DISCUSSION PAPER SERIES

IZA DP No. 7523

Peer Pressure and Productivity: The Role of Observing and Being Observed

Sotiris Georganas
Mirco Tonin
Michael Vlassopoulos

July 2013

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Peer Pressure and Productivity: The Role of Observing and Being Observed

Sotiris Georganas
Royal Holloway, University of London

Mirco Tonin
University of Southampton,
CEU, CESifo and IZA

Michael Vlassopoulos
University of Southampton

Discussion Paper No. 7523
July 2013

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 7523
July 2013

ABSTRACT

Peer Pressure and Productivity: The Role of Observing and Being Observed*

Peer effects arise in situations where workers observe each other's work activity. In this paper we disentangle the effect of observing a peer from that of being observed by a peer, by setting up a real effort experiment in which we manipulate the observability of performance. In particular, we randomize subjects into three groups: in the first one subjects are observed by another subject, but do not observe anybody; in the second one subjects observe somebody else's performance, but are not observed by anybody; in the last group subjects work in isolation, neither observing, nor being observed. We consider both a piece rate compensation scheme, where pay depends solely on own performance, and a team compensation scheme, where pay also depends on the performance of other team members. Overall, we find some evidence that subjects who are observed increase productivity at least initially when compensation is team based, while we find that subjects observing react to what they see in a non-linear but monotonic way when compensation is based only on own performance.

JEL Classification: D03, J24, M52, M59

Keywords: peer effects, piece rate, team incentives, real-effort experiment

Corresponding author:

Mirco Tonin
Economics Division
School of Social Sciences
University of Southampton
Southampton SO17 1BJ
United Kingdom
E-mail: m.tonin@soton.ac.uk

* We thank Theodore Alysandratos for research assistance. This work was supported by the British Academy. The code implementing the slider task is based on the code developed by Gill and Prowse (2012a).

1 Introduction

In a recent survey of the insights for labor economics obtained in the lab, Charness and Kuhn (2011) define "pure" peer effects as applying to "a situation where workers work, side by side, for the same firm but do not interact in any way (except that they observe each others' work activity)" (p.205). An example of such a situation is the real effort experiment by Falk and Ichino (2006), in which subjects put letters into envelopes for a remuneration that is independent of output. They find evidence of positive peer effects, as subjects who work in pairs in the same room are more productive compared to those who work alone. When working in pairs, a worker can see what the other is doing and at the same time knows that the other worker is seeing what she is doing, that is, a worker is both observing and being observed. This distinction between observing and being observed is prominent in the work of Mas and Moretti (2009). They use data from a supermarket chain and exploit the spatial orientation of the cash desks to provide evidence that the positive productivity spillovers due to the introduction of highly productive cashiers into a shift come from other workers being observed by a high productivity worker and not from observing one.[1]

The aim of the current paper is to contribute to our understanding of the behavioral mechanisms behind peer effects by disentangling these two channels of peer pressure: observing a colleague's work and being observed by a co-worker. To this end, we design a real effort experiment in which subjects are in one of three positions: they can see the work of another subject, or they work knowing that another subject can see their work, or they work in isolation without observing or being observed. In particular, subjects work on the slider task developed by Gill and Prowse (2012a). After two rounds in which we measure baseline productivity, we randomly split subjects into three treatments. Subjects in the *Control* treatment do not observe and are not observed by anyone. The rest of the subjects are in treatment *Observed*, in which their performance is observed by another subject, or in treatment *Observer*, in which they observe another subject's performance. These roles are kept for the remaining 14 rounds, with the pairs of Observer-Observed rematched in each round using a random matching protocol. Notice that the only difference between the Control and Observed treatments is that someone is anonymously observing the performance of those in the Observed treatment. By comparing productivity in these two treatments, we can establish what is the impact of being observed by a peer. Moreover, we can identify whether observing another subject's performance has an impact on own performance. To do this, we look at whether what

[1] Another paper highlighting the effect of being observed is Corgnet et al. (2013). They allow all subjects to split their time between work, leisure and monitoring (observing others remotely). They find that when subjects are aware of when they are monitored, performance under team incentives increases to the same level as under individual incentives, while this is not the case when subjects are only aware of the possibility of being monitored.

they observe has an impact on productivity of those in the Observer treatment. The fact that those in the Observed treatment never receive information about anybody else's performance means that in our design there is no reflection problem (Manski, 1993).

We start by distinguishing the role of observing and being observed in a pure peer effects setting, where compensation is on the basis of a piece rate, so there are no payoff externalities across workers. We then extend the analysis to consider peer effects when compensation depends on team performance. Besides being an empirically relevant compensation structure, we expect peer effects to be stronger when compensation is interdependent across subjects, as in the analysis by Kandel and Lazear (1992).[2] We find that both being observed and being an observer have an impact on work performance, and that the type of compensation scheme matters for when these effects are active. In particular, we find that subjects who are observed raise productivity initially, as compared to the control group, when compensation is team based. Following this initial response the control group eventually catches up with the observed, suggesting that being observed spurs subjects to reach a higher level of productivity faster in our experiment. We also find that subjects observing react to what they see in a non-linear but monotonic way when compensation is based only on own performance. Specifically, they respond negatively to observing low performers and positively to high performers.

