat wage while female choices remain on average unchanged.

Another important finding of this study is that women significantly outperform men under a flat-wage scheme. One reason often cited in the literature to explain the positive effort in flat wage is intrinsic motivation (Deci and Ryan, 1987; Gneezy and List, 2006; Cohn *et al.*, 2009; Gächter and Thöni, 2009; Charness *et al.*, 2013; Kuhnen and Tymula, 2012). It can thus be conjectured that gender differences in intrinsic motivation explain why women are more willing to perform in the absence of extrinsic motivations and associated pressures (Vallerand and Bissonnette, 1992; Lanfranchi and Narcy, 2013; Mellström and Johannesson, 2008). Let's note however that the notion of intrinsic motivation embraces various concepts in the economic literature including self-esteem, pride in one's work, enjoyment of the task, fairness considerations and a sense of duty to honor one's contractual obligations (Deci, 1975; Baron, 1988; Kreps, 1997; James, 2005; Ellingsen and Johannesson, 2008; Kuhnen and Tymula, 2012). Due to the nature of our decoding task, we do not expect self-esteem or pride in one's

²⁴ One may argue that females would be more willing to choose the flat-wage scheme due to their dislike of performing under pressure, in association with a competitive environment (Niederle and Vesterlund, 2007). Such a motivation however cannot adequately describe the fact that the gender gap in choosing a remuneration scheme drastically declines over time. Indeed there is no reason to expect that distaste for performing under pressure would decline over time.

work to matter in the decision to perform in a flat-wage setting. Enjoyment of the task and sense of duty (i.e. honoring one's contract) are more likely to play a role in our experiment.²⁵

To measure the respective roles played by each of these factors, we asked participants in a post-experimental questionnaire (see online Appendix C) to rank by order of importance the various motives that may have led them to expend effort under the exogenous flat-wage scheme (from 1, being the most important reason, to 5, the least important).²⁶ The first motive refers to sense of duty (*"I find it normal to expend effort when being remunerated"*), while the second motive concerns enjoyment of the task (*"I enjoyed the decoding task"*). We also included three alternative motives that could have played a role. The third one relates to boredom (*"I would rather work than getting bored"*). The underlying intuition is that some individuals may have simply chosen to perform rather than becoming bored by doing nothing given the absence of an interesting alternative leisure task.²⁷ We also checked whether participants may have been influenced by a pure demand effect (*"I wanted to please the experimentalist"*). Indeed, it can be reasonably argued that effort in the flat-wage scheme simply reflects the fact that subjects want to seem fair to the experimenters, who are perceived as the "authority" figure.²⁸ Lastly, we included a motive related to risk aversion (*"I was not sure of being remunerated"*). This inclusion was motivated by the unexpected finding of a positive and significant coefficient associated with the risk aversion variable in Table 1. A possible interpretation would be that the most risk-averse individuals performed more due to

²⁵ One might also argue that gender differences in inconsistency explain gender differences in effort, which however is not supported by our data; our findings report no gender difference in inconsistency in the lottery game ($\chi^2=0.089$, $p=0.7652$), ultimatum game ($\chi^2=0.3636$, $p=0.54600$) or dictator game ($\chi^2=0.2146$, $p=0.6433$) if it is reasonably assumed that inconsistent participants show a level of inconsistency across the various games.

²⁶ This questionnaire was available in only some of our experimental sessions (7-9). This sub-sample consists of 58 individuals (23 females, 35 males). The average number of decoded tasks for this sub-sample was 38 (s.d. 13.53) for women and 35.14 (s.d. 13.28) for men.

²⁷ Despite the fact that participants were allowed to read any personal documents they have brought into the lab in their bags, the absence of any real alternative to the task may have led some to perform the task. Ideally, allowing participants to leave the laboratory would have been a more suitable alternative to work (out-of-the-job leisure). This option was infeasible however in our experiment since leaving the laboratory was incompatible with successive games. For these reasons, this experiment offered an on-the-job leisure activity alternative (see Dickinson (1999) for an experimental study of on-the-job leisure). In the field, leisure at work takes multiple forms: surfing the net, long coffee breaks, office gossiping, etc.

