DISCUSSION PAPER SERIES

No. 3834

CLEAN EVIDENCE ON PEER PRESSURE

Armin Falk and Andrea Ichino

LABOUR ECONOMICS

Centre for Economic Policy Research

www.cepr.org

Available online at: **www.cepr.org/pubs/dps/DP3834.asp**

ISSN 0265-8003

CLEAN EVIDENCE ON PEER PRESSURE

Armin Falk, Universität Zurich and CEPR
Andrea Ichino, European University Institute, Firenze and CEPR

Discussion Paper No. 3834
March 2003

Centre for Economic Policy Research
90–98 Goswell Rd, London EC1V 7RR, UK
Tel: (44 20) 7878 2900, Fax: (44 20) 7878 2999
Email: cepr@cepr.org, Website: www.cepr.org

CEPR Discussion Paper No. 3834

March 2003

ABSTRACT

Clean Evidence on Peer Pressure*

While confounding factors typically jeopardise the possibility of using observational data to measure peer effects, field experiments offer the possibility of obtaining clean evidence. In this Paper we measure the output of four randomly selected groups of individuals who were asked to fill letters in envelopes, with remuneration completely independent of output. For two of these groups the output of peers was exogenously manipulated (low or high) by making individuals aware of the number of letters previously produced by artificial colleagues. In the third group individuals were set up to work one in front of the other, while the fourth group gave the baseline output for independent not manipulated work. Our first finding is that effort of the less productive workers reacts in a sizeable and statistically significant way to peer pressure. Second, there is strong evidence of peer effects when individuals work in pairs. Third, these peer effects work in the direction of making the least productive individuals work harder, thereby increasing overall productivity.

JEL Classification: D20, J20 and K40
Keywords: field experiments, incentives and peer effects

Armin Falk
Institut für Empirische Wirtschaftsfors
Universität Zürich
Blümlisalpstrasse 10
CH-8006 Zürich
SWITZERLAND
Tel: (41 1) 634 3704
Fax: (41 1) 634 4907
Email: falk@iew.unizh.ch

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=146697

Andrea Ichino
Department of Economics
European University Institute
Badia Fiesolana
I-50016 S. Domenico di Fiesole (FI)
ITALY
Tel: (39 055) 4685 922
Fax: (39 055) 4685 902
Email: andrea.ichino@iue.it

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=111104

*We would like to thank Urs Fischbacher, Michel Kosfield and Gianfranco Pieretti for useful conversations. Jasmin Gülden and Barbara Nabold provided excellent research assistance.

Submitted 20 February 2003

1 Introduction

Whether individual behavior is significantly affected by peer pressure is a question that scholars in many disciplines have since long tried to answer with limited success. The reason why answering this question is difficult, despite the apparent wealth of evidence from daily experience, is that observational data do not allow to easily separate the pure effect of peer's behavior from the effect of confounding factors. Using data from a controlled field experiment in which randomly selected subjects were paid independently of their work output, we show in this paper that the productivity of a worker is systematically influenced by the productivity of peers in the absence of confounding factors. These results provide clean evidence for the existence of peer pressure on work behavior.

In order to understand the nature of our experiment, consider two individuals working on separate tasks but one in sight of the other. Suppose that we observe them behaving in a similar way, which we suspect could be generated by peer pressure. To be precise, we define peer pressure as a situation in which the output of individual i changes when the output of j is exogenously perturbed and nothing else changes. Following Manski [1993], a first set of confounding factors is generated by the possibility that local attributes of the environments in which the two individuals operate determine their behavior. If observational data do not allow to fully control for these local attributes, even in the absence of true peer pressure we could observe the behavior of i and j changing simultaneously simply because some unobserved local attributes have changed. Second, it is possible that the two individuals have similar characteristics, which would make them behave similarly even if they were not working one in sight of the other. With respect to both these possibilities, it could also happen that i and j decide to work nearby because they like the same local attribute, which in turn affects their behavior, or because they both like to be near individuals with similar characteristics. In these cases, the supposed effect of peer pressure would instead be the result of sorting according to local or personal attributes. Finally, if i affects j and vice versa, it is in general difficult to disentangle empirically this "reflection problem" and to separate out the actual causal effects of each worker on the other.

The most recent generation of studies which try to measure peer effects with observational data has made several important steps in the direction of solving these problems.[1] However, in many of these studies, even if the

[1] See, among others, Wilson [1987], Case and Katz [1991], Crane [1991], Glaeser et al. [1996], Topa [1997], Encinosa et al. [1998], Aaronson [1998], Van Den Berg [1998], Bertrand et al. [2000], Ichino and Maggi [2000], Katz, Kling and Liebman [2001] and

setting offers an almost perfect opportunity to identify peer effects, the impossibility to fully control for local or personal confounding factors and for endogenous sorting makes the identification strategy not fully convincing. The most significant recent steps forward in this literature are offered by Sacerdote [2001] and Katz, Kling and Liebman [2001] who use data based on randomized assignments of individuals to peer groups. However, both these papers have still to confront the "reflection problem" and the consequences of local confounding factors. More specifically, Sacerdote [2001] finds evidence of peer effects among Dartmouth students randomly assigned to the same dorm but cannot exclude convincingly the possibility that these effects might be due to local time varying shocks. This is less of a problem in Katz, Kling and Liebman who analyse the consequences of changing randomly the residential neighborhood of families residing in high-poverty public housing projects and, therefore, are not primarily interested in isolating pure peer pressure from local effects. A further important difference with respect to our setting, is that in both these papers the attention is not directed to a work environment.

In our study, instead, we focus explicitly on a work environment and we aim at assessing the existence of peer effects in a fully controlled setting in which no possible confounding factor can hinder the possibility of this assessment. As in any other controlled experiment, the possibility to obtain clean evidence offers the possibility to complement in an informative way the evidence generated by observational studies.[2]

Our subjects were recruited randomly to perform a typical side-job, which was paid independently of individual or team output. The work task was to fill letters into envelopes. In total we study four treatments. In our baseline treatment subjects work alone. Output in this treatment reveals the level of productivity in the absence of any peer pressure. In the other three treatments peer pressure was possible. In the second and third treatment we exogenously manipulated the output of peers by making subjects aware of the number of letters previously completed by artificial others. This number was either low or high. Note, in particular, that precisely because of this exogenous manipulation of the behavior of peers the evidence offered by these two treatments is not subject to the "reflection problem" highlighted by Manski [1993]. In the final treatment subjects worked in pairs, i.e., the setting allowed for the possibility that the behavior of subjects were influenced by the behavior of the other member of the couple.