Peer effects can arise for a variety of reasons. In their theoretical contribution, Kandel and Lazear (1992) underline how the positive effects of peer pressure on effort can overcome free riding in environments with profit sharing. They distinguish between internal pressure (or guilt) and external pressure (or shame), with observability being the discriminant between the two, as "[a] worker feels shame when others can observe his actions. Without observability, only guilt can be an effective form of pressure" (p.806). These two feelings can also be present in settings with anonymous interaction, as discussed in Ellingsen and Johannesson (2008), and our experimental design allows us to distinguish between the two. The literature also discusses potentially negative effects of peer pressure arising, for instance, from a desire to conform (Bernheim, 1994) or discouragement. For example, Bellemare et al. (2010) find evidence of a negative peer effect in a data entry task where they provide feedback information about the total productivity of another worker in a previous session. In particular, in their fixed wages condition, they find that men decrease productivity

[2] In a recent and related paper, Babcock et al. (2012) compare individual and team incentives in a pay-for-studying experiment. In their setting, there is no free riding, as both team members need to reach a certain performance threshold for a bonus to be paid. The randomly assigned teammate could be either known or anonymous. They find strong positive effects on performance of being assigned to a team with a known partner in the pay-for-study intervention. Individuals assigned to the anonymous team treatment performed about as well as the individual treatment despite a high risk of default by the counterpart. They also find smaller but similar effects in a pay-for-exercise experiment.

when observing either very low or very high performance. Another paper that finds negative peer effects is Eriksson et al. (2009). They study individuals' performance under a piece-rate and a tournament pay scheme. In their control group, subjects receive no feedback, while in the two treatment groups individuals receive feedback on another subject's performance either halfway during the production period or in a continuous way. They find no evidence of positive peer effects in terms of performance, while they find that there is a negative quality peer effect, with feedback inducing more mistakes.[3]

There is, therefore, evidence of both positive and negative forces at play in situations involving peer effects. These forces are also present in our setting, rendering it impossible to make clear-cut predictions as to whether and how productivity will be impacted. Observed subjects might feel ashamed if they put too low effort or might internalize their impact on observers and refrain from working too hard or might even under-perform due to a choking under pressure effect (Baumeister, 1984; Ariely et al., 2009). On the other hand, observers might try to emulate high performers and raise effort or feel discouraged by them and hold back from really applying themselves. Moreover, these various motivations might interact and be amplified or diminished when the compensation structure in place generates externalities across workers. What type of motivation prevails is, therefore, an empirical question, which we address experimentally in this paper. The next section describes our experimental design and procedures. Section 3 reports the results, while the last session offers a discussion and concludes.

2 Experimental Design and Procedures

The experiment was conducted at the Royal Holloway Experimental Lab between January 2012 and June 2013. The participants were graduate and undergraduate students at Royal Holloway, University of London. Subjects were recruited using the ORSEE software (Greiner, 2004). Each session was divided into 16 periods and the experiment lasted approximately 75 minutes.

At the beginning of each session detailed instructions (available in the Appendix) appeared on the computer screens and were read aloud. The instructions provided details about the real effort task at hand and the compensation. In particular, we used the computerized slider task developed

[3]In a recent paper on rank incentives, Barankay (2012) run a randomized control trial with full-time furniture sales people and finds that providing rank feedback without pecuniary consequences has a negative impact on sales performance, also in this case concentrated among males. On the other hand, other studies like Delfgaauw et al. (2013) and Blanes i Vidal and Nossol (2011) find positive effects of providing information about the relative position vis-a-vis peers.

by Gill and Prowse (2012a).[4] The task involves a screen showing 50 sliders, which can be moved using the computer mouse and positioned anywhere between 0-100. Each slider is initially placed at 0 and the objective is to position as many sliders as possible at exactly 50. Each slider had a number to its right showing its current position. A screenshot of the task is provided in figure 1. There was no limit in how many times a slider could be moved. Subjects were instructed that their "points score" in the task would be the number of sliders positioned at exactly 50 at the end of the 90 second period.

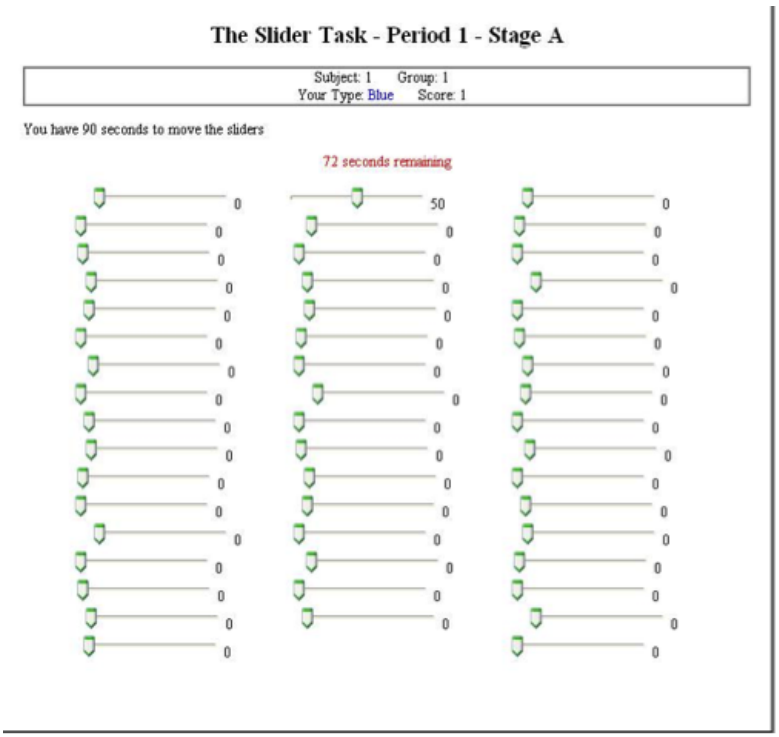

Figure 1: Screenshot of Slider Task.

