

ARE WOMEN MORE ATTRACTED TO CO-OPERATION THAN MEN?*

Peter Kuhn and Marie Claire Villeval

We conduct a real-effort experiment where participants choose between individual compensation and team-based pay. In contrast to tournaments, we find that women choose team-based pay at least as frequently as men in all our conditions and significantly more often in a well-defined subset of conditions. Women's greater attraction to co-operative incentives results in part from their more optimistic assessments of their prospective teammate's ability. Women also respond differently to alternative rules for team formation in a manner that is consistent with advantageous inequity aversion. In contrast, men show a greater responsiveness to efficiency gains associated with team production.

A considerable body of recent research has shown that women tend to shy away from competitive work environments (Niederle and Vesterlund, 2007; Sutter and Rützler, 2010; Booth and Nolen, 2012; Villeval, 2012; Datta Gupta *et al.*, 2013; Garratt *et al.*, 2013) and tend to perform worse than men when placed in those environments (Gneezy *et al.*, 2003; Gneezy and Rustichini, 2004). Women's avoidance of competition is attributed to a combination of distastes for competition and lower levels of confidence in their relative abilities.¹ This avoidance of competition has been offered as an explanation of the continuing underrepresentation of women in top economic or political positions in modern societies.

If indeed women's talents are sometimes wasted because they avoid competitive environments, it seems important to know which types of work environments attract women and how women fare relative to men in those environments. In this article we study women's choices to enter a work environment characterised by team incentives, and the role of women's perceived relative abilities in those choices. In our real-effort laboratory experiment, participants can choose to receive either an individual piece rate or an equal share of a group's output, after experiencing each compensation

* Corresponding author: Peter Kuhn, Department of Economics, University of California, Santa Barbara, 2127 North Hall, Santa Barbara, CA 93106-9210, USA. Email: pjkuhn@econ.ucsb.edu

We are grateful to Philip Babcock, David Cooper, Uri Gneezy, Dorothea Kübler, Shelly Lundberg, Matthias Sutter, Lise Vesterlund and participants at the 6th European Workshop on Experimental and Behavioral Economics in Munich, the conference of the European Association of Labour Economists at Bonn, the European meeting of the Economic Science Association in Cologne, the Broom Center conference on Gender and Family in the New Millennium in Santa Barbara, the ASFEE conference in Montpellier and seminar participants at Oxford University, THEMA in Paris, WZB Berlin, UT Austin, UN Reno, Wharton, DIW Berlin and University of Hamburg for useful comments on an earlier version of this article. We thank Sylvain Ferriol for programming this experiment. This research programme has been supported by a grant from the French National Research Agency (ANR, EMCO programme, HEIDI grant) and was performed within the framework of the LABEX CORTEX (ANR-11-LABX-0042) of Université de Lyon, within the programme 'Investissements d'Avenir' (ANR-11-IDEX-007) operated by the French National Research Agency (ANR).

¹ Culture matters also in this respect. For example, Gneezy *et al.* (2009) show that this gender difference is reversed in matrilineal societies; see also Andersen *et al.* (2008). No gender difference in competitiveness between boys and girls is found in Sweden (Dreber *et al.*, 2012) nor in Colombia (Cárdenas *et al.*, 2012).

scheme successively. In many respects, the design of our experiment is similar to Niederle and Vesterlund's (2007) study of selection into competitive environments. The main contrast is that, instead of choosing between individual pay and a situation where workers' rewards depend negatively on their co-workers' performance, in our experiment the participants choose between individual pay and an environment where workers' pay depends positively on their co-workers' performance.

While it might be tempting to imagine that women are disproportionately attracted to co-operative environments because they tend to be more altruistic or egalitarian (Andreoni and Vesterlund, 2001), our results are more complex than this. Indeed, we do find that women are significantly more likely than men to select team-based compensation in our baseline condition, where team production offers no efficiency advantages (EA) over individual production and where the individual's choice does not affect the partner's earnings. Statistically, this gap can be explained by gender differences in confidence: essentially, the same confidence deficit that pushes women out of competitions pulls women into teams, where it is beneficial to have an abler teammate. Interestingly, we also find that compared with men in the baseline condition, men become much more likely to join teams and the gender gap vanishes when we introduce an instrumental reason for joining teams, in particular an efficiency advantage to team production.² We also find that women's relative propensities to join a team are strongly affected by the rules for team formation. Specifically, when teams are formed by mutual consent, women's team formation rate increases dramatically; we show that this phenomenon is consistent with women being more advantageous inequity averse than men.

In addition to studying selection into team pay, we also present results concerning the effects of the team environment on men's and women's task performance. Notably, in contrast to the tournaments literature, we find a zero causal effect of the team environment on women's absolute and relative task performance. At the same time, we find that self-selected teams perform worse than randomly assigned teams, and worse than subjects who choose to avoid teams. We also find that voluntarily formed female teams outperform self-selected male teams. All these gaps result from the facts that:

- (i) teams are adversely selected (abler subjects tend to avoid them) and
- (ii) adverse selection is stronger among men (because men have a more optimistic view of their own relative ability).

Compared with the literature on gender and tournaments, the economics literature on gender and teams is sparse.³ The typical approach is to take the team environment as given and ask how a team's gender mix or competition between teams affect performance (Bornstein *et al.*, 2002; Ivanova-Stenzel and Kübler, 2011; Apesteguia *et al.*, 2012; Delfgaauw *et al.*, 2013). When the choice of the payment

² This result is consistent with Cooper and Jabs Saral (2010) who find no gender difference in team formation by entrepreneurs in a design with a large (50%) EA to team production.

³ There is also a literature on public goods games that does not show systematic gender differences: see the surveys in Eckel and Grossman (2008) and Croson and Gneezy (2009). Although there is a large literature in psychology and management science on gender and team performance, most of it is based on observational studies in existing teams (not self-selection into teams), and teams are rarely incentivised: see Graves and Powell (2007) for a review.

scheme is endogenous, it has been found that making subjects compete as teams mitigates women's aversion to competition (Healy and Pate, 2011; Dargnies, 2012). However, the environment created by team competition is strategically much more complex than pure team-based pay and makes it more difficult to elicit the individuals' preferences for a pure team environment.⁴ Studies of the choice between pure team pay and alternative incentive schemes are extremely rare. With a team payment scheme either belonging to a menu of several possible incentives (Wozniak *et al.*, 2010) or opposed to a flat wage option (Dohmen and Falk, 2011), the existing research designs cannot elicit a representative subject's preferences between individual and team-based incentives.⁵ In contrast, the novelty of our approach is to elicit the individual's preferences between a situation where she is rewarded for her own efforts (i.e. an individual piece rate) and a situation where she is rewarded for the performance of her group (i.e. a team piece rate). We argue that choices between these two strategically simple options shed new light on gender differences in the tendency to gravitate into situations where co-operation rather than competition is rewarded, and on the role of underlying beliefs and preferences in accounting for those gender gaps.

In addition to addressing a gap in the literature on self-selection into pay schemes, our focus on selection into team pay is motivated in part by the increasing use of explicit team-based incentives in the workplace. For example, Lawler and Mohrman (2003) report that the share of Fortune 1000 companies using work group or team incentives for more than a fifth of their workers more than doubled, from 21% to 51%, between 1990 and 2002. Well known studies of this transition towards team pay at the firm and industry level include Boning *et al.* (2007) for US Steel mini-mills and Berg (1996) and Hamilton *et al.* (2003) for the apparel industry. Teams also play an increasingly dominant role in the production of scientific information (Wuchty *et al.*, 2007). More broadly, though, our objective is to understand gender differences in selection into environments where incentives for co-operation are implicit features of the employment contract, or are deeply ingrained in corporate cultures. We hope thereby to shed light on the causes of gender differences in occupational choices and wages, and on how changes in human resource management policy and corporate culture might make work environments more female friendly.

Two additional contributions of our study are the fact that we explore the effects of adding efficiency advantages to team production and the consequences of alternative team formation processes. Allowing for efficiency advantages is important, as otherwise

⁴ In the linear revenue-sharing scheme that characterises most team pay experiments, an individual's marginal financial return from producing an extra unit of output is just the price of a unit divided by the number of team members. Computing the marginal return in a tournament requires estimating the derivative of the probability of winning with respect to own output, which depends on the distribution of output levels and abilities of all the competitors. Combining tournaments and teams extends this complexity even further.

⁵ In Wozniak *et al.* (2010), women were insignificantly more likely to choose team pay than men but the design does not provide a ranking of preferences for team *versus* individual pay for the large share of subjects who chose a tournament. Dohmen and Falk (2011) study agents' choices between fixed and variable pay; in different treatments the variable pay option was either an individual piece rate, a tournament or a group piece rate. Men chose the variable pay option more frequently than women, especially in the case of tournaments.

it is difficult to understand why many firms use group-based compensation even when good measures of individual performance are available. Similarly, in addition to Niederle and Vesterlund's (2007) group formation procedure we also implement an approach requiring mutual consent for team formation.⁶ Mutual consent strikes us as more representative of real-world team and partnership formation processes, and has large effects on choices in our experiment.