²⁸ We took several precautions to avoid or at least limit such demand effect. First, we were careful to refuse our own students in the experiment; moreover, we used no frames in the instructions and minimized interactions between players and the experimenter. Second, a demand effect could not explain all the differences observed across treatments. Despite these precautions, we acknowledge that this form of authority relationship between participants and experimenter may still exist. If such is the case however, this may simply reflect a sense of duty and mirror the field setting in which this type of vertical relationship exists between employer and employee. This consideration does not diminish the external validity of our experiment (see Zizzo, 2010 for a discussion on this point).

the fear of not being remunerated despite the fact it was explicitly mentioned in the instructions that participants would be paid a flat wage irrespective of their performance.

Table C in online Appendix C reports the percentage of participants who assigned each motive first, along with the mean ranking assigned to each motive for all participants and then for females and males separately. Table C consists of two panels. The left panel reports results on all tasks, except zero effort, while the right panel restricts the analysis to those participants who performed at least 34 tasks (i.e. we excluded those who performed a very low number of tasks, corresponding to the first quartile between 0 and 34 tasks).²⁹

Column (1) of Table C indicates that sense of duty, enjoyment of the task and distaste of boredom are the main motives explaining positive effort in a fixed-payment setting. More precisely, our data show that motive 1 (*"I find it normal to expend effort when being remunerated"*) was ranked first by 41.38% of participants (24 out of 58), followed in turn by the second motive (*"I enjoyed the particular decoding task"*) and the third (*"I would rather work than getting bored"*), each of which were ranked first by 18.97% of participants (i.e. 11 of the 58). Wilcoxon signed-rank tests show that, although motive 1 is ranked first, no significant difference is found between ranks assigned to motives 1 and 2 ($z=-1.549$, $p=0.1214$). The difference between motives 1 and 3 however is significant at the 10% level ($z=-1.712$, $p=0.0869$). The fear of not being paid and demand effect both seem to play a very weak role in the decision to expend effort. These motives were each ranked first by only 10.34% of participants (6 of the 58). A signed-rank test highlights a significant difference between the rank assigned to motives 1 and 4 ($z=-4.320$, $p=0.000$) and between motives 1 and 5 ($z=-4.087$, $p=0.000$). The right panel of Table C, focusing on top-performers, reports very similar results.³⁰ Altogether these findings and those from table 1 indicate that women exert

²⁹ Note that caution is required when interpreting these data for several reasons. First, due to the absence of monetary incentives, responses may be considered, to some extent, as pure "cheap talk". Second, subjects might use the questionnaire to 'justify' their decisions *ex post*. Third, participants may have been motivated for several other reasons than the five motives listed above. Fourth, more importantly, responses are not conditioned to the intensity of effort expended since all participants (except those who performed zero tasks) were allowed to complete this questionnaire. In other words, someone performing six tasks was considered with equal weighting as someone who accomplished 30 tasks. For these reasons, we cannot really infer individual differences in term of intensity of intrinsic motivation but only some general tendencies about the main motivations for expending positive effort under a flat-wage scheme. Despite these limitations, such data may still prove to be useful in delving deeper into the motivations behind effort in a flat-wage environment (see Croson, 2005 for a discussion).

³⁰ Regarding gender differences, men and women do not differ significantly on any of these motives. This finding does not mean that females and males are equally intrinsically motivated but simply indicates that those exerting positive effort do so for the same reasons, irrespective of their gender. Indeed as mentioned above one must be careful with the interpretation of these findings since responses to the questionnaire are not conditioned

significantly more effort than men under a fixed wage and that such positive effort is primarily motivated by a sense of duty, enjoyment of the task and distaste of boredom.