After the end of the second period new instructions were provided for the second part of the experiment consisting of the remaining 14 periods. In particular, subjects were randomly assigned to one of three treatments: Observed, Observers and Control (we used neutral language in the instructions). From that point on, periods consisted of two stages of 90 seconds each. In stage A subjects in the Observed and Control treatments were working on the task, while the observers were seeing on their screen a copy of the screen of their partner's. In Stage B the observers were performing the task, while the other players were seeing a time-out screen. While subjects were performing the task, they could see their current score, the time remaining and, after session 2,

[4]Gill and Prowse (2011) provide details of how to implement the task and discuss its advantages. This task has been used in many experiments, listed in Gill and Prowse (2012b).

their type. In addition, the observers' screen reported the number of points their partner scored in stage A of the same period. In each period subjects in the Observed and Observers treatments were randomly matched with each other.

Our experimental design also involved two conditions: the piece rate condition and the team compensation condition. In the first condition, subjects received 0.35 pence for each point they scored. In the team compensation condition, in each period - from period 3 onwards - one observer, the subject that she observed and one subject in the control group formed a team. They were informed that their payments for the period would be based on the average performance of their team in that period. No information regarding the identity of their teammates was provided at any point.

To summarize, in the first 2 periods of the experiment all subjects worked in isolation. We use these periods as a gauge of individual ability on the task. From period 3 onwards, we implement a 3 x 2 between-subject design, whereby subjects are either in the role of control, observed or observer and are compensated either on an individual or a team basis. We conducted 10 sessions of the piece rate condition and 7 sessions of the team compensation condition, with a total of 179 participants, 107 and 72 respectively. In particular, in the piece rate condition we had a total of 107 subjects (38 observed, 38 observers and 31 control), while in the team condition we had a total of 72 subjects (24 of each type). At the end of the experiment, one period was randomly selected and the points score in that round was used to determine experimental earnings, which averaged £10 per person.

3 Results

3.1 Descriptive Statistics and Nonparametric Analysis

We begin the presentation of our experimental results with some overview of the data. Average productivity across all periods and treatments was 27.5 and 28.4 sliders in the piece rate and team condition respectively. A histogram of the per period productivity by condition is provided in Figure 2, which illustrates that the bulk of observations lies in the range of 20-35 sliders per period.

We next examine the baseline productivity by treatment, defined as the average productivity in periods 1-2, that is, before assignment into treatment took place. Table 1, columns 1-2, contains this information by treatment and condition. Baseline productivity in the control treatment appears to be slightly lower in both conditions. However, when performing pairwise comparisons between control and each of the two treatments within each condition using a two-sided Mann-Whitney

U-test we find no significant differences, except between observed and control in the piece rate condition (p-value=0.072). There are no significant differences in initial productivity between observed and observer in either condition according to the same test.

We next consider productivity after assignment to treatment - periods 3 onwards. To gain some sense of how productivity on the task evolves, Figures 3 and 4 plot average per-period productivity in the two conditions by treatment. What is evident in these two figures is the steep increase in productivity that takes place in the course of the experiment, in particular in periods 1-7, while subjects seem to reach a plateau in later periods. We compute for every subject a productivity index: the ratio between their mean productivity in the last 14 periods and their mean productivity in the baseline two periods (see columns 3-4 of Table 1). Note that for the control group in the piece rate condition the average productivity improvement is 56%.[5] In the team condition the average productivity improvement is much less dramatic (17%) and the difference between the two control groups is statistically significant (MW-test; p-value=0.005).

For a visualization of the distribution of this productivity improvement by treatment see the box and whisker plots in Figure 5. In the piece rate condition, the difference in productivity improvement between observed and control is significant (MW-test; p-value=0.03). Our interpretation for this difference, when combined with the fact that the observed are somewhat more productive initially, is that subjects in the control catch up in the course of the experiment. In the team condition, productivity improvement is very similar across the three groups, in fact the difference between each of the treatments and the control is not significant (MW-test p-value is 0.52 and 0.36 for observed and observer respectively).

The initial increase in productivity we see in figures 3 and 4 is reflecting the fact that learning is an important determinant of performance on the task. The subsequent flattening out, that is common across treatments, might be indicative that there is a ceiling in subjects' ability to perform the task, which is reached after a few periods of learning. In the last period average productivity in the three treatments is indeed remarkably similar.

The evidence so far is not supporting the presence of any peer effects. However, given that learning appears to be quick and important for this task and may be dominating any treatment effects, it is of interest to examine the initial response to treatments in early periods. We measure the productivity improvement in the early periods by computing a similar productivity ratio as

[5] We note that in this particular group there in an extreme outlier who improves productivity more than nine-fold. In fact, the median productivity increase is a more moderate 27%. If we exclude the outlier, average productivity improvement is 30%.

above, using only the first two periods after assignment to treatment (see Figure 6 and columns 5-6 of Table 1). What we see this time is that the performance of the control group in the piece rate condition improves by 30%, whereas this is not the case in the team condition.[6] This difference is statistically significant (MW-test; p-value=0.02) and suggests that early on some free-riding takes place in the team condition. In order to check for whether assignment into the observed group impacts productivity, we compare productivity improvement across observed and control within each condition. In the piece rate condition we do not find significant differences, whereas, in the team condition, both observed and observers increase productivity by 10% and are significantly different from control (MW-test; p-value is 0.04 and 0.01 respectively).[7]

To summarize, behavior of subjects on the task is characterized by significant learning, with subjects reaching their capacity in later stages of the session. This makes of particular interest to look at early periods, to see whether treatments have an impact on how fast subjects learn. In the piece rate condition we find no evidence of peer effects, either in the short run or the long run. In the team condition, there is some evidence of free-riding taking place in the short run, captured by the fact that the control group in this condition improves performance less than their counterpart group in the piece rate condition. Furthermore, in the team condition subjects in both the observed and observer treatments improve productivity in the short run relative to the control group. As mentioned above, in the long run subjects reach their capacity, so that learning overcomes any treatment effects. We, therefore, conclude that in the team condition, peer pressure leads subjects to overcome the problem of moral hazard in teams and allows subjects to reach their potential faster.