The remainder of this article is organised as follows. Section 1 develops the experimental design and procedures. Sections 2–5 present the results. Section 6 discusses these results and concludes.

1. Experimental Design

At the beginning of each session, we elicit the participants' risk attitudes using the Holt and Laury (2002) procedure.⁷ Then, each participant enters his/her first name on the computer before being paired with another participant who is located in another room; in essentially all cases this revealed the participant's gender.⁸ Participants remain paired with the same co-participant for the entire session.

In a session, participants have to perform a task during sequences of 4 minutes. This task consists of decoding numbers into letters according to a code which changes repeatedly (see instructions in online Appendix C). Two features of this task made it well suited for the current experiment. First, the task is gender neutral: Unlike, for example, some sport activities, it is not commonly associated with any particular gender. Second, previous experiments with this task show no evidence of learning by doing after a first trial (Charness *et al.*, 2013). Although our main results are based on between-subject comparisons, this lack of learning simplifies the interpretation of the within-subject comparisons we make.

Before the experiment begins, participants are given 3 minutes to practise the task. At any time, they have the option to read magazines that are available in their cubicle or to surf the Internet instead of performing the task (this was made common information in the instructions but only one participant used this opportunity). The rest of the design is partly inspired by Niederle and Vesterlund (2007). Each session consists of six parts, always in the same order. One of these six parts is randomly selected for payment at the end of the session. Participants observe their own outputs in all parts but do not learn their co-participant's output in any part until the very end

⁶ Experimental approaches include matching each team chooser with the past output of a random 'mandated' team player (Niederle and Vesterlund, 2007), randomly pairing team choosers with various 'fixes' for solitary choosers and odd numbers of choosers (Dohmen and Falk, 2011; Dargnies, 2012) and randomly implementing the decision of one prospective team member (Healy and Pate, 2011).

⁷ The participants have to make 10 successive choices between two paired lotteries, 'Option A' and 'Option B'. The pay-offs for Option A are either €2 or €1.60, whereas the riskier Option B pays either €3.85 or €0.10. In the first decision, the probability of the high pay-off for both options is 1/10. In the second decision the probability increases to 2/10. Similarly, the chance of receiving the high pay-off for each decision increases as the number of the decision increases. A risk neutral participant should crossover from Option A to Option B at the fifth decision.

⁸ We designed the experiment to reveal the partner's gender (in an unobtrusive way) because we expected team formation decisions to depend on it. As we document below, however, we detected no effects of the partner's gender on either task performance or selection into compensation schemes.

of the session. Below, we describe the baseline (B) treatment. Aspects that were changed for our EA treatment are described after that.

1.1. *The Baseline Treatment*

Parts 1 and 2 of the experiment are designed to measure the gender gap in participants' task performance in the individual and team environments, respectively, in a situation where subjects are assigned to each pay scheme by the experimenter. Specifically, in Part 1 participants are paid a piece rate: each participant's pay for this part (if this part is selected for actual payment) is given by $Y_i^1 = r^1 Q_i^1$, where Q_i^1 is his own output. We set $r^1 = 20$ euro cents. In Part 2, participants are teamed with their co-participant to perform the task; they share the output of the team equally. In other words, individual i is paid $Y_i^T = r^T (Q_i^1 + Q_i^2)/2$ for her work during this part, where Q_i^2 is her co-participant's output. Throughout the baseline treatment, we set $r^T = r^1 = 20$ euro cents; thus, there is no efficiency advantages to team production. For any convex disutility-of-effort function, individually rational behaviour implies that participants should exert less effort in the team setting than the individual piece rate.

Part 3 is the first of two key elements in our experiment. Its goal is to study participants' revealed preference for teamwork in the simplest possible environment. To this end, participants choose between being paid an individual piece rate (as in Part 1) or according to a team-based payment scheme (as in Part 2). Then, they perform the task. If they have chosen teamwork, their performance in this part is added to the output of their co-participant in Part 2; this approach follows Niederle and Vesterlund (2007) and provides a guaranteed 'co-worker' for all participants who choose the team environment. Comprehension tests indicate that this procedure is well understood. Thus, Part 3 measures participants' responses to an important feature of team production: the fact that their pay will depend on their partner's performance while that partner is working under team incentives. At the same time, Part 3 does not tell us about participants' responses to another typical feature of team production: the fact that choosing to join a team could also affect other workers' monetary pay-offs. This is because, by construction, neither subjects' team formation nor their effort decisions in Part 3 have any effect on their partners' pay-offs.⁹

In Part 4, participants do not perform the task; instead they simply choose the payment scheme that will apply to their Part 1 performance: individual pay *versus* team pay based on their partner's Part 1 performance. Our motivation was to test for subjects' expectations of free riding by their partner: if they expected their partner to free ride when on a team, they should be more willing to choose team production based on their partner's Part 1 output (when he is paid individually) than on his Part 2 output.¹⁰ As expectations about relative ability are crucial in decision making, between Parts 4 and 5 we administer a short interim questionnaire. Participants are asked to

⁹ This follows directly from the fact that every subject is paired with their partner's previous-period actions, and applies also to Niederle and Vesterlund's experiment. In their context, it has the (perhaps surprising) implication that all four contestants can win the tournament if they enter: to win, each player simply has to beat the highest output achieved by others in their group in the previous period.

¹⁰ Of course, if preferences for the team environment also depend on whether the work is actually performed, Part 4 behaviour could differ from Part 3 for that reason as well.

estimate the number of problems they believe their co-participant solved correctly in Parts 1 and 2. They are rewarded 50 euro cents for each correct answer (plus or minus 1 unit).

Part 5 is the second key element of our experimental design. Its purpose is to study participants' team preferences in a richer context that more closely mimics real-world processes of team formation. In contrast to Part 3, participants in a team are paid based on their teammate's actual output in the current period; as this requires an active teammate, teams are formed only when both partners agree to form a team. If a team is formed, subjects are informed of this, then both participants work and are paid equally according to the team formula. If no team is formed, they are informed of that, then both partners work and are paid on an individual basis.¹¹ In addition to the factors affecting team choice explored in Part 3, subjects' team choices can now depend on their social preferences, as their actions now affect their partner's pay-offs.

Finally, Part 6 is the same as Part 5, except that after teams are formed (but before production occurs), participants who have agreed to form teams can exchange instant messages for 2 minutes. Its goal was to test whether the opportunity to socialise after the choice has been made affects team membership and performance.¹² That said, the interpretation of Part 6 is complicated by the fact that this is the only part of the experiment where participants have received information about any of their partner's decisions: specifically, at the start of Part 6, the participants who picked 'team' in Part 5 will know whether or not their partner picked 'team' in Part 5. Thus, participants could, in principle, use their Part 6 team-choice decisions either to reward or punish their partner's Part 5 team-choice decision.¹³ In fact, we found no significant differences between rates of team choice in Parts 5 and 6, although we place more weight on our Part 5 results due to this complication affecting Part 6.

Table 1 summarises the time structure of a session.

1.2. *The Efficiency Advantage Treatment*

This treatment is identical to the B treatment, with the exception that team production has a 10% productivity advantage over individual production. Specifically, the individual piece rate remains the same at $r^I = 20$ euro cents but the team piece rate is raised to $r^T = 22$ euro cents. The purpose of the EA treatment is to study selection into teams in a setting where technological factors favour production in groups (Lazear, 1999).

¹¹ If compensation is to be based on current performance of all team members, some team formation rule is required. While it might be interesting to explore the effects of other rules (such as a rule that makes teams the default unless both parties exit, or majority rule in the case of larger teams), the mutual consent rule seems the most realistic to us for the formation of two-person partnerships. Note that analogous rules are also required to form 'real' teams or tournaments, where pay-offs are based on the current performance of those who have agreed to join.

¹² To keep participants' choices confidential within rooms, we gave all participants an option to type text on their computer during this period. All participants had to wait till the communication period was ended before beginning Part 6 production: choosing individual compensation did not allow individuals to finish earlier.

¹³ Recall that partners' effort levels are not revealed till the very end of the experiment.

Table 1
Summary of the Experimental Design

Part	1	2	3	4	5	6
Is work performed?	Yes	Yes	Yes	No	Yes	Yes
Can participants choose the pay scheme?	No	No	Yes	Yes	Yes	Yes
Pay scheme	Individual	Team	Individual or team	Individual or team	Team if both agree, otherwise individual	Team if both agree, otherwise individual
Partner's performance in this part is used to compute team pay	–	Part 2	Part 2	Part 1	Part 5	Part 6
Communication	No	No	No	No	No	Yes

1.3. Procedures

The experiment consisted of 10 sessions conducted at the laboratory of the GATE (Groupe d'Analyse et de Théorie Economique) institute in Lyon. We invited undergraduate students from the local engineering and business schools via the ORSEE software (Greiner, 2004). Due to no shows, between 14 and 20 individuals actually took part in each session, for a total of 174 participants. The B (baseline) treatment was implemented in five sessions involving 86 participants and the EA treatment in five sessions with a total of 88 participants.¹⁴ We organised only mixed gender sessions. To guarantee a balance between genders, the number of participants of each gender in each session could not deviate by more than two from the other gender. In the B treatment, we have collected 16 observations of women paired with women, 14 observations of men paired with men and 56 observations of persons in mixed pairs. In the EA treatment, the corresponding values are 22, 24 and 42 respectively. We used our two contiguous laboratories. To preserve anonymity, upon arrival the first nine participants were assigned to a room and the next ones were directed to the other room and we proceeded to the necessary adjustments before distributing the instructions.¹⁵ As each participant was paired with someone from the other room, friends showing up at the same time could not be paired together.