To disentangle these aforementioned explanations we designed a new experiment in which participants were asked to perform the same decoding task under a flat wage scheme but with the opportunity to leave the experiment sooner if they decided to work less.³¹ Participants could solve as many tasks (between 0 and 100) as they wish without any time limit being explicitly informed.^{32,33} One may reasonably expect that less effort should be observed both for males and females since participants may have less “intrinsic” motives to perform in terms of relative enjoyment of the decoding task (*"I enjoyed the particular decoding task"*) or in terms of avoidance of boredom (*"I would rather work than getting bored"*).³⁴ We conducted four sessions with a total of 82 participants (36 females and 46 males) who were recruited from undergraduate classes in business, sciences, law and economics.

Figure D1 (in online appendix D) displays the performance for women and men respectively, as well as measures of central tendency. While performances of females do not significantly differ from males in average (females' effort: 28.89 and males' effort: 27.63; $z=0.740$; $P=0.459$), they unambiguously differ in distribution. Indeed women have a higher median performance (28 tasks against 16.5 tasks for males) and a higher fourth quartile (50 tasks against 40 tasks for males). These findings are supported by the estimations shown in table D1 in online Appendix D. The first column reports an OLS model of the determinants of effort while the three remaining columns report probit models on the probability of performing effort in the lowest quartile, above the median or in the highest quartile, respectively. Although less significant, these findings confirm our previous results on the existence of a gender gap under a flat wage. They also shed light on the fact that motive 1 (*"I*

to the intensity of effort and thus in term of intensity of intrinsic motivation. Yet we know from table 1, that women exert significantly higher effort than men.

³¹ Instructions of this experience are available upon request.

³² To avoid any contamination effect, (i.e. the fact that observing someone leaving the lab may incite individuals to also leave the lab), we asked participants to choose how many tasks they wanted to exert before performing them. Participants were informed that they would be paid only if all tasks declared were realized.

³³ We varied the amount of the incentive across two conditions. In the *low payment* condition, participants receive 5 euros, whereas they receive 10 euros in the *high payment* condition. Gender differences in provision effort did not vary with the condition implemented. Data are pooled here and this dimension is not investigated in this current paper.

³⁴ This seems to be indeed the case since motive 2 and 3 were ranked first by 11.69% and 12.69% of participants respectively instead of 18.97% in the first experience. In sharp contrast, motive 1 (*"I find it normal to expend effort when being remunerated"*) is ranked first by 50.65% of participants against 41.38% in the first experience. The relatively high percentages associated to motives 2 and 3 may simply reflect the fact that the trade-off between leisure out of the job and effort is imperfect in a lab for several reasons.

find it normal to expend effort when being remunerated") seems to be an important factor of intrinsic motivation.

6. Conclusion

Research in behavioral economics has recently addressed gender differences in attitudes toward competition in order to explain gender inequalities in the labor market. We contribute to this literature through an original experimental setting that differs from previous studies in several aspects. In particular, we have measured gender differences in both performance and remuneration preferences in the context of a flat-wage scheme vs. a tournament, whereas most of the existing literature is focused on tournament vs. piece-rate schemes.

Four key findings have emerged from our work. First, women exert significantly more effort than men under a flat wage despite the absence of any penalty for shirking. This gender gap continues to remain highly significant even after controlling for several individual and social preferences. As such, we believe that an interpretation in terms of intrinsic motivation is the most consistent with our experimental findings. Second, we identified no gender difference in performance under the tournament scheme, due to a combination of two effects. Consistent with previous findings, we observed that men more significantly increased their effort when switching from a flat wage to a tournament setting. On the other hand, women had less margin to increase their performance when switching from flat wage to tournament since their effort had already been relatively high under the flat-wage scheme. Third, we found that women were less likely to enter into competition than men. This gender gap is robust in controlling for performance as well as for several individual and social attitudes. This finding seems to be partly explained by men being more overconfident than women in their abilities. Fourth and last, experience tends to narrow the gender gap with respect to tournament entry; this is mainly due to initially overconfident, lower-performing males revising their beliefs in their own ability.