In the next subsection, we look closer at individual behavior using regressions.

3.2 Regression Analysis

We split the presentation of regression results into two parts. We first replicate the nonparametric analysis above and investigate whether there are differences in the productivity of the observed treatment relative to that of the control, for each condition separately. We then turn attention to the behavior of the observers and ask whether their productivity is influenced by the productivity of the observed they are being matched with.

[6] Again, this 30% increase is influenced by the presence of the earlier-mentioned outlier. The median increase is 12.5%, while excluding this subject the average increase is 7.5%.

[7] If we look at productivity improvement beyond period 4, convergence between observed and control takes place. In particular, average productivity improvement in periods 5-6 is 12.6% and 7.5% for observed and control respectively, with the difference not being statistically significant (MW-test; p-value is 0.29).

3.2.1 Behavior of those being observed

Table 2 presents OLS regressions of productivity on a dummy for being in the observed treatment, focusing on the post treatment part of the experiment (periods 3 onwards). We consider two alternative specifications: one in which the dependent variable is the number of sliders, and another one in which the dependent variable is the percentage change in productivity relative to the baseline productivity (average productivity in periods 1 and 2). The second specification accounts for differences in starting ability to perform the task. The regressions also include a full set of unreported period and session dummies. Standard errors are clustered at the individual level.

What we see in column 1 is that for the piece rate condition there is an overall positive but insignificant effect on the productivity of the observed. If we consider changes in productivity relative to the baseline in column (2) the coefficient of being observed becomes negative, but is again not statistically significant. When we focus on the initial response to the treatment, in columns 3-4, we see a positive effect of being observed on the number of sliders, however, there is no statistically significant difference in the reaction of the treated group relative to the control when we consider productivity change in column 4. These regression results thus confirm the lack of a significant overall peer effect on the observed also found in the nonparametric analysis.

Results for the team condition are presented in Table 3. For this condition again there is no evidence of an overall treatment effect. We do see a positive and significant effect if we concentrate in periods 3-4. In those periods, being observed is associated with a productivity improvement that is 11% larger than that of being in the control. Again these regression results confirm the pattern of an initial positive effect on productivity of being observed also highlighted in the previous analysis.

We have also investigated whether being observed leads to a "choking under pressure" effect (Baumeister, 1984; Ariely et al., 2009), whereby a person's performance may deteriorate when scrutinized. To do this we define a measure of the "accuracy" of a subject's effort on the task: the number of sliders positioned at exactly 50 over number of sliders positioned between 48 and 52. Using this measure we find no evidence that the observed's accuracy suffers in either condition.

3.2.2 Behavior of those observing

We next consider the behavior of those assigned to the role of observer. In particular, in table 4 we present OLS regressions of the observer's productivity on a linear and a quadratic term of the matched observed's productivity, plus a full set of period and session dummies. We report

results for two specifications: number of sliders in levels and percentage change relative to baseline productivity. In the piece rate condition there appears to be a significant nonlinear relationship between the productivity of the observed and that of the observer. In particular, the effect of observing somebody whose productivity is higher by 1 slider is -0.14 at the 10th percentile of the observed's productivity, at the median it is -0.03 and at the 90th percentile it is 0.08, from column 1. Thus, the number of sliders completed in one session by observers decreases when observing a low productivity subject and increases when observing a high productivity one. In the team condition, the signs of the coefficients are the same as above, but smaller in magnitude and statistically insignificant. Thus, we find no evidence of any significant relationship between the productivity of the observed and that of the observer in this condition.

The pattern that emerges regarding the reaction of observers' in the piece rate condition is therefore one of responding to extreme observations, with a slight tendency to respond stronger to low values. The next section provides some interpretation for this finding.

4 Discussion and Concluding Remarks

Peer effects arise in situations where workers observe each others' work activity. In this paper we disentangle the role played by observing a peer from that of being observed by a peer by setting up a real effort experiment in which we manipulate the observability of effort. We look at both a piece rate compensation scheme, where pay depends solely on own performance, and a team compensation scheme, where pay also depends on other team members' performance. Overall, we find some evidence that subjects who are observed work harder when compensation is team based, while we find that subjects observing react to what they see in a non-linear but monotonic way when compensation is based only on own performance. An increase in productivity when being observed is consistent with what was postulated by Kandel and Lazear (1992), namely, that subjects work harder to avoid feeling shameful about low performance when others are observing them. The fact that we find this in the team compensation condition, but not in the piece rate one, is also consistent with the idea that peer effects are stronger when compensation across subjects is linked.

In the piece rate condition, we find a non-linear effect of peer pressure on the observer. The impact on productivity is, however, monotonic within the range of our observations, with a negative impact of observing a subject with low productivity and a positive impact of observing a subject with high productivity, while observing a subject with median productivity has essentially no effect.

Bellemare et al. (2010) also find evidence of non-linearity in peer effects, albeit only in their fixed wage condition and not in their piece rate one. Also, in their case the effect is non-monotonic, with the observation of both high and low productivities inducing a negative effect on males. In the conclusions, the authors argue that peer effects may be difficult to find under a piece rate condition as people work at their maximum level of effort and focus their attention on the piece rate, disregarding other information about the performance of peers. One feature of our design that could explain why we do instead find evidence of peer effects in the piece rate condition is that in our setting observers actually see the screen of the observed, while she is working, observing each single movement of the slider. In the Bellemare et al. (2010) setup, instead, subjects are just informed about the other subject's score. Being able to see another subject's work activity, and not just the outcome, may be more conducive to the emergence of peer effects in the lab.