The experiment was computerised, using the REGATE software (Zeiliger, 2000). The participants first participated in the Holt and Laury (2002) test to elicit their risk preferences, as the uncertainty regarding the potential partner's ability could affect the choice of the team payment scheme. Then, the instructions for the main task were

¹⁴ Due to a technical breakdown, we had to stop one B session after Part 4. In our results, we include data from Parts 1–4 of that session, which results in 10 more observations for Parts 1–4 than for Parts 5 and 6 of the B treatment. We have replicated all our results excluding all the data from this session; the results are virtually identical.

¹⁵ After 18 participants had shown up, we directed the 19th participant to the first room and the 20th to the second room. If fewer than 18 participants showed up, we moved participants from the first to the second room to make sure that we had the same number of participants in both rooms. Aside from asking participants to use their real first names, all of our instructions were carefully designed to draw as little attention to gender as possible.

distributed. They specified that there would be six parts and that one of these parts would be selected for payment at the end of the session but only the instructions for Part 1 were included. A quiz was used to check the understanding of the instructions and answers were checked individually. Participants practised for 3 minutes to familiarise themselves with the task. Then, they were required to type their first name and after being randomly paired with another participant, they were informed of the first name of this co-participant; they knew that they would be paired with the same participant throughout the session.¹⁶ The instructions for each new part were distributed after completion of the previous part. At the end of Part 6 and after completion of an exit questionnaire, the participants of the first laboratory were allowed to proceed to the payment room. Once these were paid, the participants from the other laboratory moved to the payment room.

On average a session lasted 75 minutes and participants earned €16.66 in the baseline and €17.23 in the EA treatment, including a €3 show-up fee and the payment of correct predictions.

2. Gender and Performance in Exogenously Assigned Teams

In this Section, we present results from Parts 1 and 2, where all participants were assigned to work first under an individual piece rate, then under team compensation.

RESULT 1.

- (i) *When individuals cannot choose their compensation scheme, there is no gender difference in performance in any scheme and in any treatment.*
- (ii) *Neither men nor women free ride.*

Support for Result 1.

Table 2 shows participants' mean output levels in Parts 1 and 2 by gender and treatment. It also displays the p-values from t-tests for differences between the means.¹⁷

Table 2 confirms that our task is gender neutral: there is no significant difference in output between men and women when they receive individual piece rates or team piece rates, irrespective of whether team production has EA over individual production. Taken together, these two features eliminate gender differences in workers' performance under the two incentive schemes as possible explanations of any gender gap in team choices.

¹⁶ We chose to pair players at the beginning of Part 1 so as to hold the environment constant throughout the session apart for the payment schemes. If we had paired subjects only at the beginning of Part 2, we would have introduced then two changes at the same time: the matching with a partner (including information on his/her gender) and the switch from the individual to the team piece rate. We acknowledge that we cannot exclude that some players may have increased their performance in Part 1 because they thought (wrongly) that we would compare the performance of the two paired players. However, nothing in the instructions of Part 1 indicated that there would be such a comparison. Note also that subjects clearly understood that their Part 1 pay depended only on their own performance and that no subject learned his/her partner's performance in any part till the very end of the experiment.

¹⁷ Kolmogorov–Smirnov tests of equality of distributions were also done, with similar outcomes. In all these tests, each participant is considered as one independent observation.

Table 2
Task Performance under Mandatory Individual and Team Pay

	(1) Part 1: individual pay	(2) Part 2: team pay	(3) p-value for difference between Parts 1 and 2
<i>(a) Baseline treatment</i>			
Women	56.93	57.95	0.065
Men	54.83	55.98	0.073
p-value for gender gap	0.093	0.127	
<i>(b) Efficiency advantage treatment</i>			
Women	54.88	55.84	0.089
Men	55.49	55.29	0.813
p-value for gender gap	0.624	0.685	

Notes. Sample sizes are 44 women and 42 men in the B treatment and 43 women and 45 men in the EA treatment. p-values are from two-sided t-tests for differences between means.

The other key finding from Table 2 is that, despite the anonymous nature of interactions and the fact that participants do not learn their partner's performance in any part until the conclusion of the experiment, participants do not free ride on their partners in teams. In fact, if anything the data show an (insignificant) increase in task performance between Parts 1 and 2.¹⁸ The most likely explanation of the lack of free riding is the nature of the task. Our results are consistent with a scenario in which effort costs are low and the time available to work in each part is short. In such situations, participants' individually rational decisions can be to simply exert the maximum possible effort for a wide range of marginal financial incentives.¹⁹ Fortuitously, focusing on a low moral hazard scenario such as ours not only simplifies the interpretation of the results but also enhances the applicability of our results to real workplaces. As a number of field studies have shown (Knez and Simester, 2001; Hamilton *et al.*, 2003; Boning *et al.*, 2007; Babcock *et al.*, 2011), free riding is absent in most workplace teams that have been studied.

Our next result refers to the players' beliefs and expectations.

RESULT 2.

- (i) *Neither men nor women expect free riding in teams.*
- (ii) *Women have much higher expectations regarding their partner's ability than men.*

¹⁸ This increase could reflect a small amount of residual task learning between Parts 1 and 2. We think this is unlikely, however, based on Charness *et al.*'s (2013) results with the same task. Without a practice period, Charness *et al.* observed learning only between the first and second 2-minute periods of their experiment. As we included a 3-minute practice period, their results would suggest zero learning between our Parts 1 and 2. Like us, Charness *et al.* found no gender difference in task performance.

¹⁹ Consistent with this interpretation, we find no significant performance differentials when the group piece rate is changed from 20 to 22 euro cents. Also, as we show below, participants' team choice responses to own and expected partner ability are strongly consistent with own pay-off-maximising behaviour; this argues against strong social preferences as an explanation of the lack of free riding. Indeed, note that under this interpretation even strong social preferences would not affect task performance levels, although they may certainly affect the choice of compensation scheme (individual *versus* team).

Table 3

Beliefs Regarding Partner's Task Performance under Mandatory Individual and Team Pay

	(1) Part 1: individual compensation	(2) Part 2: team compensation	(3) p-value for difference between Parts 1 and 2
<i>(a) Baseline treatment</i>			
Women	55.14	56.39	0.045
Men	50.26	52.10	0.002
p-value for gender gap	0.000	0.001	
<i>(b) Efficiency advantage treatment</i>			
Women	53.88	55.16	0.013
Men	51.47	52.78	0.002
p-value for gender gap	0.031	0.040	

Notes. Sample sizes are 44 women and 42 men in the B treatment and 43 women and 45 men in the EA treatment. p-values are from two-sided t-tests for differences between means.

Support for Result 2.

Table 3 shows participants' mean beliefs concerning their co-participant's performance in Parts 1 and 2.

Table 3 shows that, not only was there no free riding in teams but participants did not expect free riding in teams either. Specifically, if they expect their partner to free ride in the team setting, they should expect a lower level of output from him in Part 2 than in Part 1. This is decidedly not the case; in fact, they expect a small, but statistically significant improvement in their partner's performance in the team setting. This increase could reflect some anticipated learning, or even an expected motivational benefit of the team environment.²⁰

The other key finding from Table 3 is a highly significant gender gap in expectations of the partner's ability: as a number of other studies have found, both men and women expect their partner to be less able than themselves. Strikingly, however, men have much lower expectations of their partner's ability than women. This gender gap in expectations is quantitatively large. For example, in the B treatment, women's mean assessment of their partner's Part 1 performance (55.14 units) is at the 47th percentile of participants' actual Part 1 performance. For men, the corresponding figure (50.26 units) is at the 16th percentile of actual performance.²¹

²⁰ In addition, behavioural evidence that subjects did not expect moral hazard is provided in the next Section: participants did not choose teams more often in Part 4 than Part 3, which suggests that they did not expect their partner to be less productive when team pay was imposed.

²¹ One might wonder whether participants' perceptions of their partner's ability depend on the partner's gender as well. This issue is explored in Table D1 in the online Appendix, which shows that neither actual performance nor perceived partner performance depends on the partner's gender. The only partner gender effect that approaches statistical significance is a tendency for men to rate male partners higher than female partners ($p = 0.149$). These results contrast with some studies on competitiveness showing that the partner's gender can influence the performance in competition (Gneezy *et al.*, 2003; Iriberry and Rey-Biel, 2011) or the decision to enter a tournament (Datta Gupta *et al.*, 2013). For example, Iriberry and Rey-Biel (2011) show that information on the rival's gender decreases women's performance in competition but only when the task is perceived to favour men because of stereotype threat.