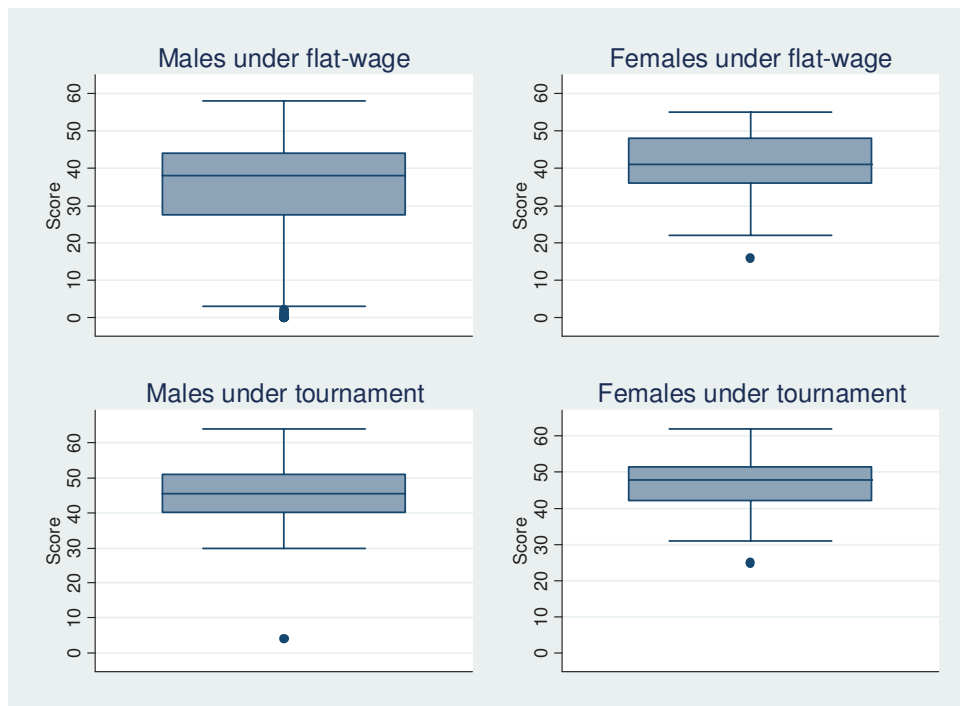
A natural extension of this study would consist in exploring in greater detail the gender differences in effort provided under a flat wage and, more specifically, to distinguish between the underlying motivations to perform in a fixed-payment setting. Although our experimental setting has offered some insight on this topic, a replication and development of this work would be required to fully understand all relevant mechanisms.