On the other hand, in the team compensation condition we do not find any effect, while, as mentioned above, one could have expected the effect to be even stronger. This could be due to some form of income targeting (Camerer et al., 1997; Fehr and Goette, 2007). Recall that the compensation is calculated on the basis of a randomly selected session. When paired with a high performer, an observer knows that in any case the compensation will be adequate if that session is selected for compensation and this could counteract the heightened feelings of guilt when slacking. On the other hand, when paired with a low performer, the desire to make up for this low performance (and thus insure for herself a minimum level of experimental earnings) could counteract any reduced feeling of guilt (or enhanced feelings of spite) arising from the low performance of the team mate. We should also underline that we simply look at the impact of performance that has just been observed. This is of course a very relevant and prominent measure, particularly in the team compensation condition, where that productivity will contribute to determine pay in case the session is selected. Potentially, however, there could be some persistency, and a subject may be influenced by the whole history of what has been observed so far. To explore this convincingly would require a larger dataset and we leave this for future research.

What we find in this paper is that both observing and being observed by a co-worker matter for work effort, albeit, under different incentive structures. A fruitful avenue for future research would be to consider the size and direction of these peer effects under other compensation schemes, such as, a fixed pay regime and a relative performance scheme.

References

[1] Ariely, D., U. Gneezy, G. Lowenstein, and N. Mazar (2009): "Large Stakes and Big Mistakes," *Review of Economic Studies*, 76, 451-469.

[2] Babcock, P., Bedard, K., Charness, G., Hartman, J., and H. Royer (2012): "Letting Down the Team? Evidence of Social Effects of Team Incentives," Working Paper UCSB.

[3] Barankay, I. (2012): "Rank incentives: Evidence from a randomized workplace experiment," Working Paper, Wharton School, University of Pennsylvania.

[4] Baumeister, R.F., (1984): "Choking under pressure: self-consciousness and paradoxical effects of incentives on skillful performance," *Journal of Personality and Social Psychology*, 46 (3), 610-620.

[5] Bellemare, C., P. Lepage, and B. Shearer (2010): "Peer Pressure, Incentives, and Gender: an Experimental Analysis of Motivation in the Workplace", *Labour Economics*, 17, 276-283.

[6] Bernheim, B. D. (1994): "A Theory of Conformity," *Journal of Political Economy*, 102, 841-877.

[7] Blanes i Vidal, J., and M. Nossol (2011): "Tournaments without Prizes: Evidence from Personnel Records," *Management Science*, 57, 1721-1736.

[8] Camerer, C., L. Babcock, G. Loewenstein, and R. Thaler (1997): "Labor supply of New York City cabdrivers: One day at a time," *Quarterly Journal of Economics*, 112(2), 407-441.

[9] Charness, G., and P. Kuhn (2011): Lab Labor: What Can Labor Economists Learn in the Lab?, *Handbook of Labor Economics*, Volume 4a, 229-330.

[10] Corgnet, B., Hernan-Gonzalez, R., and S. Rassenti (2013): "Peer Pressure and Moral Hazard in Teams: Experimental Evidence," Chapman University Working Paper.

[11] Delfgaauw, J., R. Dur, J. Sol and W. Verbeke (2013): "Tournament incentives in the field: Gender differences in the workplace," *Journal of Labor Economics*, 31(2), 305-326.

[12] Ellingsen, T., and M. Johannesson (2008): "Pride and Prejudice: The Human Side of Incentive Theory," *American Economic Review*, 98(3), 990-1008.

[13] Eriksson, T., A. Poulsen, and M. C. Villeval (2009): "Feedback and Incentives: Experimental Evidence," *Labour Economics*, 16(6), 679-688.

[14] Falk, A., and A. Ichino (2006): "Clean Evidence on Peer Effects," *Journal of Labor Economics*, 24, 39-58.

[15] Fehr, E., and L. Goette (2007): "Do Workers Work More if Wages Are High? Evidence from a Randomized Field Experiment," *American Economic Review*, 97(1), 298-317.

[16] Gill, D., and V. Prowse (2011): "A Novel Computerized Real Effort Task Based on Sliders," IZA Discussion Paper 5801.

[17] Gill, D., and V. Prowse (2012a): "A Structural Analysis of Disappointment Aversion in a Real Effort Competition," *American Economic Review*, 102(1), 469-503.

[18] Gill, D., and V. Prowse (2012b): "Gender Differences and Dynamics in Competition: The Role of Luck," Oxford Department of Economics Discussion Paper 564.

[19] Greiner, B. (2004): An online recruitment system for economic experiments. In Forschung und Wissenschaftliches Rechnen, eds. Kurt Kremer and Volker Macho, 7993. GWDG Bericht 63: Gottingen.

[20] Kandel, E., and E. Lazear (1992): "Peer Pressure and Partnerships," *Journal of Political Economy*, 100, 801-817.

[21] Manski, C. F. (1993): "Identification of Endogenous Social Effects: The Reflection Problem," *Review of Economic Studies*, 60(3), 531-42.

[22] Mas, A., and E. Moretti (2009): "Peers at Work," *American Economic Review*, 99(1), 112-145.