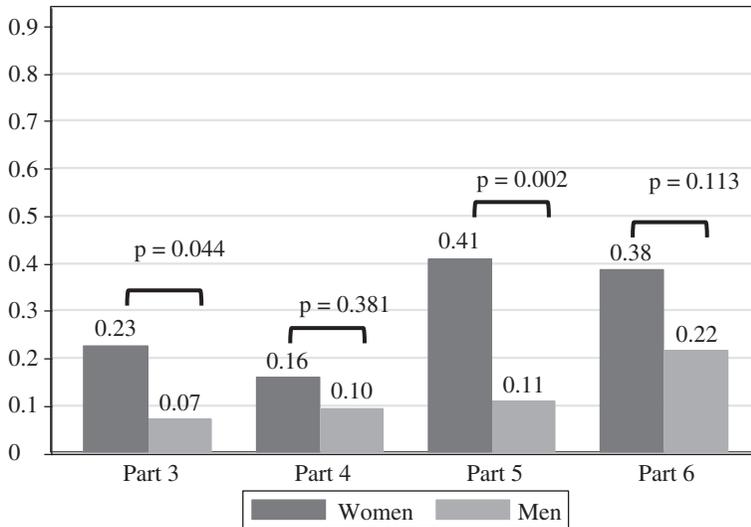


Fig. 1. *Share of Participants Choosing Team Compensation, by Gender, B Treatment*
 Note. p-values are from two-tailed t-tests.

3. Gender and Team Choice: Basic Results

This Section summarises participants' choices in the four 'team choice' parts of the experiment (Parts 3–6) in which subjects were free to choose their compensation scheme. We begin with the B treatment.

RESULT 3. *In the absence of efficiency advantages, women are much more likely to choose the team payment scheme than men.*

Support for Result 3.

The share of men and women who choose team compensation in the B treatment is shown in Figure 1.

According to Figure 1, female participants elected to receive team-based pay more frequently than male participants in all the team choice parts. Specifically, in Part 3 – which provides the simplest test of selection into teams – more than three times as many women as men (22.73% *versus* 7.14%; $p = 0.044$) chose to be paid on a team basis; recall that this choice was made despite the lack of any efficiency advantages to team production and despite the fact that the team situation could in principle expose the participant to a risk of free riding by a self-interested teammate. Perhaps surprisingly, despite designing the experiment to ensure that participants knew their partner's gender, we found no evidence that the partner's gender affected the decision to receive team-based compensation.²²

²² Table D2 in the online Appendix shows regression coefficients (with no controls) that parallel the statistics reported in Figures 1 and 2 of the study, broken down by the partner's gender. Table D3 in the online Appendix presents regression results that replicate Table 4 of the study. It clearly shows higher team formation rates if the participant is female (unless a control for beliefs is introduced), regardless of whether her prospective teammate was male or female.

Comparing Parts 3 and 4 strongly confirms our (incentivised) questionnaire-based evidence that subjects did not expect free riding in the team environment: in fact, women selected team compensation less frequently in Part 4, although Part 4 protects them against free riding by pairing them with their co-participant's output under the individual piece rate. Men, on the other hand, selected teams slightly more in Part 4, although for both men and women the difference between their Parts 3 and 4 choice is statistically insignificant ($p = 0.262$ for women and $p = 0.570$ for men, two tailed). Comparing Parts 5 and 6, there is no indication that women were more attracted to teams when the team experience was more interactive. In fact, women selected team compensation less frequently in Part 6 but insignificantly so ($p = 0.767$), despite the fact that Part 6 allows for a period of communication between the team members prior to production. Men, on the other hand, selected teams significantly more often in Part 6 than in Part 5 ($p = 0.044$). This difference could be due to gender differences in communication preferences or in the search for information (during the chat interactions, many try to learn the performance of their teammate in the previous parts).

RESULT 4. Women are much more likely to choose the team pay option when choosing 'team' affects their partner's pay-off (Part 5).

Support for Result 4.

Comparing Parts 3 and 5, we find no difference in men's behaviour: team pay was chosen 10.81 *versus* 7.14% of the time ($p = 0.160$). Women, on the other hand, are much more likely to choose teams (41.03 *versus* 22.73% of the time, $p = 0.033$) in Part 5, which captures more features of the team production environment.²³ Section 4 of our study argues that the distribution of the gap among our subjects is consistent with advantageous inequity aversion in women.

Turning next to the EA treatment, we have:

RESULT 5. The gender gap in the willingness to form a team vanishes when efficiency advantages are introduced. Both genders increase their team choices, but men's increase is much larger.

Support for Result 5.

Figure 2 displays the share of participants who choose team compensation in the EA treatment.

Two features of the results are immediately apparent: first, despite only a small improvement in efficiency associated with team production, the share of both men and women choosing team compensation is much higher than in the B treatment. In all cases the new rates of team choices are above 50%.²⁴ Second, the gender gap

²³ At 6.8%, the share of women who ultimately formed a team was, not surprisingly, actually lower in Part 5 than in Part 3. For men, team formation was the same in Parts 3 and 5, at 7.1%.

²⁴ It is perhaps worth noting that, with a 10% efficiency advantage to team production, the marginal private return to effort in teams rises from 50% to 55% of the marginal return to effort under individual compensation. Thus, in a 'standard' model, we should still expect high levels of free riding in teams and would therefore still expect most if not all participants rationally to avoid the team environment.

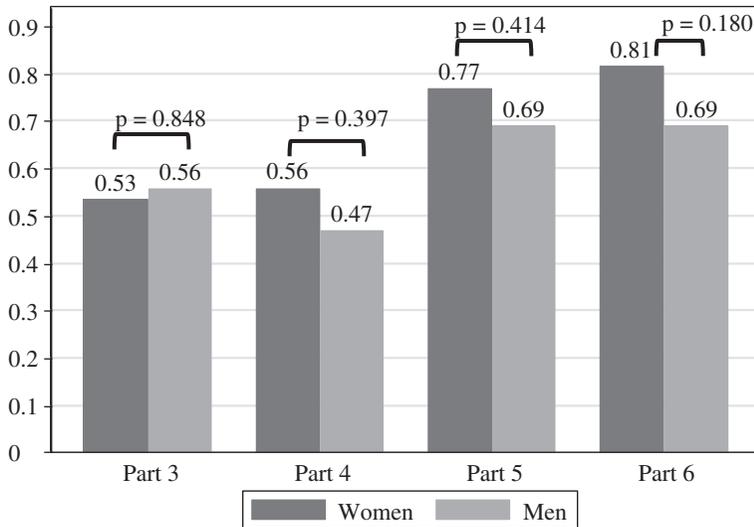


Fig. 2. *Share of Participants Choosing Team Compensation, by Gender, EA Treatment*

Note. p-values are from two-tailed t-tests.

in team selection vanishes: Although women still choose teams more frequently than men, the gap is much smaller in magnitude and insignificant as men's propensity to choose teams rises much more between the B and EA treatments. It thus appears that men are more responsive to the introduction of these extrinsic benefits of being on a team. This finding echoes those of Andreoni and Vesterlund (2001) who found that men's demand curve for altruism is more reactive to price changes.²⁵

Comparing Parts 3 and 4 of the EA treatment, there is once again no indication that participants expected moral hazard in the team environment (t-tests, $p = 0.743$ for women and $p = 0.253$ for men, two tailed). Comparing Parts 5 and 6, there is now no indication that participants of either gender were more attracted to teams when the team experience was more interactive ($p = 0.421$ for women and $p = 1.0$ for men).²⁶ Finally, comparing Parts 3 and 5, we find additional support for stronger social preferences among women: Women are significantly more likely to choose teams (76.74 versus 53.49% of the time) when subjects' team production decisions affect their partners' pay-offs ($p = 0.003$). For men, this difference is smaller (68.89 versus 55.56% of the time) and marginally significant ($p = 0.083$) (it was insignificant in the B case).

Our next result concerns the role of abilities and beliefs in accounting for gender differences in team formation:

²⁵ If participants interpret the EA treatment's higher piece rate for teams as the experimenter's preferred choice, then their responses to the EA treatment could include an experimenter demand effect. To account for our results using experimenter demand effects, however, men would need to be more eager to please the experimenter than women. We are not aware of any gender dimension in experimenter demand effects (Zizzo, 2010).

²⁶ All the differences mentioned are small in magnitude and none are statistically significant at conventional magnitudes. Perhaps an effect of interaction would be found if we allowed for collaboration on the work task itself.

RESULT 6.

- (i) *Regardless of the institutional environment (i.e. the rules for team formation) and the economic environment (i.e. the efficiency gains associated with team production), abler participants of both genders are always less likely to choose team compensation.*
- (ii) *Similarly, participants who thought their co-participant was abler were more likely to choose team compensation.*
- (iii) *Statistically, women's more generous beliefs about their partner's ability account for all the gender gap in team choices in Part 3 of the experiment.*

Support for Result 6.