References

- Andreoni, J., Vesterlund, L., 2001. Which is the fair sex? Gender differences in altruism. *Quarterly Journal of Economics*, 116, 293-312.
- Altonji, J., Blank, R., 1999. Race and gender in the labor market. *Handbook of labor economics*, 3, 3143-3259.
- Balafoutas, L., Kerschbamer, R., Sutter, M., 2012. Distributional preferences and competitive behavior. *Journal of Economic Behavior and Organization*, 83, 125-135.
- Baron, J., 1988. The employment relation as a social relation. *Journal of the Japanese and International Economies*, 2, 492-525.
- Bartling B. and F. A. von Siemens, 2010. The intensity of incentives in firms and markets: Moral hazard with envious agents. *Labour Economics*, 17 598–607.
- Bartling, B., Fehr, E., Maréchal, M.A., Schunk, D., 2009. Egalitarianism and competitiveness. *American Economic Review*, 99, 93-98.
- Bertrand, M., Hallock, K., 2001. The gender gap in top corporate jobs. *Industrial and Labor Relations Review*, 55, 3-21.
- Blanco, M., Engelmann, D., Normann, H., 2010. A within-participant analysis of other-regarding preferences. *Games and Economic Behavior*, 72, 321-338.
- Charness, G., Masclet, D., Villeval, M.C., 2013. Competitive preferences and status as an incentive: experimental evidence. Forthcoming *Management Science*.
- Christoffersen, S., Sarkissian, S., 2002. Location overconfidence. Working Paper, McGill University.
- Cohn A., Fehr, E., Goette, L., 2009. Fairness and effort – evidence from a field experiment. Working paper.
- Cotton, C., McIntyre, F., Price, J., 2013. Gender Differences in Repeated Competition: Evidence from School Math Contests. *Journal of Economic Behavior and Organization*, 86, 2013.
- Croson R., 2005. The Method of Experimental Economics. *International Negotiation*, 10, 131–148
- Croson, R., Gneezy, U., 2009. Gender differences in preferences. *Journal of Economic Literature*, 47, 448-474.
- Datta Gupta, N., Poulsen, A., Villeval, M.C., 2013. Gender matching and competitiveness: Experimental evidence. *Economic Inquiry*, 51, 816-835.
- Deci, E., 1975. *Intrinsic Motivation*. New York, Plenum Press.
- Deci, E., Ryan, R., 1987. The support of autonomy and the control of behavior. *Journal of Personality and Social Psychology*, 53, 1024-1037.
- Dickinson, D., 1999. An experimental examination of labor supply and work intensities. *Journal of Labor Economics*, 17, 637-670.
- Ding, D. Z., L. Ge and M. Warner, 2001. A new form of Chinese human resource management? Personnel and labour-management relations in Chinese township And village enterprises: A case-study approach. *Industrial Relations Journal*, 32, 327-345.
- Dohmen, T., Falk, A., 2010. Performance pay and multidimensional sorting: productivity, preferences, and gender. *American Economic review*, 101, 556-590.
- Dreber, A., Von Essen, E., Ranehill, A., 2014. Gender and Competition in Adolescence: Tasks Matter. *Experimental Economics*, 17, 154-172.
- Eckel, C., Grossman, P., 2008. Men, women and risk aversion: experimental evidence. *Handbook of Experimental Economics Results*, Volume 1, ed. C. Plott and V. Smith, 1061–73. New York Elsevier.

- Ellingsen, T., Johannesson, M., 2008. Pride and prejudice: the human side of incentive theory. *American Economic Review*, 98, 990-1008.
- Fehr, E., Naef, M., Schmidt K., 2006. Inequality aversion, efficiency, and maximin preferences in simple distribution experiments: comment. *American Economic Review*, 96, 1912-1917
- Fischbacher, U., 2007. Z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, 10, 171-178.
- Franceschelli, S. Galiani, and E. Gulmez, 2010. Performance pay and productivity of low- and high-ability workers. *Labour Economics*, 17, 317-322. Cambridge University Press.
- Gächter, S., Thoenig, C., 2009. Social comparisons and performance: experimental evidence on the fair-wage effort hypothesis. *Journal of Economic Behavior and Organization*, 76, 531-543.
- Georgellis Y., Iossa, E., Tabvuma, V., 2011. Crowding Out Intrinsic Motivation in the Public Sector. *Journal of Public Administration Research and Theory*, 21, 473-493.
- Gneezy, U., Niederle, M., Rustichini, A., 2003. Performance in competitive environments: gender differences. *Quarterly Journal of Economics*, 118, 1049-74.
- Gneezy, U., Rustichini, A., 2004. Gender and competition at a young age. *American Economic Review*, 94, 377-81.
- Gneezy, U., List, J., 2006. Putting behavioral economics to work: testing for gift exchange in labor markets using field experiments. *Econometrica* 74, 1365-1384.
- Greiner, B., 2004. An online recruitment system for economic experiments. Forschung und Wissenschaftliches Rechnen, ed. by K. Kremer and V. Macho. GWDG Bericht 63. Gesellschaft für Wissenschaftliche Datenverarbeitung Göttingen: Datenverarbeitung, 79-93.
- Günther, C., Arslan Ekinici, N., Schwieren, C., Strobel, M., 2009. Women can't jump? An experiment on competitive attitudes and stereotype threat. *Journal of Economic Behavior and Organization*, 75, 395-401.
- Holt, C., Laury, S., 2002. Risk aversion and incentive effects. *American Economic Review*, 92, 1644-1655.
- James, H., 2005. Why did you do that? An economic examination of the effect of extrinsic compensation on intrinsic motivation and performance. *Journal of Economic Psychology*, 26, 549-566.
- Kamas, L., Preston, L., 2012. Are women really less willing to compete than men? Gender Stereotypes, Confidence, and Social Preferences. Working paper.
- Kreps, D., 1997. Intrinsic motivation and extrinsic incentives. *American Economic Review*, 87, 359-364.
- Kuhnen, C., Tymula, A., 2012. Feedback, self-esteem, and performance in organizations. *Management Science*, 58, 94-113.
- Lanfranchi J. and M. Narcy, 2013. "What are the factors that lead so many women to choose a job in public and nonprofit sectors? Evidence from a French national survey", working paper Centre d'Etudes de l'Emploi.
- Lazear, E. P. 2000. "Performance Pay and Productivity." *American Economic Review*, 90, 1346 - 61.
- Locke, P., Mann, S., 2001. House money and overconfidence on the trading floor. Working Paper, George Washington University.
- Lundeberg, M., Fox, P., Puncochar, J., 1994. Highly confident but wrong: gender differences and similarities in confidence judgments. *Journal of Educational Psychology*, 86, 114-121.
- Mellström C. and M. Johannesson, 2008. Crowding out in Blood Donation: Was Titmuss Right? *Journal of the European Economic Association*, Vol. 6, No. 4 pp. 845-863.