Appendix A - Figures

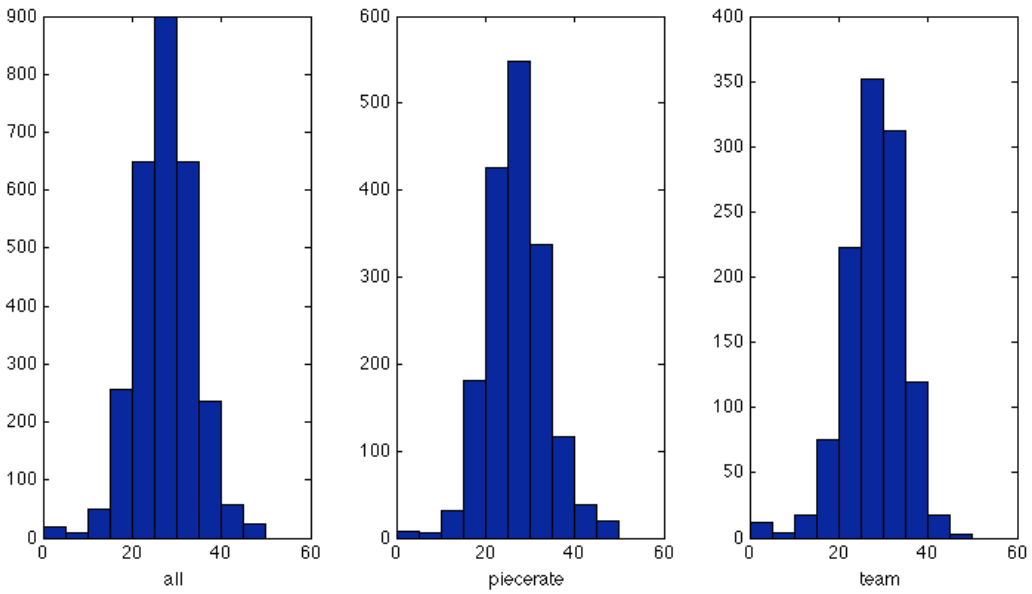

Figure 2: Histogram of productivity by condition

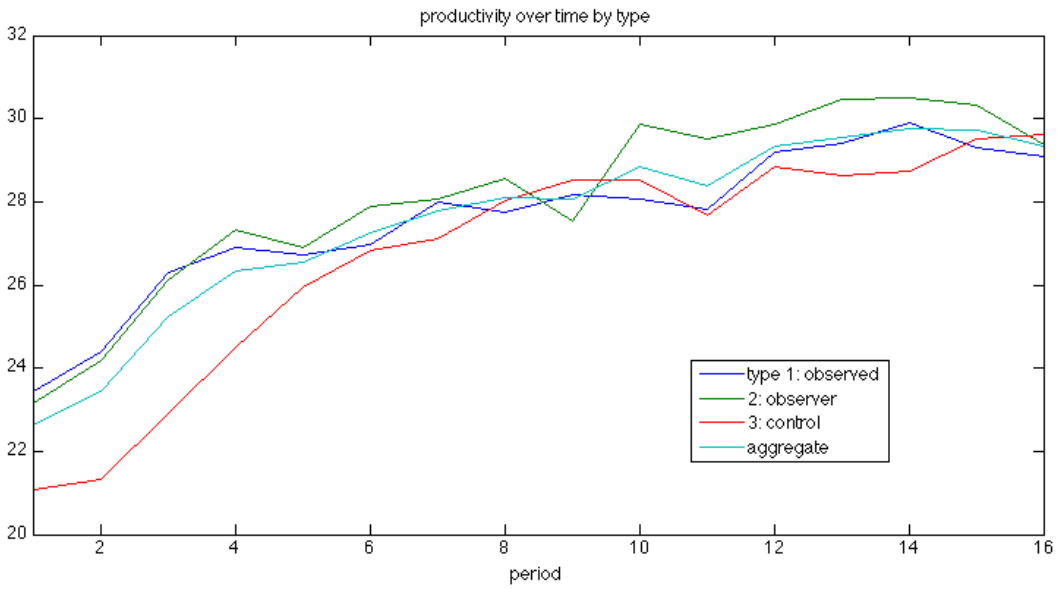

Figure 3: Productivity over time in condition piece rate

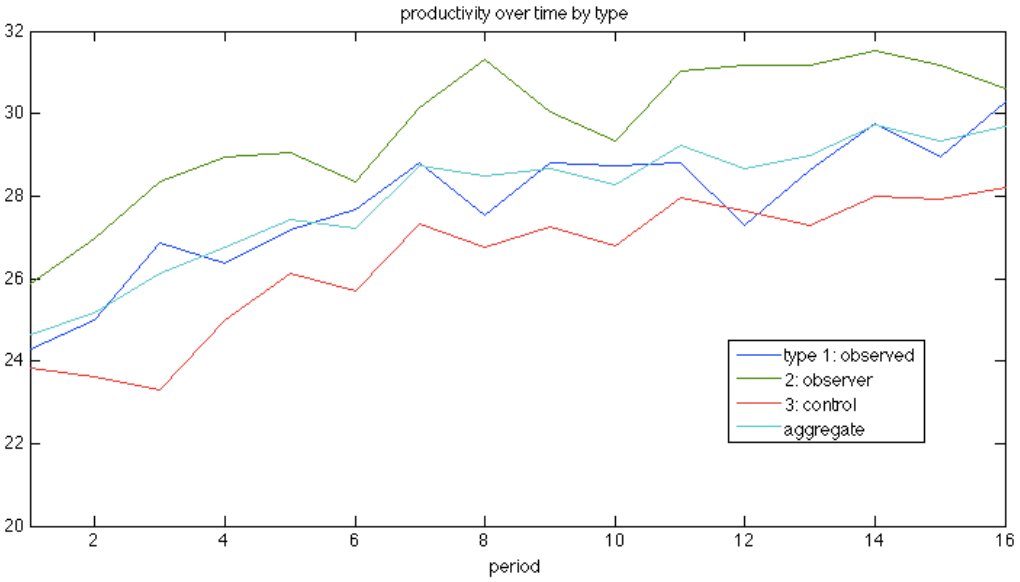

Figure 4: Productivity over time in condition team

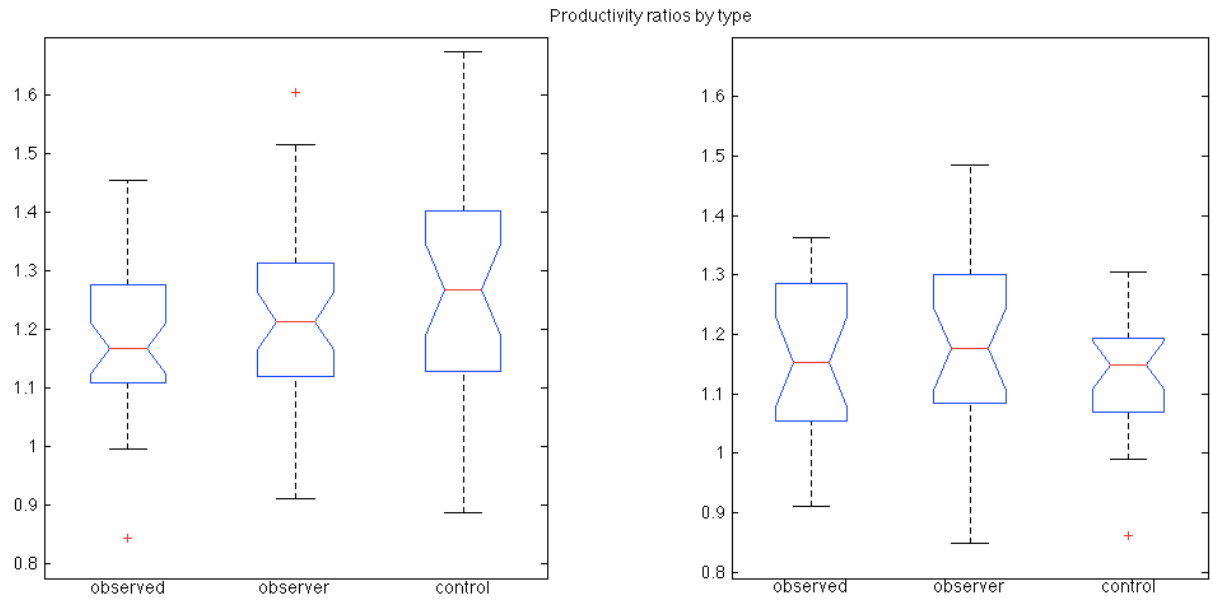