To explore the role of abilities and beliefs on choices, Table 4 regresses a team choice indicator on measures of individual ability, risk attitudes and beliefs about partner ability.²⁷ The sample for Table 4 pools both the B and EA treatments but, to focus on the cleanest tests of preferences for team pay, restricts attention to the two key team choice parts of the experiment: Parts 3 and 5.²⁸ All the regressions use the subject's own performance in Part 1 as a measure of ability and his estimate of the partner's Part 1 performance to measure expected partner ability, and include gender-specific fixed effects for the parts and treatments.²⁹ As Parts 3 and 5 were administered to the same subjects, all the regressions are clustered at the individual level.

With no controls except the parts and treatments interacted with gender, column (1) of Table 4 replicates the significant positive gender effect of about 15 percentage points in the regression's baseline treatment and condition (Part 3 of the B treatment). It also shows that workers of both genders elect to form teams more often when team choices affect one's partner's pay-offs (Part 5), although the effect is twice as large and statistically more significant for women ($p = 0.053$ for men; $p < 0.001$ for women). Both genders also respond positively to efficiency advantages in team production, although here the effect is much larger for men than women.³⁰ These effects of Part 5 and the EA treatment on each gender are robust to the addition of controls for ability, risk aversion and expected partner ability, as indicated in columns 2–4.

Row 2 of Table 4 introduces a control for the participant's own Part 1 performance. As claimed, abler participants are less likely to choose the team environment in all specifications. As abler participants will earn a lower pay-off when

²⁷ Risk aversion could affect decisions, although its predicted effects are ambiguous in sign. On the one hand, if there is a lot of part-to-part variation in performance, greater risk aversion might lead participants to prefer teams, as being paid the average of the two workers' performance adds an element of insurance. On the other hand, uncertainty about the ability of one's teammate works in the opposite direction. We also detected no statistically significant gender gap in risk aversion in our subject pool.

²⁸ Recall that interpretation of Part 4 – whose main purpose was to test for expectations of moral hazard – is complicated by the fact that it is the only part where no production actually occurs. Interpretation of Part 6 is complicated by the fact that it is the only part where subjects have received any information about their partner's past choices. That said, the main results of Table 4 are confirmed when we pool data from all four team choice parts by simply using the total number of times (from zero to four) a participant chose team in those parts as our outcome variable. Please see Table D4 of the online Appendix.

²⁹ Arguably we could use expectations of their partner's Part 2 output as this reflects his performance in a team environment. In practice, given the lack of free riding it makes no difference which measure we use.

³⁰ The p-value for a test of no gender difference in the Part 5 coefficient is 0.092, whereas the p-value for no gender difference in the EA coefficient is 0.080.

Table 4
Regressions for the Probability of Choosing Team Pay in Parts 3 and 5

	(1)	(2)	(3)	(4)
1. Female	0.1681*	0.1985**	0.1864*	0.0720
	(0.0750)	(0.0716)	(0.0720)	(0.0731)
2. Own Part 1 output		-0.0151**	-0.0143**	-0.0289**
		(0.0051)	(0.0050)	(0.0054)
3. Holt–Laury switch point			-0.0364	-0.0350
			(0.0204)	(0.0187)
4. Multiple switches			0.1400	0.1164
			(0.1200)	(0.1133)
5. Beliefs re partner's Part 1 output				0.0301**
				(0.0061)
6. Part5 × Male	0.0882	0.0906*	0.0948*	0.0956*
	(0.0453)	(0.0452)	(0.0447)	(0.0449)
7. Part5 × Female	0.2083**	0.2080**	0.2058**	0.2092**
	(0.0543)	(0.0544)	(0.0545)	(0.0544)
8. EA × Male	0.5308**	0.5383**	0.5202**	0.4936**
	(0.0740)	(0.0688)	(0.0694)	(0.0702)
9. EA × Female	0.3316**	0.3009**	0.2996**	0.3048**
	(0.0855)	(0.0867)	(0.0853)	(0.0757)
10. R-squared	0.242	0.272	0.288	0.352
p-values for				
11. Part5 × Male = Part5 × Female	0.092	0.098	0.117	0.109
12. EA × Male = EA × Female	0.080	0.032	0.045	0.067

Notes. * $p < 0.01$, ** $p < 0.05$. Estimates are from linear probability models. Standard errors, clustered on participants, in parentheses. $N = 338$ in all columns.

matched with a random teammate than when working independently, this pattern is consistent with a simple, own pay-off–maximising adverse selection hypothesis. Also consistent with adverse selection, column 4 of Table 4 shows that participants who thought their co-participant was abler were more likely to choose team compensation. Notably, the effect of higher estimated partner ability is equal in magnitude, but opposite in sign to the effect of own ability. Another robust result from both Tables is that regardless of specification, risk attitudes do not predict the choice of team compensation.³¹

Finally, the gender gap in team choice, although robust to controls for own ability and risk aversion, falls sharply in magnitude and becomes insignificant when controls for perceived partner ability are added in column 4. This demonstrates our claim that gender differences in beliefs about others' ability are a major determinant of the gender gap in team choices.³²

³¹ Table D3 in the online Appendix replicates Table 4, allowing for teammate gender to affect team choices. No significant effects are detected.

³² To quantify the effects of men's more biased beliefs on their pay-offs, we compared the participants' actual team formation rates and part-specific pay-offs to what they would have been if the participants picked the compensation scheme that maximised their earnings given unbiased expectations. While unbiased expectations generate much higher rates of team choice, especially among men, they did not increase men's relative pay-offs much. For example, in Part 3 of the B treatment, unbiased expectations would raise men's team choices from 0.071 to 0.524 and women's from 0.227 to 0.432. Pay-offs, however, only rise from €11.55 to €11.63 for men and from €11.71 to €11.75 for women due to the compressed distribution of performance on this task among our participants.

4. Social Preferences and the ‘Part 5 Effect’

One of our most unexpected findings is the large effect of apparently minor changes in the rules of the team formation process between Parts 3 and 5 on the frequency with which participants, especially women, choose the team option. Why would requiring mutual consent for team formation almost double the share of female subjects choosing the team option in our baseline treatment, and raise it by over 20 percentage points in the EA treatment? In this Section we briefly sketch and test a simple social preference-based explanation of this phenomenon. Specifically, Appendix A shows formally that if the subjects in our experiment have Fehr and Schmidt (1999) inequity aversion and no moral hazard is present, their rates of team choice should increase between Parts 3 and 5 but only for subjects who perceive they are abler than their partners. Noteworthy features of the model are that:

- (i) all predicted behaviour takes the form of strictly dominant strategies in the simultaneous move game where both partners are choosing between the individual and team pay options; and
- (ii) all predicted behaviour is independent of subjects’ degree of disadvantageous inequity aversion (i.e. aversion to having less income than one’s partner).

The intuition behind the model’s predictions is simple: in Part 3 (by construction), subjects’ team choices have no impact on their partner’s pay-offs. Thus, all subjects behave selfishly. In Part 5, inequity-averse subjects who perceive themselves as less able than their partner will also behave selfishly: they should all choose the team option because it both raises their own income and reduces the income gap between them and their partner. Thus, the only type of agent who should change their team choices between Parts 3 and 5 is someone who has advantageous inequity aversion (i.e. dislikes having more income than her partner) and believes she is abler than her partner (because only then can she help her partner by teaming up with them). Thus, if the increase in women’s team choices between Parts 3 and 5 of our experiment is confined to the subset of women who think they are abler than their partner, our results are consistent with a model in which women have more advantageous inequity aversion than men. As it happens, this is exactly what we find:

RESULT 7. The stronger attraction of team compensation on women in Part 5 when mutual consent is required to form a team, compared with Part 3, is compatible with advantageous inequity aversion of women.

Support for Result 7.

The rates of team choice in Parts 3 and 5, by a subject’s own perceived relative ability and gender, are presented in Table 5.

According to Table 5, women who thought they were less able than their partner increased their team formation rates by 9 percentage points (0.73–0.64) between Parts 3 and 5. This increase is not statistically significant. Among women who thought they

Table 5
Team Choice Rates in Parts 3 and 5, by Perceived Relative Ability

	(1) Chose team in Part 3	(2) Chose team in Part 5	(3) N
<i>(a) Women</i>			
<i>Less able than partner</i>			
Share	0.64	0.73	
Number	21	24	33
<i>Abler than partner</i>			
Share	0.22	0.51	
Number	11	25	49
<i>(b) Men</i>			
<i>Less able than partner</i>			
Share	0.47	0.47	
Number	7	7	15
<i>Abler than partner</i>			
Share	0.30	0.42	
Number	20	28	67

Notes. p-values for (Part 5 – Part 3 among abler) = (Part 5 – Part 3 among less able): Women: $p = 0.086$; Men: $p = 0.273$. p-values are for the significance of an interaction term between Part 5 and relative ability in a gender-specific regression of team formation on Part, relative ability and their interaction. Robust standard errors are clustered on participant id. Statistics are for the B and EA treatments combined.

were more able than their partner, this increase was much larger, at 29 percentage points. The difference is statistically significant at the 10% level ($p = 0.086$) and is consistent with advantageous inequity aversion. For men who thought they were less able than their partner, there was no difference in team choices between Parts 3 and 5 (0.47 in both cases), whereas the share choosing the team payment rose by 12 percentage points among men who thought they were abler than their partner. This difference is statistically insignificant ($p = 0.273$). Overall, the fact that the increase in women's team formation rates between Part 3 (where abler women cannot help their partner by choosing 'team') and Part 5 (where they can) is concentrated among women who thought they were abler than their partner is consistent with our predictions.