- Müller, J., & Schwieren, C., 2012. Can personality explain what is underlying women's unwillingness to compete? *Journal of Economic Psychology*, 33, 448-460.
- Niederle, M. Vesterlund, L., 2007. Do women shy away from competition? Do men compete too much? *Quarterly Journal of Economics*, 122, 1067–1101.
- Niederle, M, Vesterlund, L., 2010. Explaining the gender gap in math test scores: the role of competition. *Journal of Economic Perspectives*, 24(2), 129–44.
- Shurchkov, O., 2012. Under pressure: gender differences in output quality and quantity under competition and time constraints. *Journal of the European Economic Association*, 10(5), 1189-1213.
- Teyssier, S., 2008. Experimental evidence on inequity aversion and self-selection between incentive contracts. GATE Working paper.
- Vallerand, R., Bissonnette, R., 1992. Intrinsic, extrinsic, and amotivational styles as predictors of behavior: A Prospective Study. *Journal of Personality*, 60, 599-620.
- Vandegrift, D. Brown, P., 2005. Gender differences in the use of high-variance strategies in tournament competition. *Journal of Socio-Economics*, 34, 834–849.
- Wozniak, D., Harbaugh, W., Mayr, U., 2010. The menstrual cycle and performance feedback alter gender differences in competitive choices. University of Oregon Economics Department Working Papers 2010-2.
- Zahidi, S. Ibarra, H., 2010. The corporate gender gap report 2010. World Economic Forum. Geneva, Switzerland.
- Zizzo D., 2010. Experimenter demand effects in economic experiments. *Experimental Economics*, 13, 75–98.

Figure 1: Performance distribution under exogenous remuneration schemes (per gender)

Notes: The box in the middle of each boxplot describes central tendencies of the distribution. The solid lines inside the boxes are the median (Q_2). The bottom and top of each box are the first and third quartiles (Q_1 and Q_3), respectively. The upper adjacent value is the highest value below the upper inner fence ($Q_3 + 1.5 * (Q_3 - Q_1)$). The lower adjacent value is the smallest value above the lower inner fence ($Q_1 - 1.5 * (Q_3 - Q_1)$). Points correspond to outliers.