Sacerdote [2001]. See also the literature based on the classic Hawthorne experiments (e.g., Whithead [1938] and more recently Jones [1990]).

[2]For related literature on laboratory experiments aimed at measuring peer effects see Falk and Fischbacher [2002] and Falk, Fischbacher, and Gächter [2002].

Our main findings are the following: First, output reacts in a significant way with respect to the exogenous introduction of peer pressure. When the output of artificial others is low (high), subjects' output is lower (higher) compared to the output in the baseline treatment. Second, there is strong evidence for peer pressure when subjects work in pairs. This can be inferred from the fact that output within pairs is very similar while output differs substantially between pairs. By comparing the standard deviation of output within and between pairs we show that these peer effects are highly significant. Third, the peer effects observed in our study are mainly improving the output of the less productive subjects, with a positive effect on total output.

In the next Section we present the design of our experiment. Section 3 discusses the behavioral hypotheses. Section 4 contains our results. Section 5 concludes.

2 Design of the experiment

The goal of this paper is to study potential peer pressure effects on work behavior. We therefore conducted a field experiment where subjects who performed a simple task in a highly controlled environment were exogenously sorted into different treatments. Before discussing our treatments in detail, we describe the recruitment process, the work task and the procedures.

2.1 Recruitment

All our subjects were high-school students who were recruited from different schools in the area of Winterthur, a city in the canton of Zurich (Switzerland). In announcements posted on blackboards students were asked whether they wanted to do a simple side job, which requires no previous knowledge. In the announcement it was stated that the job was a one-time four hour job, which was paid 90 Swiss Francs (1 Swiss Franc $\approx$ .70 US or $\approx$ .70 EURO). This was obviously an attractive payment since within 24 hours, we were able to recruit all the 40 subjects we had planned to recruit. Students applied by email. After receiving their application we informed them about the precise date and location where they were expected to do the job. The experiment took place during the spring vacations of 2002, which cover two weeks. It was performed in a high-school building in Winterthur.

2.2 Procedure and task

When a subject arrived to do the job, he or she was welcomed and carefully informed about the task and the procedural details. Subjects were told that they had to work for four hours without a break (except of course if they had to go to the bathroom) and that at the end of this time they would receive their payment.

We wanted to implement a work task, which is simple, requires no previous knowledge and which can be easily measured. Therefore students had to fill letters into envelopes and they were told that the purpose of their work was to prepare a questionnaire study. The cubicle of a subject was a simple desk. On this desk we put the material necessary to perform the task: staples of two different sheets of paper, a staple of gummed envelopes, rubber bands, a stamp, a rubber-stamp pad, pencils, a sponge and A-priority-stickers (Figure 1 displays a picture of a subject's cubicle).

The concrete steps of their task were as follows: First subjects had to stamp and number both different sheets of paper. Second they had to put both sheets into one envelope and to close the envelope. Third they had to number the envelope and to put an A-priority sticker on it (see Figure 2 for a picture of the two letters and the envelope). When 25 envelopes had been completed they had to be tied up with a rubber band and put in a box. Each subject's cubicle always looked exactly the same, including, e.g., the same type of desk and chair and the same large number of envelopes and sheets (much more than a single person could ever complete). Payment was explicitly independent on output and paid in cash. Across all treatments the whole procedure was exactly the same.

2.3 Treatments

In total we study four treatments (see Table 1). The first treatment is our "base" treatment. It serves as a control for all other treatments. In this treatment a single subject was working in a room without any contact with another subject. There was also no information about other subjects' output. The base treatment therefore rules out any potential peer effect that could come from a co-worker.[3]

In the "low" and the "high" treatment everything was exactly as in the base treatment except that subjects were given some information about the

[3]Of course we cannot rule out that a subject in the base treatment felt some pressure from the experimenters. Note, however, that this potential effect holds equally in *all* conditions. Moreover the payment was explicitly independent of output and we tried to keep the interaction between experimenters and subjects as neutral and limited as possible.

output of three previous students. We provided this information in each condition by putting three staples of completed envelopes in a transparent box, which was placed in a way such that subjects could easily see it. The subjects were told that this was the output of three previous students who had done the exact same job (Figure 3 displays this box). In the low treatment this output was low while in the high treatment it was high. In order to determine a low and a high output respectively, we first collected output data from some subjects who participated in the base treatment. For the low treatment we took the lowest output and subtracted 70 envelopes. Likewise in the high treatment we added 70 envelopes to the highest output. This calculation results in an output of about 90 in the low treatment and 326 in the high treatment. The boxes presented to the subjects actually contained staples equal to 106, 90 and 75 in the low treatment (with an average equal 90) and 339, 330 and 310 in the high treatment (with an average equal 326).

The fourth treatment is called the "pairs" treatment. This treatment was exactly as the base treatment except that now two subjects worked at the same time in the same room. The two cubicles were standing next to each other such that subjects could easily realize the output of the other student (the position of the second cubicle can be seen in the background of Figure 1). We did not impose any restriction on their interaction, i.e., they were free to communicate. However they were not allowed to engage in teamwork or some sort of division of labor, i.e., they were told that both of them had to do the task described above. To minimize the possibility that two subjects in the pairs treatment knew each other we invited only students from different high-schools to participate in this treatment.

In total 40 subjects participated in our study, eight subjects in each of the base, the low and the high treatment and 16 subjects (eight pairs) in the pairs treatment. The subjects were randomly allocated to the treatments. No subject participated in more than one treatment.

From a methodological point of view some aspects about the design are worth pointing out: Unlike most lab experiments that study work behavior, our subjects performed a 'real' task. In a typical lab experiment the choice of work effort is represented by an increasing monetary function, i.e., instead of choosing real effort subjects choose a costly number. This procedure has been used in tournament experiments, e.g., Bull, Schotter and Weigelt [1987], or in efficiency wage experiments, e.g., Fehr, Kirchsteiger and Riedl [1993]. Recently some authors have conducted so-called "real effort" experiments to study incentive mechanisms and efficiency wages. In these experiments subjects have to do a real task instead of choosing a number. In Fahr and Irlenbusch [2000] subjects had to crack walnuts, in van Dijk et al. [2001] subjects performed cognitively demanding tasks on the computer (two-variable optimization problems) and in Gneezy [2002] subjects had to solve mazes at the computer. However, at least in the latter two studies the task is not perceived as an economically valuable output, which means that an important dimension of usually performed work is missing. Furthermore, in all these studies subjects were well aware of the fact that they acted in an experimental environment. Our study is unique insofar as subjects performed a task that is realized as a regular and economically valuable job. Moreover, we never mentioned the expression “experiment”. Subjects applied for a typical side-job and were paid according to the announced conditions. They were never aware of the fact that we had a scientific interest in their work output.