Figure 5: Differences in the productivity ratios. Left panel is treatment piecerate, right panel treatment team. The scale of the y-axis has been restricted.

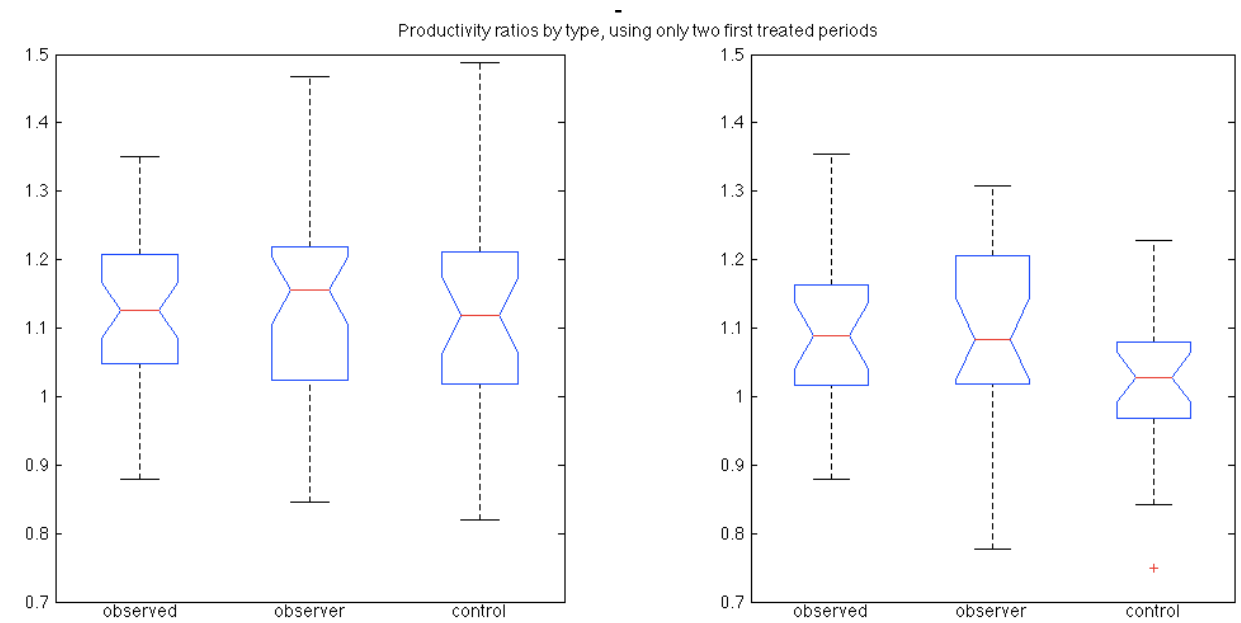

Figure 6: Differences in the productivity ratios, using only the first two treated periods. Left panel is the piecerate condition, right panel condition team. The scale of the y-axis has been restricted.