Figure 2: Frequency of compensation choices during the "choice" treatments (choice1 - choice2)

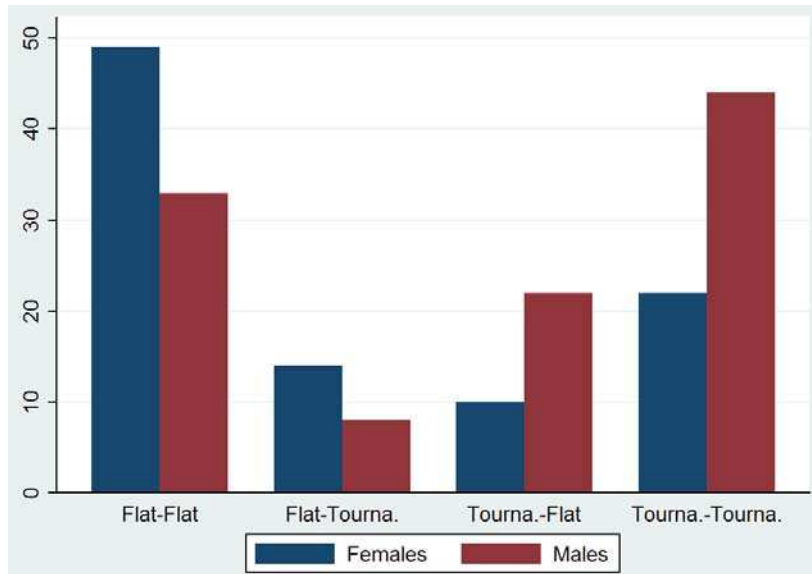
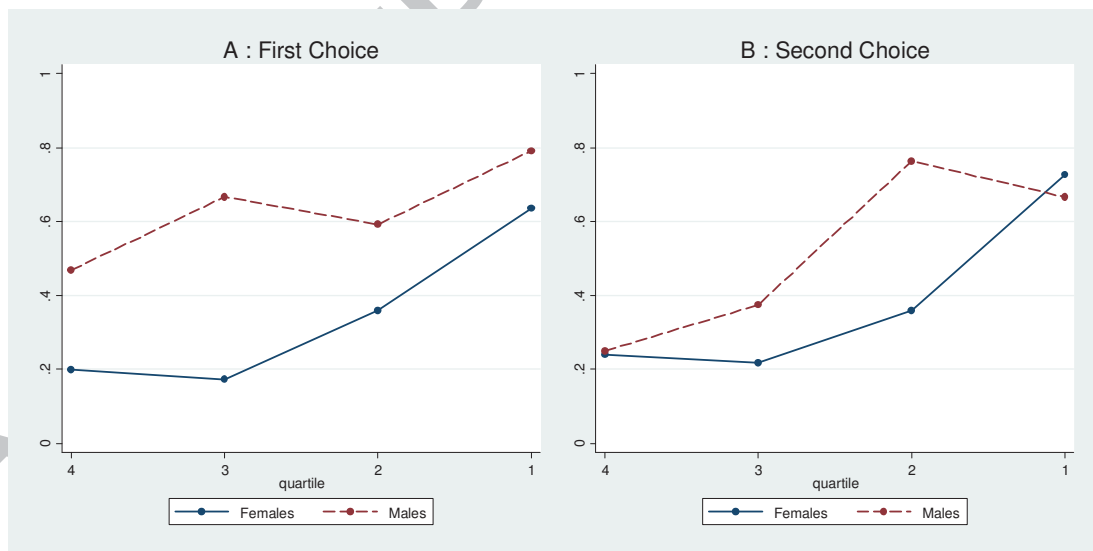


Figure 3: Frequency of tournament entry in choice games conditional on performance in the exogenous tournament game



Interpretation: 4 (worst quartile) to 1 (best quartile)