3 Behavioral hypotheses

In the absence of peer pressure the distribution of completed envelopes (i.e, the output) should be the same across treatments. This is so because the economic incentives are identical across conditions: Each subject receives 90 Swiss Francs for four hours of work independent of output. Of course there might be individual differences because some subjects are, e.g., more talented than others or feel more obliged than others to perform well. But since subjects are randomly allocated to the different treatment conditions individual differences should cancel out across conditions.

To illustrate what we expect to happen in the presence of peer pressure we present in this section a slightly modified version of the model proposed by Ichino and Maggi (2000). Consider a generic subject i of our experiment who chooses a level of output denoted by $X_i \in [0, X^{\max}]$. We assume that the gain from producing X_i is given by $G(X_i, Y^e, \theta_i)$, with $G_1 > 0$ and $G_{11} < 0$, where θ_i is a preference parameter (the subject's “type”) and Y^e is a vector of characteristics describing the local environment e in which i operates. A higher value of θ_i indicates a worker with a higher marginal gain from producing X_i. This amounts to assuming $G_{13} > 0$.

We introduce the possibility of peer pressure by modelling the cost of producing X_i as given by $L(X_i, \bar{X}_i^e)$, where $\bar{X}_i^e$ is the average output of peers in the environment in which i operates. In the absence of peer pressure $L_{12} = 0$, while if instead $L_{12} \leq 0$ there is peer pressure in the sense that the cost of producing X_i is lower when average production is higher. This specification is on purpose rather general since our goal is not to understand the determinants of peer pressure, but just to understand what we should see in the data generated by our experiment if peer pressure exists. By allowing L to depend on $\bar{X}_i^e$ we suggest the possibility that, for whatever reason, it may be costly for a subject not to keep effort in line with what the others do.[4]

Considering a team of subjects like i we can characterize the Nash equilibria of this game. The first step is to derive an individual subject's optimal choice given the other subjects' choices. Each subject chooses X_i to maximize the individual utility of production,

$$U^i = G(X_i, Y^e, \theta_i) - L(X_i, \bar{X}_i^e)$$

Therefore, the optimal output level will be a function of $\bar{X}_i^e$, Y^e and θ_i:

$$X_i = g(\bar{X}_i^e, Y^e, \theta_i). \tag{1}$$

Given our assumptions, we have $\partial X_i / \partial \bar{X}_i^e \geq 0$ with strict inequality if peer pressure exists, and $\partial X_i / \partial \theta_i > 0$. Note that equation (1) is a structural condition, because $\bar{X}_i^e$ is typically endogenous in a real setting. In our experiment, however, it is endogenous in the pairs treatment but exogenous in the other treatments in which the behavior of peers is controlled by us.

Using (1) and denoting with $f^e(\theta)$ the distribution of types in the environment e, we can write

$$\bar{X}_i^e = \int g(\bar{X}_i^e, Y^e, \theta) df^e(\theta). \tag{2}$$

The solutions of this equation in $\bar{X}_i^e$ represent the equilibrium average output levels. Note that, if g is linear, there is a unique equilibrium, but if g is nonlinear, multiple equilibria are possible.

Coming closer to the setup of our experiment, consider for simplicity a linearized version of equation (1):

$$X_i = Y + \beta \bar{X}_i^e + \theta_i \tag{3}$$

where e now denotes the four treatments $\{b, l, h, p\}$ for, respectively base, low, high and pairs. Note that the environment is constant for all subjects and treatments, so $Y_i^e = Y$. Moreover, random assignment of subjects to treatments ensures that $E\{\theta_i \mid e\} = 0$. Finally, given our assumptions, if peer pressure exists then $\beta > 0$.

In the low and in the high treatment $\bar{X}_i^e$ is controlled exogenously by us.[5] It is therefore straightforward to detect the existence of peer pressure on the basis of the following proposition, which will be tested in Section 4.

[4]For a discussions of possible determinants of peer pressure see, among others, Kandel and Lazear [1992], Akerlof [1997] and Huck, Kübler and Weibull [2002]. Note also that, for what is needed in this paper, we could have introduced the effect of peers in the gain function G, but, once again, we are interested in modelling just the effects of peer pressure, not its determinants.

[5]This is indeed a major advantage of our setting with respect to the existing literature in which the estimation of equations like (3) requires solutions to the problem of the endogeneity of $\bar{X}_i^e$.

Proposition 1 *Comparing the low and the high treatment, if peer pressure exists and $\bar{X}^l < \bar{X}^h$, then*

$$E\{X_i \mid e = l\} < E\{X_i \mid e = h\}.$$

In the pairs treatment, for each couple of subjects i and j we have that, respectively, $\bar{X}_i^e = X_j$ and $\bar{X}_j^e = X_i$. In this case the behavior of peers is evidently endogenous. However, it is easy to see that within each pair the difference between the outputs of the two subjects is equal to

$$X_i - X_j = \frac{\theta_i - \theta_j}{1 + \beta}, \tag{4}$$

while the average output of each pair is determined as one of the possible multiple equilibria implied by equation 2.

As a result, peer pressure can be detected in the pairs treatment according to the following proposition.

Proposition 2 *In the pairs treatment, the higher is the effect of peer pressure (i.e., the larger is β), the smaller is the difference of the output levels within pairs.*

In Section 4 we test also this second prediction. Note that in the pairs treatment the existence of peer pressure does not predict a particular level of output given the possibility of multiple equlibria implied by equation 2. All we hypothesize is that output levels should be similar within pairs but it is not clear whether subjects 'coordinate' on a rather high or low level. Output in the pairs treatment may therefore be lower or higher compared to the base treatment.

4 Results

In this section we present our results and test our behavioral predictions. Figure 4 shows the average number of letters produced in each of the four treatments described in the previous section and it is immediate to see that individuals react to the behavior of peers. The lowest average output (185 letters) is observed in the low treatment in which individuals are confronted with an artificially manipulated behavior of peers equal to 90 letters. In the base treatment, in which individuals work independently and without the constraint of any perceived norm, the output is only slightly higher (190 letters). There is instead a clear upward jump in the average output of the remaining two treatments. When the behavior of peers is set artificially to the high level of 327 letters, individuals fill in 217 envelopes and when, instead, they work in pairs we observe the highest average output of 221 letters.