Appendix B - Tables

Table 1: Summary Statistics of Baseline Productivity and Productivity Improvement

	Baseline Productivity		Productivity Improvement I		Productivity Improvement II	
	Piece Rate	Team	Piece Rate	Team	Piece Rate	Team
	(1)	(2)	(3)	(4)	(5)	(6)
Observed	23.9	25.7	1.18	1.16	1.12	1.09
	(0.57)	(0.97)				
Observer	23.7	26.3	1.23	1.17	1.14	1.10
	(0.59)	(0.69)				
Control	21.2	23.7	1.56	1.17	1.30	0.98
	(0.83)	(0.89)				
Subjects	107	71	107	71	107	71

Notes: Columns (1) and (2) contain average productivity in periods 1-2 (Standard deviations in parentheses). Columns (3) and (4) contain the ratio of average productivity in periods 3-16 over average productivity in periods 1-2. Columns (5) and (6) contain the ratio of average productivity in periods 3-4 over average productivity in periods 1-2.

Table 2: Effect on productivity of being observed: Piece rate Condition

	Periods 3-16		Periods 3-4	
	Sliders	% Change	Sliders	% Change
	(1)	(2)	(3)	(4)
Observed	1.014	-0.333	4.575**	-0.073
	(1.59)	(0.20)	(1.74)	(0.18)
Intercept	28.1***	0.332***	24.9***	0.174
	(3.36)	(0.12)	(3.06)	(0.12)
Obs	966	966	138	138

Notes: *,** and *** denote, respectively, significance at the 10%, 5% and 1% levels. Standard errors clustered at the individual level are in parentheses. A full set of session and period dummies are included in all columns.

Table 3: Effect on productivity of being observed: Team Condition

	Periods 3-16		Periods 3-4	
	Sliders	% Change	Sliders	% Change
	(1)	(2)	(3)	(4)
Observed	2.647	-0.004	3.576*	0.110**
	(1.73)	(0.05)	(1.85)	(0.05)
Intercept	19.913***	0.167	20.07***	-0.141
	(4.60)	(0.13)	(5.20)	(0.19)
Obs	658	658	94	94

Notes: *,** and *** denote, respectively, significance at the 10%, 5% and 1% levels. Standard errors clustered at the individual level are in parentheses. A full set of session and period dummies are included in all columns.

Table 4: Effect of productivity of observed on matched observer

	Piece Rate		Team	
	Sliders	% Change	Sliders	% Change
	(1)	(2))	(3)	(4)
Observed	-0.421***	-0.020***	-0.143	-0.003
Prod.	(0.129)	(0.006)	(0.265)	(0.008)
Observed	0.007***	0.000***	0.002	0.000
Prod.^{2}	(0.002)	(0.000)	(0.004)	(0.000)
Intercept	28.5***	0.328***	27.9***	0.194
	(2.189)	(0.104)	(4.286)	(0.143)
Obs	532	532	322	322
Subjects	38		24	

Notes: *,** and *** denote, respectively, significance at the 10%, 5% and 1% levels. Standard errors clustered at the individual level are in parentheses. A full set of session and period dummies are included in all columns.

Appendix

Experimental Instructions

1 – Initial Instructions

This session consists of 16 rounds. In each round you will undertake an identical task lasting 90 seconds. The task will consist of a screen with 50 sliders. Each slider is initially positioned at 0 and can be moved as far as 100. Each slider has a number to its right showing its current position. You can use the mouse in any way you like to move each slider. You can readjust the position of each slider as many times as you wish. Your "points score" in the task will be the number of sliders positioned at exactly 50 at the end of the 90 seconds. Here is an example of a slider (feel free to play with it)

[SLIDER HERE]

While performing the task, a number of pieces of information will appear at the top of your screen, including the time remaining, the Period number, and your points score in the Period. Your pound earnings for the session are determined as follows: At the end of round 16 the computer will randomly select one of the rounds and use your points score in that round to determine payments. You will be paid 35 pence for each point accumulated in the round that is randomly selected to be implemented. It makes good sense, therefore, to act in each round as though it will actually be carried out.

Please note that communication between participants is strictly prohibited during the experiment. If you have any questions during the experiment please ask the experimenter.

2 – Instructions after period 2 – Piece Rate Condition

This is the Second Stage of the experiment. All players are now randomly separated in three roles:
Blue, Red and Green.

Red players observe the sliders of a randomly chosen Blue player in their screen before they play. This of course means Blue players' moves are being observed.

Green players do not observe and are not observed by anyone.

[CLICK TO CONTINUE BUTTON]

[EXAMPLE OF ROLE ASSIGNMENT:] Your type is Blue. Blue players are observed by Red while moving their sliders

3 – Instructions after period 2 – Team Condition

This is the Second Stage of the experiment. All players are now randomly separated in three roles:
Blue, Red and Green.

Red players observe the sliders of a randomly chosen Blue player in their screen before they play. This of course means Blue players' moves are being observed.

Green players do not observe and are not observed by anyone.

In each round the computer will randomly assign everyone to a team comprising of a Red player, the Blue player that he or she will observe, and one Green player.

In any given round, no information on the identity of the team members or the team performance will be displayed.

At the end of round 16 the computer will randomly select one of the 16 rounds and use that round to determine payments. If it is one of the rounds following the current one, the computer will calculate the AVERAGE points scored by your team. Your payment will be based on that average and, in particular, you will be paid 35 pence for each point.

For instance, if in the round selected for implementation, one team member has scored X points, another Y, and the third Z, the average points score is (X+Y+Z)/3 and you will be paid based on that.

[CLICK TO CONTINUE BUTTON]

[EXAMPLE OF ROLE ASSIGNMENT:] Your type is Blue. Blue players are observed by Red while moving their sliders