Tables

Table 1: Determinants of performance under both compensation schemes

Dep. var.	Number of tasks correctly decoded					
	Flat-wage All sessions			Tournament All sessions		
	(1) Pooled	(2) Exogenous	(3) Endogenous	(4) Pooled	(5) Exogenous	(6) Endogenous
Female	6.271*** (1.78)	6.369*** (1.87)	6.560*** (1.97)	0.329 (1.08)	0.617 (1.15)	0.082 (1.05)
Ability	1.131 (0.87)	1.337 (0.88)	-0.272 (0.97)	0.817* (0.47)	0.653 (0.51)	1.254*** (0.47)
Level of education	-0.062 (0.53)	-0.147 (0.56)	0.216 (0.57)	0.117 (0.37)	0.128 (0.39)	0.054 (0.41)
Risk aversion	8.914* (4.59)	10.993** (4.55)	1.979 (4.50)	2.021 (2.62)	0.680 (2.71)	5.634* (3.33)
Disadvantageous inequality aversion	0.222 (0.60)	0.405 (0.63)	0.024 (0.63)	-0.062 (0.35)	0.067 (0.37)	-0.026 (0.38)
Advantageous inequality aversion	4.126 (3.31)	4.079 (3.42)	3.479 (3.24)	-3.716* (2.06)	-3.454 (2.25)	-2.180 (1.89)
Choice game (Games 1 and 4)	-3.051*** (0.63)	-	-	-0.789* (0.41)	-	-
Round	0.179 (0.43)	-	-0.024 (0.61)	2.310*** (0.21)	-	2.409*** (0.33)
Constant	14.860 (10.67)	11.792 (10.67)	30.058** (11.84)	30.665*** (5.66)	39.830*** (6.03)	24.67*** (5.70)
Control for inconsistency	Yes	Yes	Yes	Yes	Yes	Yes
Observations	420	202	218	388	202	186

Note: *** Significant at the 0.01 level; ** at the 0.05 level; * at the 0.10 level. Robust standard errors in parentheses. Models (1) and (4) are GLS regressions with individual random effects. Models (2), (3), (5) and (6) are OLS regressions.

Table 2: Determinants of tournament scheme choice (probit models)

	Dep. var.		Prob. of choosing the tournament compensation scheme						
	First choice					Second choice			
	All sessions			Sessions 7-9 only		All sessions			Sess. 7-9 only
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Female	-0.280*** (0.068)	-0.305*** (0.07)	-0.322*** (0.07)	-0.376*** (0.13)	-0.414** (0.18)	-0.110 (0.075)	-0.120 (0.08)	-0.126 (0.08)	0.129 (0.15)
Perf. in exogenous tourn.		0.020*** (0.05)	0.021*** (0.05)	0.037*** (0.01)	0.045*** (0.02)		0.010* (0.01)	0.012* (0.01)	0.016* (0.01)
Level of education	0.003 (0.027)	0.001 (0.03)	-0.006 (0.03)	0.092 (0.10)	0.164 (0.11)	-0.006 (0.029)	-0.007 (0.83)	-0.011 (0.03)	-0.283* (0.14)
Risk aversion		-	-0.403** (0.21)	-	-0.037 (0.47)		-	-0.331 (0.22)	-0.225 (0.41)
Disadv. ineq. aversion		-	-0.037 (0.03)	-	-0.110 (0.07)		-	-0.016 (0.03)	0.022 (0.04)
Adv. ineq. aversion		-	0.082 (0.15)	-	0.235 (0.27)		-	0.002 (0.15)	-0.016 (0.23)
Expected Rank		-	-	-	-0.095*** (0.04)		-	-	-0.042* (0.02)
Won previous tournament		-	-	-	-	0.474*** (0.062)	0.408*** (0.06)	0.416** (0.08)	0.061 (0.15)
Control inconsist.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Log likelihood	-131.15	-123.03	-119.84	-30.77	-22.52	-112.14	-110.76	-109.19	-29.63
Observations	202	202	202	58	58	202	202	202	58

Note: *** Significant at the 0.01 level; ** at the 0.05 level; * at the 0.10 level. Robust standard errors in parentheses.

Highlights

- We present a new experiment that explores gender differences in both performance and compensation choices.
- Females exert a significantly higher effort than men in fixed payment schemes.
- No gender difference is found in performance under the tournament scheme
- Females are more likely than males to choose a flat-wage scheme than a tournament
- The gender gap narrows dramatically when feedback on previous experience is provided.

ACCEPTED MANUSCRIPT