The difference between the average output of the pairs treatment and of the low treatment is not only sizeable in percentage terms (almost 20%) but also statistically significant despite the very small sample size. This is shown in Table 2, which reports the probability (p-value) that the observed average outputs of two treatments differ when the true average outputs are instead identical (null hypothesis).[6] The cell corresponding to the pairs and low treatments indicates that, having observed average outputs of 221 and 190 letters respectively, the probability of no true difference is 3.4%. The p-values for the differences between the high and low treatments and between the pairs and base treatments are higher (9.7% and 7.3% respectively) but, given the small sample, still indicative of the likelihood that outputs really differ across treatments. Note that the average output difference between the high and the low treatments amounts to a sizeable 17%, which supports the prediction of Proposition 1.

There is no reason to expect that peer pressure should influence in the same way individuals with different levels of productivity. This is indeed the result shown in Figure 5, which plots the 10th, 25th, 50th, 75th and 90th quantiles of the output distribution for each treatment. While the output of the 90th quantile is basically unaffected by the different treatments, the output of the 10th quantile changes considerably, reaching an increase of almost 80% between the low and high treatments (this difference is statistically significant with a p-value of 0.005). More generally the sensitivity to peer pressure appears to decrease monotonically going from lower to higher quantiles of the output distribution. As a result output variability changes across treatments displaying the lowest value when peers output is artificially set to the high level. The evidence presented in this figure clearly suggests that low productivity workers are significantly more sensitive to the behavior of peers.

While the comparison of the low and the high treatments suggests that the exogenous manipulation of peer behavior may have substantial effects on the effort exerted by less productive workers, the pairs treatment offers information on the endogenous formation of peer pressure and on its effects. Consider the standard deviation of output within and between pairs. If working in pairs has no effect on individual behavior these standard deviations should be identical to the ones generated by any random configuration of pairs constructed from the same group of people. Moreover, there should be no reason to expect that the between and within standard deviations ob-

[6]The tests are based on regressions of output on treatment dummies.

tained with the true pairs should differ in any specific direction. Therefore, by comparing the standard deviations generated by the true pairs of our experiment with the standard deviations generated by a random set of hypothetical configurations of pairs we can construct a test for the endogenous formation of peer pressure. This comparison is shown in Figures 6, 7 and 8.

The first of these figures plots the kernel density of the hypothetical within pairs standard deviations computed for 20,271 randomly chosen different configurations of pairs of the 16 individuals involved in the pairs treatment. To be more precise, we generated all the 2,027,025 possible configurations of 8 pairs with these 16 individuals[7] and for 1 every 100 of these configurations we computed the within pairs standard deviation.[8]

The range of variation of these hypothetical within standard deviations goes from 9.6 to 34.8 letters. The vertical line in Figure 6 identifies the standard deviation within true pairs, i.e., the one computed for the pairs who actually worked together in our experiment. This standard deviation is equal to 14.6 letters and only 1.17% of the hypothetical configurations originated a lower value. As predicted by Proposition 2, this evidence suggests that on average the output levels of two individuals working in the same room on separate tasks, are significantly more similar than the output levels of two individuals working separately. Another way to say it, is that the probability (p-value) of observing a within-pairs deviation as low as 14.6 in the absence of any peer pressure is less than 1.17%. Hence, we can reject with a high level of confidence the hypothesis of absence of endogenous peer effects.

Given Figure 6, it is not surprising to find, in Figure 7, that the observed standard deviation between the true pairs in the experiment (which is equal to 33.7 letters) is higher than 98.85% of the between standard deviations generated by the hypothetical configurations of pairs. Also in this case, the chance (p-value) that such a high between standard deviation could be generated in the absence of peer pressure is extremely low (in particular smaller than 1.15%). Moreover, Figure 8 plots the kernel density of the between minus within difference for each hypothetical configuration of pairs and it is evident that this difference is not systematically positive or negative since it is approximately symmetric around zero. Note that this is exactly what one would expect in the absence of peer effects, while in the presence of these effects the between standard deviation should be larger than the

[7] It is easy to calculate that this number of configurations is in general equal to $\prod_{i=0}^{(N-2)/2}(N-2i-1)$, where N is the (even) number of individuals, i.e. 16 in our case.

[8] We would have liked to compute the within pairs standard deviations for all the 2,027,025 configurations but this calculation would have required approximately 3,400 hours at the speed of our workstation without any major gain from the viewpoint of the reliability of our results.

within. This is indeed what we find for the true pairs of our experiment: the between minus within difference is equal to 19.0 letters, as indicated by the vertical line in the figure. For only less than 1.17% of the hypothetical configurations the analogous difference reaches a higher value. Hence, while in the absence of peer effects there would be no reason to expect the within standard deviation to be smaller than the between standard deviation or vice versa, Figure 8 suggests that when two individuals are in the same room the between pairs deviation is significantly larger than the within pairs deviation. This means that, ceteris paribus, working in pairs induces more similar output levels than working separately.[9]

The evidence presented so far strongly supports the existence of endogenous peer effects, but in principle, as far as average productivity is concerned, these effects could go both ways, i.e., they could generate vicious or virtuous interactions. In our specific case, there is no reason to expect that within each pair either the potentially most productive worker should slow down to the pace of the potentially least productive worker, or that the opposite should happen. However, if in the light of this question we go back to Figures 1 and 2, the evidence offered by our experiment does not leave space to doubts. The average output of the pairs treatment is the highest because the lowest quantiles of the output distribution are higher in this treatment. In other words, "bad apples", far from damaging "good apples", seem instead to gain in quality when paired with these latter. At least in this sample, peer effects work in the direction of raising significantly the overall average productivity, by increasing the output of the least productive workers[10].

[9] A potential problem in figures 6, 7 and 8 is that the random combinations of pairs were constructed with the same individuals who interacted in the pairs treatment. However, results do not change, qualitatively, when the 8 individuals who participated in the base treatment are used instead. Note that these individuals worked independently without being influenced by any artificially manipulated norm. In this case there are 105 possible configuration of 4 pairs with 8 individuals (see footnote 7), and only one of these configurations originated an hypothetical within standard deviation lower than the one obtained with the true pairs of the pairs treatment. The other results and figures based on this comparison are omitted to save space but are readily available from the authors.