Importantly, the fact that less-able women's team choices do not fall between Parts 3 and 5 in Table 5 also casts doubt on a simpler model of social preferences, based on a higher level of pure altruism among women. In such a model, altruistic agents care about their partner's income irrespective of who has the higher income. As demonstrated in Appendix A, in this model sufficiently altruistic agents who think they are less able than their partner should reduce their team choice rates between Parts 3 and 5, to avoid hurting their (more able) partner. This is not the case in our data, either for women or for men, suggesting that inequity aversion rather than pure altruism is a more appropriate model here.

Additional possible interpretations of the Part 5 effect include the possibility that by choosing team compensation in Part 5:

- (i) participants will learn the level of their partner's pay;
- (ii) participants gain an opportunity to help their partner by raising their own effort levels;

- (iii) selfish participants will have an abler teammate if a team is actually formed; and
- (iv) participants can guarantee that they will not frustrate their partner's possible desire to form a team.

Our design does not allow us to test whether there is an intrinsic value to knowing one's partner's pay (*i*). Possibility (*ii*) seems unlikely given the fact that effort levels do not vary between team and individual pay environments.³³ As for possibility (*iii*), although we did not elicit the participants' beliefs regarding how others will self-select into team compensation, Results 6 and 8 show that actual selection into teams is negative, not positive. Finally, recall from the proofs in Appendix A that an inequity aversion-based model provides a 'built-in' explanation of why only participants who think they are abler than their partner should switch to team pay in Part 5: It is only those participants who would help their partner by sharing output with them. Importantly, this automatic connection does not hold for an intentions-based explanation: there is no *a priori* reason to expect more confident participants to be more averse to frustrating a prospective teammate's intentions than less confident participants.³⁴ Accordingly, we think that an inequity aversion-based social preferences interpretation seems to be the most adequate to account for the observed 'Part 5 effect'.

5. Gender and Performance in Self-selected Teams

Result 1 stated that, when assigned to work in a team environment, men and women perform equally well, both in absolute terms and relative to their performance when compensated individually. However, this zero treatment effect of teams does not necessarily imply that men and women will perform equally on teams that are voluntarily formed. One reason is ability bias or selection on levels: indeed, as adverse selection is stronger among men, we would expect voluntarily formed female teams to outperform voluntarily formed male teams. The second reason may not work in this direction. Suppose that subjects who expect that they will respond better to the team environment disproportionately select into teams (i.e. there is selection on slopes, or on responsiveness to the treatment). If the amount or direction of selection on slopes differs between men and women, then voluntarily formed male teams could perform better or worse than voluntarily formed female teams. Fortunately, our design allows us to non-

³³ The fact that effort levels do not vary between the team and individual environments is consistent with two scenarios: in one, subjects who wish to help their partner simply choose to do so only on the team choice margin, not the effort margin. In the other, subjects are unable to adjust on the effort margin because they are already working at or near capacity in all parts and treatments. Given that effort is also unresponsive to variations in the piece rate within treatments – see footnote 19 – we think the second scenario is more likely. In neither case, however, can a desire to help the partner by working harder explain the empirical Part 5 effect in our experiment.

³⁴ We also have a small amount of direct evidence on the role of intentions from the exit survey. In that survey, 21.2% of the women choosing the team compensation in Part 5 in the EA treatment justify their choice by the desire not to disappoint the partner in case s/he wanted to form a team; only 16.1% of men invoke this reason (the difference is, however, not significant, $p = 0.609$). Unfortunately, the reasons proposed to choose the team compensation did not explicitly include an inequity aversion-based motivation (see Table D5 in the online Appendix for a complete tabulation of responses).

parametrically estimate causal effects of the team environment on the persons who chose that environment (i.e. the treatment on the treated (ToT) effect) without any assumptions on the form of unobserved heterogeneity. Thus, we can determine not only whether self-selected teams perform differently from mandated teams but also why.

Our main findings are as follows:

RESULT 8.

- (i) *Self-selected teams perform worse than mandated teams due to strong selection on levels, but little evidence is found of selection on slopes.*
- (ii) *Because adverse selection is worse among men, self-selected male teams perform worse than self-selected female teams.*

Support for Result 8.

Appendix B presents performance information for team joiners and non-joiners in Parts 3 and 5, showing that performance on teams was worse than performance under an individual payment scheme. It also presents information on the Part 1 performance of persons who voluntarily joined teams in Parts 3 and 5, which demonstrates a strong role for negative self-selection into teams. Finally, Appendix B uses information on each participant's Part 1 output to partition the performance gap between voluntary team joiners and non-joiners into two components: selection on levels and selection on slopes. It shows that selection on levels is the dominant effect.

These findings suggest that firms offering voluntary teamwork should be wary of adverse selection and especially so in male workplaces: Because men more strongly underestimate their co-workers' ability, only men with low productivity will tend to sign up for the teams.

6. Discussion and Conclusion

In sharp contrast to the literature on gender and competition, this study finds that women's task performance is not adversely affected by entering an environment where rewards are team based and that women do not avoid entering team environments. Indeed, in most of our experimental treatments, women are more likely than men to enter team-based environments, although men's and women's attraction to those environments is shown to depend in important ways on subjects' perceptions of others' abilities, the EA of team production and subjects' social preferences. Taken together, we hope that these results might help shed light on the substantial and continuing gap in the occupational distribution of men and women, even in societies where a great deal of equality of opportunity exists. Indeed, women are still highly overrepresented in the non-profit sector and in helping occupations (Powell and Steinberg, 2006), both of which arguably involve co-operative production with little marginal financial reward to individual effort. Our results also suggest that replacing tournaments by team-based incentives in highly paid jobs (at least where appropriate given the production technology) might increase women's representation in those jobs.

At the same time, we also recognise that our concept of a team is a mere shadow of what it means to be on a real workplace team. Our team members interact only a few times, never meet in person and our linear production process does not exhibit any complementarities. Moreover, moral hazard in teams plays no role in our results and we deliberately eliminated any element of competition between teams. Finally, our task is gender neutral, while a task perceived by women as favouring men might increase their attraction for the team payment scheme when paired with a male partner. Addressing all these issues would seem to be useful directions for additional laboratory and field-based studies of selection into both competitive and co-operative work environments.

Appendix A. Proofs

Following Fehr and Schmidt (1999), let individual i 's utility be given by:

$$U_i = Y_i - \alpha(Y_j - Y_i)(1 - P) - \beta(Y_i - Y_j)P, \quad (\text{A.1})$$

where Y_i is i 's experimental income, Y_j is his partner's income, P is an indicator for whether $Y_i > Y_j$, $\alpha \in [0, 1)$ denotes disadvantageous inequity aversion and $\beta \in [0, 1)$ denotes advantageous inequity aversion. In what follows, we derive subjects' predicted choices in the B treatment. Proofs (available from the authors) are exactly parallel for the EA treatment and the only difference in results is that the threshold perceived relative ability level above which inequity-averse subjects might switch their choice from individual to team pay between Parts 3 and 5 changes from 1 to 11/9.

In Part 3 of the baseline treatment, i 's income if he chooses 'team' is $20\bar{Q}$, where $\bar{Q} = (Q_i + Q_j)/2$ with Q_i the subject i 's performance. His income if he chooses individual compensation is $Y_i = 20 Q_i$. Similarly, j 's income is $20\bar{Q}$ if j chooses 'team' and $20 Q_j$ otherwise. Thus, in Part 3, each player's income – by design – depends only on his own actions. This does not, however, imply that the parties' utilities are independent of each others' actions; in fact they are interdependent whenever the players are inequity averse. Thus, for example, player i 's utilities as a function of both i 's and j 's choices in Part 3 of the B treatment are shown below, for the case where i believes he is abler than j .

Player i 's utilities if i believes he is abler than j : Part 3 of the B treatment

i 's choice	j 's choice	
	Team	Individual
Team	$20\bar{Q}$	$20[(1 - \beta)\bar{Q} + \beta Q_j]$
Individual	$20[\beta\bar{Q} + (1 - \beta)Q_i]$	$20[(1 - \beta)Q_i + \beta Q_j]$

As $Q_i > \bar{Q} > Q_j$ and $0 < \beta < 1$, given these beliefs person i 's strictly dominant strategy is to pick individual compensation, whether or not person i is inequity averse. The intuition is obvious in the absence of inequity aversion – abler workers sacrifice income if they share team revenue with a less-able worker. Inequity-averse workers should behave the same way as selfish workers in Part 3 because i 's choice has no effect on j 's income, so i might as well just maximise her own income.