[10] The finding that output in the pairs treatment exceeds output in the base treatment is in line with the psychological paradigm of 'social facilitation' (see Zajonc [1965]): There is a large body of evidence suggesting that the sheer presence of others improves performance of simple tasks. This holds for human as well as for animal behavior. In Allport [1920], e.g., performance of subjects doing simple tasks (like chain word association) was much better in groups than if subjects did the tasks alone. In a more recent study, Towler [1986] takes the time cars need to reach a 100-yd mark from a standing start at traffic lights. He reports that the if there are two cars at the traffic light the time to travel the 100 yds. is significantly shorter than if there is just one car.

5 Summary

In this paper we have presented clear and unambiguous evidence in favor of peer pressure effects. In a controlled field experiment where subjects are randomly allocated to different treatments we find that the exogenous variation of peer output systematically changes work behavior of subjects. We also show that the less productive workers are those who react more significantly to peer pressure. As a result, in our sample peer effects work in the direction of raising significantly the overall average productivity. This raises the interesting question of how to optimally allocate low and high productivity workers. In the light of our results the output maximizing strategy might be to group low and high producitivity workers instead of grouping workers of similar productivity.

As a second test we show that behavior in pairs is strongly affected by peer effects. The standard deviations within groups are significantly smaller than between groups. We also find that work output is highest in the pairs condition, a finding, which is in line with the so-called social facilitation paradigm in social psychology.

Note that in our study the presence of peer effects is robust and quantitatively important even though subjects interacted only once and did not know each other. This suggests the possibility that the effects measured in our study are a lower bound for the effects that prevail in actual labor relations. In contrast with this conclusion, however, it can also be argued that a setting of repeated interactions on a longer horizon might generate effects which cannot be easily predicted on the basis of our evidence. For example, while in the short run the least productive workers seem to react to the higher producivity of their peers, in the long run the opposite might be true if it becomes clear that, as in our setting, low levels of productivity do not have consequences on rewards. To shed light on these issues, the next step on our research agenda is to collect evidence on peer pressure when interaction between peers is repeated over longer horizons.

References

Aaronson, Daniel, "Using Sibling Data to Estimate the Impact of Neighborhoods on Children's Educational Outcomes," *Journal of Human Resources*, XXIII (1998), 915–946.

Akerlof, G., "Social Distance and Social Decisions." Econometrica 65(4), (1997), 1005-1027.

Allport, F.H., "The influence of the group upon association and thought," Journal of Experimental Psychology 3 (1920), 159-182.

Bertrand, Marianne, Erzo F.P. Luttmer and Sendhil Mullainathan, "Network Effects and Welfare Cultures," *Quarterly Journal of Economics* 115 (2000), 1019-1055.

Bull, C., Schotter, A. and Weigelt, K, Tournaments and Piece Rates: An ExperimentaStudy, Journal of Politiclal Economy 95 (1987), 1-33.

Case, Anne, and Lawrence Katz, "The Company You Keep: the Effect of Family and Neighborhood on Disadvantaged Youth," NBER Working Paper 3705 (1991).

Crane, Jonathan, "The Epidemic Theory of Ghettos and Neighborhood Effects on Dropping Out and Teenage Childbearing," *American Journal of Sociology* XCVI (1991), 1226–1259.

Encinosa, William E., Martin Gaynor and James B. Rebitzer, "The Sociology of Groups and the Economics of Incentives: Theory and Evidemce on Compensation Systems," Case Western Reserve University (1998).

Fahr, René, and Bernd Irlenbusch, "Fairness as a constraint on trust in reciprocity: earned property rights in a reciprocal exchange experiment," Economics Letters 66 (2000), 275-282.

Falk A. and U. Fischbacher, "Crime in the Lab-detecting social interaction" *European Economic Review* 46 (2000), 859-869.

Falk, A., Fischbacher, U. and Gächter, S. "Isolating Social Interaction Effects - An Experimental Investigation", IEW Working Paper, University of Zurich. (2002)

Fehr, E., Kirchsteiger, G. and Riedl, A., "Does fairness prevent market clearing? An experimental investigation", *Quarterly Journal of Economics* 108 (2), (1993)437-459.

Glaeser, Edward L., Bruce Sacerdote and Jose A. Scheinkman, "Crime and Social Interactions," *Quarterly Journal of Economics* CXI (1996), 507–548.

Huck, Steffen, Dorothea Kubler and Jörgen Weibull, "Social Norms and Incentives in Firms", (2002), mimeo

Hunt, Peter J., and Joseph M. Hillery, "Social facilitation in a coaction setting: An examination of the effects over learning trials", Journal of Experimental Social Psychology 9 (1973), 563-571.

Kandel, E. and Lazear, E. P. (1992): "Peer Pressure and Partnerships", Journal of Political Economy 100 (4), 801-817.

Ichino A. and G. Maggi, "Work Environment and Individual Background: Explaining Regional Shirking Differentials in a Large Italian Firm" *Quarterly Journal of Economics* 115 (2000), 1057-1090.

Jones, Stephen R.G., "Worker Interdependence and Output: The Hawthorne Studies Reevaluated," *American Sociological Review* LV (1990), 176–190.

Katz Lawrence, Jeffrey R. Kling and Jeffrey B. Liebman, " Moving to Opportunities in Boston: Early results of a Randomized Mobility Experiment." Quarterly Journal of Economics (2001) 116(2), 607-654.

Manski, Charles F., "Identification of Endogenous Social Effects: The Reflection Problem, *Review of Economic Studies* LX (1993), 531–542.

Sacerdote, Bruce "Peer Effects with Random Assignment: Results for Dartmouth Roommates." Quarterly Journal of Economics (2001) 116(2), 681-704.

Topa, Giorgio., "Social Interactions, Local Spill–overs and Unemployment," Working Paper, New York University (1997).

Towler, Gary, "From zero to one hundred: Coaction in a natural setting," Perceptual and Motor Skills 62 (1986), 377-378.

Van den Berg, Gerard, Bas Van der Klaauw and Jan van Ours, "The Effect of Neighborhood Characteristics on the Labor Supply of Welfare Recipients," CentER, Tilburg, mimeo (1998).

Van Dijk, Frans, Joep Sonnemans and Frans van Winden, "Incentive systems in a real effort experiment," European Economic Review 45 (2001), 187-214.

Whitehead, Thomas N., *The Industrial Worker: A Statistical Study of Human Relations in a Group of Manual Workers* (Cambridge, MA: Harvard University Press, 1938).

Wilson, William J., *The Truly Disadvantaged* (Chicago, IL: The University of Chicago Press, 1987).

Zajonc, Robert B., "Social Facilitation," Science 149 (1965), 269-274

Table 1: Treatments

Treatment	Number of subjects in each room	Information displayed
Base	1	none
Low	1	low output of three previous workers
High	1	high output of three previous workers
Pair	2	none (endogenous by output of other)

Table 2: Output differences between treatments

	Baseline	Low	High	Pairs	Output
Baseline	1				190
Low	0.757	1			185
High	0.172	0.097	1		217
Pairs	0.073	0.034	0.811	1	221
Output	190	185	217	221	

Note: The table reports the probability (p-value) that the observed average outputs of two treatments differ when the true average outputs are instead identical (null hypothesis). The tests are based on regressions of output on treatment dummies. There are 8 couples of players in the pairs treatment; in each of the other treatments there are 8 players.

Fig. 1: The cubicle

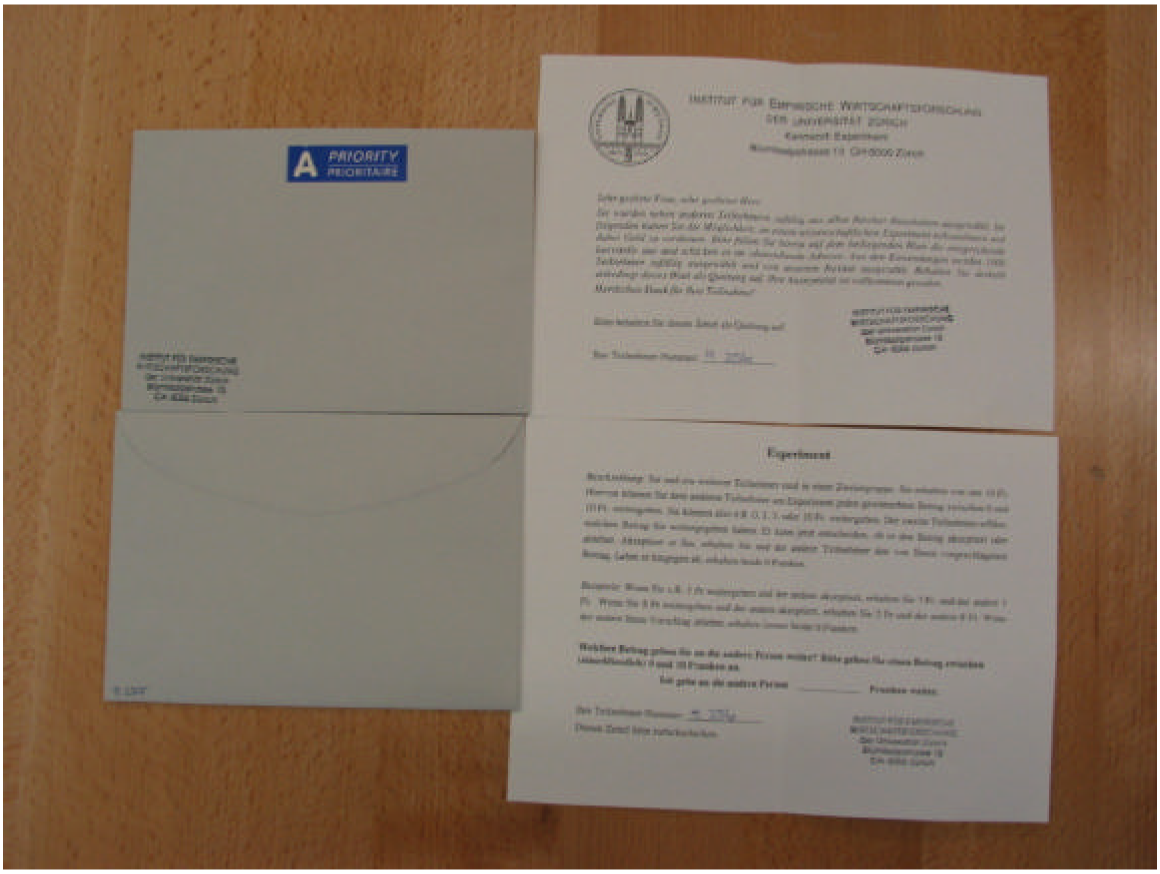

Fig. 2: The letters and the envelopes

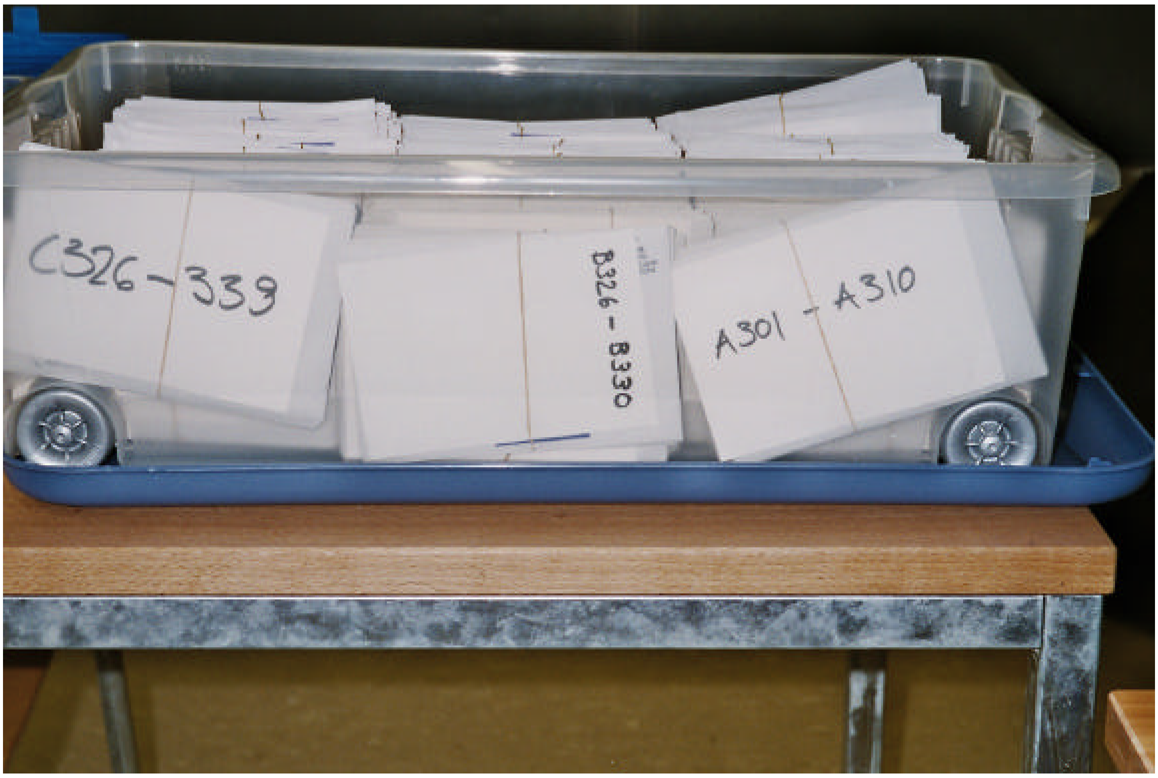

Fig. 3: The box with (high) output