Turning to the case where i believes he is less able than j , player i 's utilities are given by

Player i 's utilities if i believes he is less able than j : Part 3 of the B treatment

i 's choice	j 's choice	
	Team	Individual
Team	$20\bar{Q}$	$20[(1 - \alpha)\bar{Q} - \alpha Q_j]$
Individual	$20[Q_i - \alpha(\bar{Q} - Q_i)]$	$20[(1 - \alpha)\bar{Q} - \alpha Q_j]$

Now, i 's dominant strategy is to pick team pay, regardless of his level of inequity aversion. The intuition is that selfish and inequity-related motives now both push i in the same direction: picking 'team' both raises i 's absolute pay, and reduces the gap between i 's pay and his partner's. This is true regardless of j 's decision.

Next, we turn to Part 5 of the B treatment. Here, i 's income equals $Y_i = 20 Q_j$ unless both i and j choose team compensation. If both players pick 'team', they each receive an income of $20 \bar{Q}$. So, each player's income now depends on the other's choices. Player i 's utilities as a function of i 's and j 's choices and i 's beliefs are now as follows:

Player i 's utilities if i believes he is abler than j : Part 5 of the B treatment

i 's choice	j 's choice	
	Team	Individual
Team	$20\bar{Q}$	$20[(1 - \beta)Q_i + \beta Q_j]$
Individual	$20[(1 - \beta)Q_i + \beta Q_j]$	$20[(1 - \beta)Q_i + \beta Q_j]$

Player i 's utilities if i believes he is less able than j : Part 5 of the B treatment

i 's choice:	j 's choice:	
	Team	Individual
Team	$20\bar{Q}$	$20[(1 - \alpha)Q_i - \alpha Q_j]$
Individual	$20[(1 - \alpha)Q_i + \alpha Q_j]$	$20[(1 - \alpha)Q_i - \alpha Q_j]$

Now, player i 's choices have no effect on either player's incomes (and hence on i 's own utility) if j chooses individual compensation. Thus, i again has a dominant strategy: he should pick whichever choice maximises his utility in the event that j chooses 'team'. If i believes he is less able than j that choice is always team compensation because, as before, selfish and inequity-related motives work in the same direction. If i believes he is abler than j , basic algebra reveals that i picks individual pay if he is not too inequity averse (specifically, if $\beta < 0.5$); otherwise i will pick team compensation.

Finally, consider how the predictions of the above model contrast with a pure altruism model where agents care positively about their partner's income, irrespective of which person has the higher income. Formally, this corresponds to the case where – instead of both α and β being

positive $-\alpha$ is negative and $\alpha = -\beta$. Applying the above pay-off tables to this case of the model changes none of the predictions for players who think they are able than their partner and changes none of the predictions for behaviour in Part 3. However, 'less able' players who are sufficiently altruistic (specifically, $\beta > (\bar{Q} - Q_i)/2\bar{Q}$) will now switch from choosing 'team' in Part 3 to individual pay in Part 5. They do this to avoid hurting their more-able partner.

Appendix B. Selection on Slopes *versus* Levels and the Performance of Self-selected Teams

For Parts 3 and 5 of the experiment, Table B1 shows the mean performance of participants who chose team compensation and those who did not, by gender and treatment.³⁵

In seven of eight cases, team joiners performed worse on average than non-joiners; this gap is statistically significant (at $p < 0.001, 0.004$ and 0.058) in three of those cases. Furthermore, the joiner–non-joiner output gap is always larger for men than for women, significantly so ($p = 0.001$ and 0.028) in two of the four cases. A naïve interpretation of these results might, of course, be that team production causes free riding and that men are more prone to free riding than women. On the other hand, we have already shown that self-selection affects these comparisons; this is confirmed by Table B2, which shows Part 1 output separately for individuals who chose teams in Parts 3 and 5, and for individuals who did not. Overall, the patterns are very similar to Table B1: team joiners were, on average, less productive than non-joiners even in Part 1 of the experiment, when everyone was compensated individually. This joiner–non-joiner 'ability' gap is always greater for men than women, significantly so the same

Table B1
Mean Performance of Participants, by Choice of Team Pay and Treatment

	Part 3				Part 5			
	(1) Non-joiners	(2) Joiners	(3) Gap (2) – (1)	(4) p-value for gap	(5) Non-joiners	(6) Joiners	(7) Gap (6) – (5)	(8) p-value for gap
<i>(a) Baseline treatment</i>								
Women	59.32	56.90	–2.42	0.203	60.48	58.38	–2.10	0.119
Men	58.18	47.33	–10.85	0.004	60.00	49.50	–10.50	0.000
p-value for gender gap	0.385	0.020	0.001		0.704	0.003	0.028	
<i>(b) Efficiency advantage treatment</i>								
Women	57.70	56.43	–1.27	0.418	57.40	57.70	0.30	0.865
Men	58.80	54.80	–4.00	0.058	58.36	56.52	–1.84	0.372
p-value for gender gap	0.536	0.380	0.272		0.724	0.379	0.438	

Notes. In Part 3, joining denotes the participant's decision to be paid via team compensation; in Part 5 this only results in team compensation being paid if the co-participant also chose team compensation; however, performance levels are very similar for the (smaller) sample who actually formed teams. All p-values are from two-sided t-tests for differences between means, except those for the gender gap in joiner–non-joiner difference. These come from a Chi-squared test that accounts for the fact that the underlying joiner–non-joiner difference is itself a difference in sample means.

³⁵ Recall that, in Part 5, choosing team compensation does not necessarily result in being paid on a team basis. When we repeat the analysis only for participants who successfully formed teams (because their partner also chose team compensation), they are very similar, though with a smaller sample size.

Table B2
Mean Part 1 Performance of Participants, by Choice of Team Pay in Parts 3 and 5

	Part 3				Part 5			
	(1) Non-joiners	(2) Joiners	(3) Gap (2) – (1)	(4) p-value for gap	(5) Non-joiners	(6) Joiners	(7) Gap (6) – (5)	(8) p-value for gap
<i>(a) Baseline treatment</i>								
Women	57.15	56.20	-0.95	0.624	57.83	55.56	-2.26	0.218
Men	55.49	46.33	-9.15	0.011	56.30	45.75	-10.55	0.001
p-value for gender gap	0.212	0.011	0.009		0.311	0.004	0.035	
<i>(b) Efficiency advantage treatment</i>								
Women	56.05	53.87	-2.18	0.214	53.90	55.18	1.28	0.539
Men	57.05	54.24	-2.81	0.110	57.64	54.52	-3.13	0.097
p-value for gender gap	0.527	0.838	0.789		0.200	0.615	0.146	

Note. See Table B1.

two of four times as for current period output ($p = 0.009$ and 0.035). This gap is never statistically significant for women. Thus, confirming our earlier results, not only is there adverse selection into teams but this selection is more pronounced among men than women.

Finally, to isolate the effects of selection on levels and selection on slopes, note that the output gap between voluntary team joiners and non-joiners in (say) Part 3 can be expressed as:

$$\begin{aligned}
 E(Q^3|J^3) - E(Q^3|NJ^3) &= E(Q^1|J^3) - E(Q^1|NJ^3) \\
 &+ [E(Q^3|J^3) - E(Q^1|J^3)] - [E(Q^3|NJ^3) - E(Q^1|NJ^3)]
 \end{aligned}
 \tag{B.1}$$

where Q^i denotes the individual's performance in Part i , and J^i and NJ^i , respectively, denote whether the participant joined or did not join a team in Part i .

The first term on the RHS of (B.1) gives the pure selection component of the performance gap between joiners and non-joiners in Part 3; it is simply the difference in Part 1 output between the participants who chose team in Part 3 and those who did not. These numbers are already shown in Table B2, which are predominantly negative, showing adverse selection into teams. The remaining two terms in (B.1) are the output change between Parts 1 and 3 for Part 3 joiners, minus the output change for Part 3 non-joiners. Only the first of these changes (for joiners) is affected by a change in the production environment from an individual to a team situation, whereas both are affected by any task learning or pure period effects that distinguish Parts 1 and 3. Therefore, if these part and task learning effects are the same for persons who eventually join teams and those who do not – this is our identifying assumption – then the last two terms in (B.1) estimate the pure ToT effect of the team environment.

Table B3 decomposes the team–non-team output gaps identified in Table B1 using (B.1). While the magnitudes of ToT and selection effects vary substantially across treatments and parts of the experiment, Table B3 shows clearly that selection plays the dominant role in explaining the gaps observed in Table B1: in five of the eight possible cases, selection accounts for more than 100% of the team–non-team output gaps; in seven of eight cases it accounts for more than half of the output gap. Also noteworthy is the fact that, in four of the eight cases (two involving

Table B3

Decomposition of Performance Gap between Joiners and Non-Joiners, Parts 3 and 5

	Part 3			Part 5		
	(1) Gap due to selection	(2) Gap due to behavioural change	(3) Total gap	(4) Gap due to selection	(5) Gap due to behavioural change	(6) Total gap
<i>(a) Baseline treatment</i>						
Women	-0.95	-1.48	-2.42	-2.26	0.16	-2.10
%	39.08	60.92	100.00	107.62	-7.62	100.00
Men	-9.15	-1.69	-10.85	-1.55	0.05	-1.50
%	84.40	15.60	100.00	100.51	-0.51	100.00
<i>(b) Efficiency advantage treatment</i>						
Women	-2.18	0.92	-1.27	1.28	-0.98	0.30
%	172.34	-72.34	100.00	431.64	-331.64	100.00
Men	-2.81	-1.19	-4.00	-3.13	1.29	-1.84
%	70.25	29.75	100.00	169.84	-69.84	100.00

men and two involving women), the ToT causal effect of team production is positive, not negative. This provides further support for our finding that moral hazard is essentially absent in this two-person team environment.