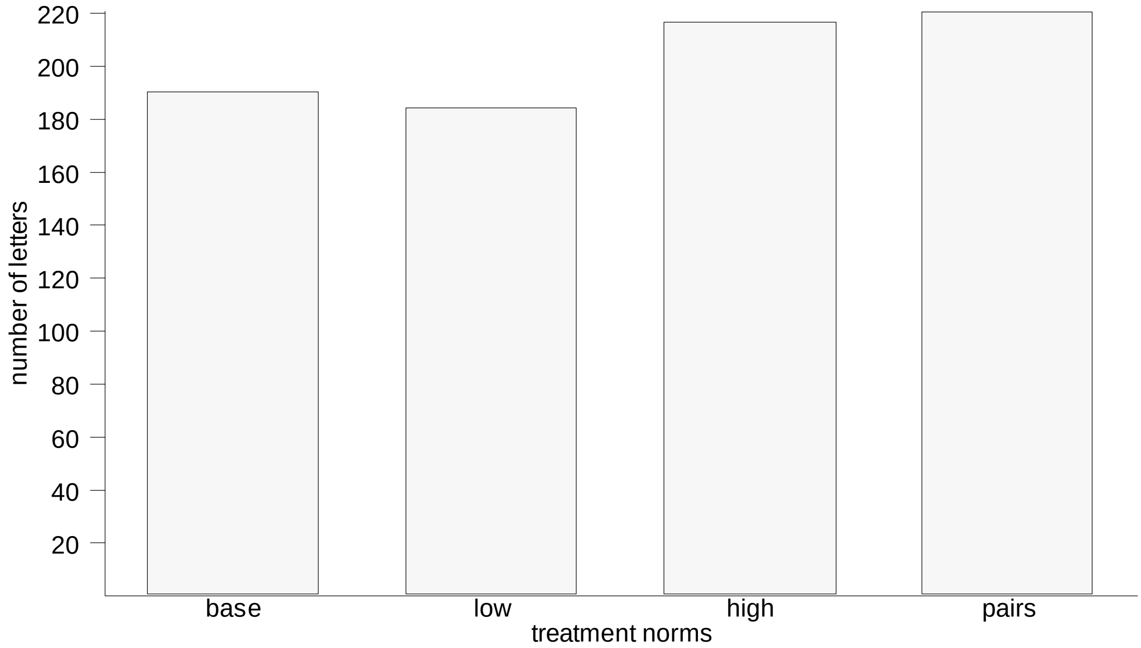

Fig. 4: Average number of letters produced in each treatment

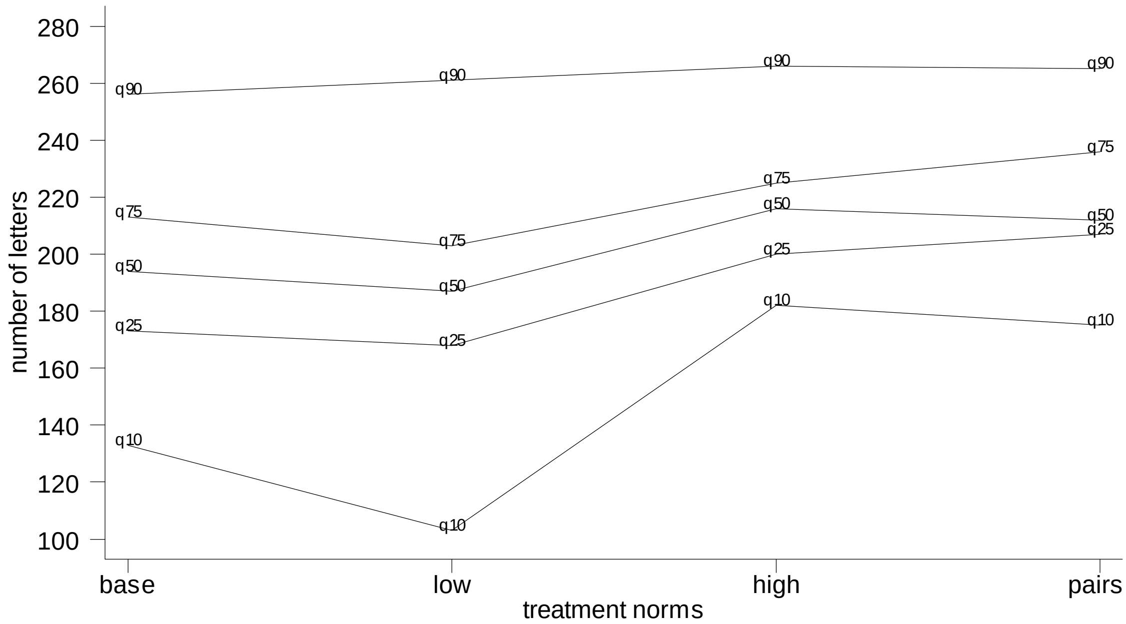

Fig. 5: Quantiles of the output distribution in each treatment

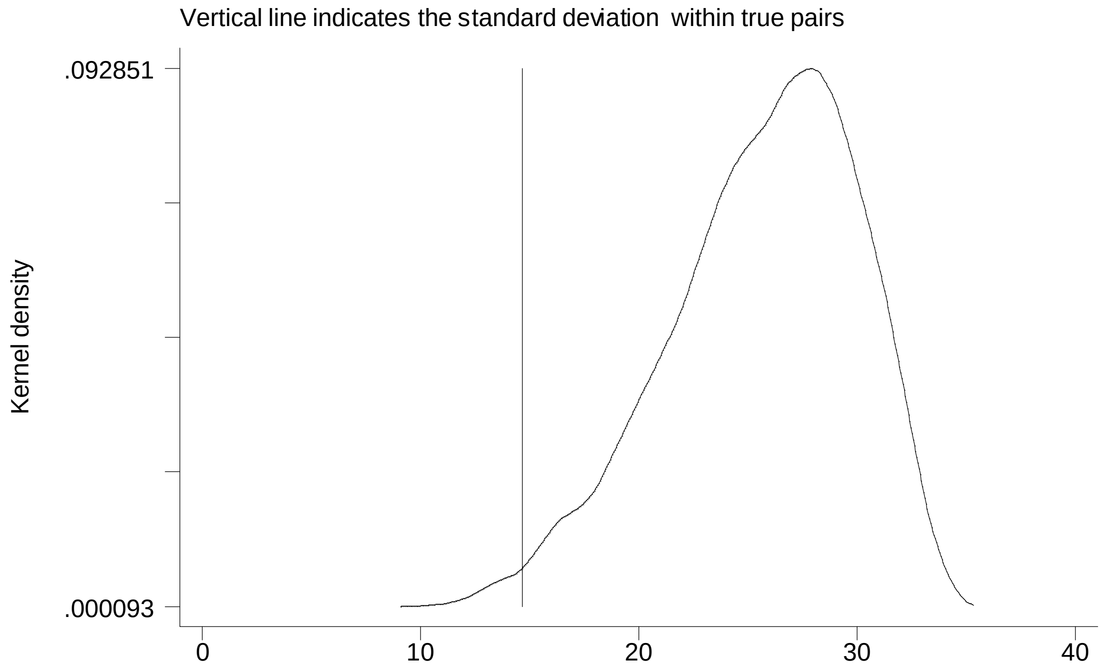

Fig. 6: St. dev. within true and hypothetical pairs in pair sample

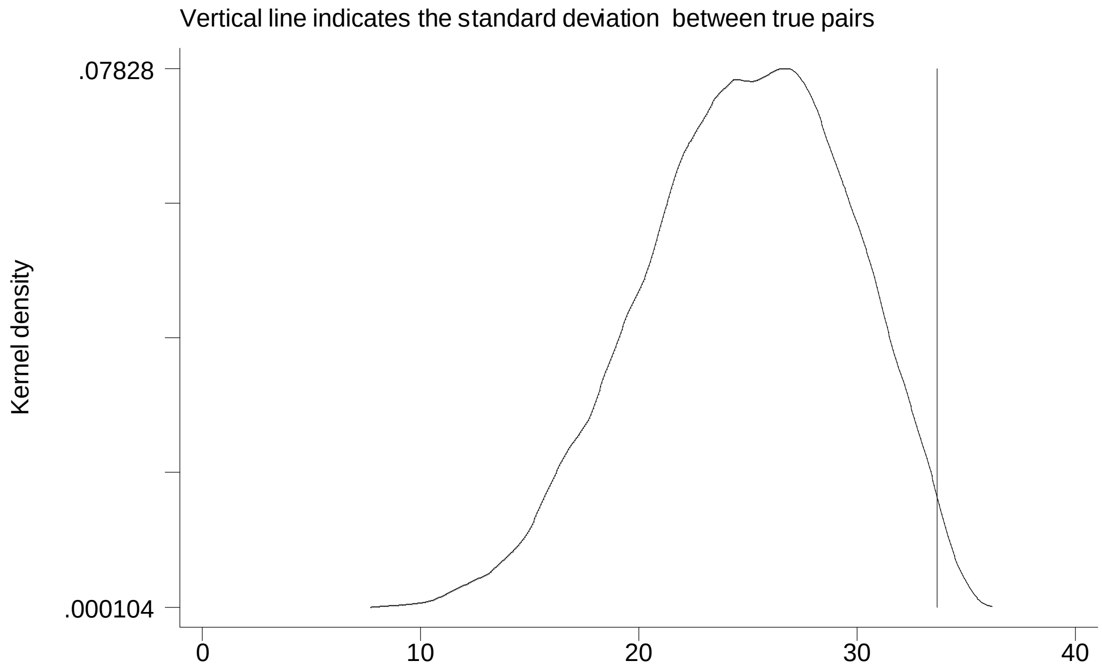

Fig. 7: St. dev. between true and hypothetical pairs in pair sample

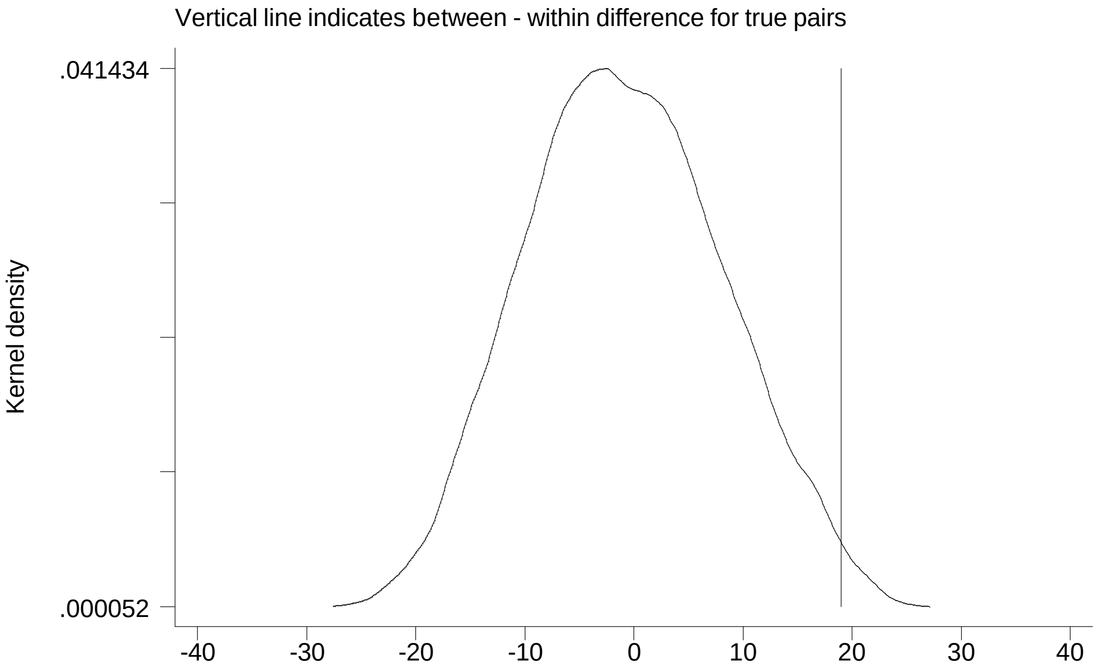

Fig. 8: Between - within st. dev. for true and hypothetical pairs