University of California, Santa Barbara, NBER, IZA and CES

Université de Lyon, CNRS-GATE and IZA

Submitted: 18 April 2013

Accepted: 30 October 2013

Additional Supporting Information may be found in the online version of this article:

Appendix C. Instructions for the B Treatment.

Appendix D. Supplementary Tables.

Data S1.

References

- Andersen, S., Bulte, E., Gneezy, U. and List, J.A. (2008). 'Do women supply more public goods than men? Preliminary experimental evidence from matrilineal and patriarchal societies', *American Economic Review: Papers & Proceedings*, vol. 98(2), pp. 376–81.
- Andreoni, J. and Vesterlund, L. (2001). 'Which is the fair sex? Gender differences in altruism', *Quarterly Journal of Economics*, vol. 116(1), pp. 293–312.
- Apesteguia, J., Azmat, G. and Iriberry, N. (2012). 'The impact of gender composition on team performance and decision-making: evidence from the field', *Management Science*, vol. 58(1), pp. 78–93.
- Babcock, P., Bedard, K., Charness, G., Hartman, J. and Royer, H. (2011) 'Letting down the team? Evidence of social effects of team incentives', Working Paper no. 16687, NBER.
- Berg, P. (1996). 'The performance effects of modular production in the apparel industry', *Industrial Relations*, vol. 35(3), pp. 3563–73.
- Boning, B., Ichniowski, C. and Shaw, K. (2007). 'Opportunity counts: teams and the effectiveness of production incentives', *Journal of Labor Economics*, vol. 25(4), pp. 613–50.
- Booth, A.L. and Nolen, P.J. (2012). 'Choosing to compete: how different are girls and boys?', *Journal of Economic Behavior & Organization*, vol. 81(2), pp. 542–55.

- Bornstein, G., Gneezy, U. and Nagel, R. (2002). 'The effect of intergroup competition on intragroup coordination: an experimental study', *Games and Economic Behavior*, vol. 41(1), pp. 1–25.
- Cárdenas, J.-C., Dreber, A., von Essen, E. and Ranehill, E. (2012). 'Gender differences in competitiveness and risk taking: comparing children in Colombia and Sweden', *Journal of Economic Behavior & Organization*, vol. 83(1), pp. 11–23.
- Charness, G., Masclot, D. and Villeval, M.C. (2013) 'The dark side of competition for status', *Management Science*, forthcoming.
- Cooper, D. J. and Jabs Saral, K. (2010) 'Entrepreneurship and team participation: an experimental study', mimeo, Florida State University.
- Crosen, R. and Gneezy, U. (2009). 'Gender differences in preferences', *Journal of Economic Literature*, vol. 47(2), pp. 448–74.
- Dargnies, M.-P. (2012). 'Men too sometimes shy away from competition: the case of team competition', *Management Science*, vol. 58(11), pp. 1982–2000.
- Datta Gupta, N., Poulsen, A. and Villeval, M.C. (2013). 'Gender matching and competitiveness: experimental evidence', *Economic Inquiry*, vol. 51(1), pp. 816–35.
- Delfgaauw, J., Dur, R., Sol, J. and Verbeke, W. (2013). 'Tournament incentives in the field: gender differences in the workplace', *Journal of Labor Economics*, vol. 32(2), pp. 305–26.
- Dohmen, T. and Falk, A. (2011). 'Performance pay and multidimensional sorting: productivity, preferences, and gender', *American Economic Review*, vol. 101(1), pp. 556–90.
- Dreber, A., von Essen, E. and Ranehill, E. (2012). 'Outrunning the gender gap – boys and girls compete equally', *Experimental Economics*, vol. 14(4), pp. 567–82.
- Eckel, C. C. and Grossman, P. J. (2008) 'Differences in the economic decisions of men and women: experimental evidence', in (C.R. Plott and V.L. Smith, eds.), *Handbook of Experimental Economics Results*, vol. 1, pp. 509–19, Amsterdam: Elsevier.
- Fehr, E. and Schmidt, K.M. (1999). 'A theory of fairness, competition, and cooperation', *Quarterly Journal of Economics*, vol. 114(3), pp. 817–68.
- Garratt, R.J., Weinberger, C. and Johnson, N. (2013). 'The state street mile: age and gender differences in competition aversion in the field', *Economic Inquiry*, vol. 51(1), pp. 806–15.
- Gneezy, U., Leonard, K.L. and List, J.A. (2009). 'Gender differences in competition: evidence from a matrilineal and a patriarchal society', *Econometrica*, vol. 77(5), pp. 1637–64.
- Gneezy, U., Niederle, M. and Rustichini, A. (2003). 'Performance in competitive environments: gender differences', *Quarterly Journal of Economics*, vol. 118(3), pp. 1049–74.
- Gneezy, U. and Rustichini, A. (2004). 'Gender and competition at a young age', *American Economic Review*, vol. 94(2), pp. 377–81.
- Graves, L. M. and Powell, G. N. (2007) 'Sex, sex similarity and sex diversity effects in teams: the importance of situational factors', in (D. Bilimoria and S.K. Piderit, eds.), *Handbook on Women in Business and Management*, pp. 217–31, Cheltenham: Elgar.
- Greiner, B. (2004). 'An online recruitment system for economic experiments', in (K. Kremer and V. Macho, eds.), *Forschung und wissenschaftliches Rechnen*, GWDG Bericht 63, pp. 79–93, Göttingen: Gesellschaft für Wissenschaftliche Datenverarbeitung.
- Hamilton, B., Nickerson, J. and Owan, H. (2003). 'Team incentives and worker heterogeneity: an empirical analysis of the impact of teams on productivity and participation', *Journal of Political Economy*, vol. 111(3), pp. 465–97.
- Healy, A. and Pate, J. (2011). 'Can teams help to close the gender competition gap?', *ECONOMIC JOURNAL*, vol. 121(555), pp. 1192–204.
- Holt, C.A. and Laury, S.K. (2002). 'Risk aversion and incentive effects', *American Economic Review*, vol. 92(5), pp. 1644–55.
- Iriberry, N. and Rey-Biel, P. (2011). 'Let's (not) talk about sex: the effect of information provision on gender differences in performance under competition', Department of Economics and Business, Universitat Pompeu Fabra, Economics Working Paper.
- Ivanova-Stenzel, R. and Kübler, D. (2011). 'Gender differences in team work and team competition', *Journal of Economic Psychology*, vol. 32(5), pp. 797–808.
- Knez, M. and Simester, D.I. (2001). 'Firm-wide incentives and mutual monitoring at Continental Airlines', *Journal of Labor Economics*, vol. 19(4), pp. 743–72.
- Lawler, E. E. III, and Mohrman, S. A. (2003). 'Pay practices in Fortune 1000 corporations', Center for Effective Organizations, publication G 03-20 (448).
- Lazear, E. P. (1999) 'Globalisation and the market for team-mates', *ECONOMIC JOURNAL*, vol. 109(454), pp. C15–40.
- Niederle, M. and Vesterlund, L. (2007) 'Do women shy away from competition? Do men compete too much?', *Quarterly Journal of Economics*, vol. 122(3), pp.1067–101.
- Powell, W.W. and Steinberg, R. (2006). *The Non-Profit Sector: A Research Handbook*, 2nd edition, New Haven, CN: Yale University Press.

- Sutter, M. and Rützler, D. (2010). 'Gender differences in competition emerge early in life: three-year old girls compete as much as boys, but older girls don't', mimeo, University of Innsbruck.
- Villeval, M.C. (2012) 'Steady, ready, compete', *Science*, vol. 335(6068), pp. 544–5.
- Wozniak, D., Harbaugh, W. T. and Mayr, U. (2010) 'The menstrual cycle and performance feedback alter gender differences in competitive choices', mimeo, University of Oregon.
- Wuchty, S., Jones, B.F. and Uzzi, B. (2007). 'The increasing dominance of teams in production of knowledge', *Science*, vol. 316(5827), pp. 1036–9.
- Zeiliger, R. (2000). 'A presentation of Regate, internet based software for experimental economics', GATE, available at: <http://www.gate.cnrs.fr/~zeiliger/regate/RegateIntro.ppt> (last accessed: 17 June 2011).
- Zizzo, D. (2010). 'Experimenter demand effects in economic experiments', *Experimental Economics*, vol. 13(1), pp. 75–98.