Clean Evidence on Peer Effects
Author(s): Armin Falk and Andrea Ichino
Reviewed work(s):
Source: *Journal of Labor Economics*, Vol. 24, No. 1 (January 2006), pp. 39-57
Published by: The University of Chicago Press on behalf of the Society of Labor Economists and the NORC at the University of Chicago
Stable URL: http://www.jstor.org/stable/10.1086/497818
Accessed: 16/03/2012 04:52

Clean Evidence on Peer Effects

Armin Falk, *University of Bonn, IZA, CEPR, and CESifo*

Andrea Ichino, *European University Institute, IZA, CEPR, and CESifo*

We study subjects who were asked to fill letters into envelopes with a remuneration independent of output. In the "pair" treatment, two subjects worked at the same time in the same room, and peer effects were possible. In the "single" treatment, subjects worked alone, and peer effects were ruled out. We find evidence of peer effects in the pair treatment because the standard deviations of output are smaller within pairs than between pairs. Moreover, average output is higher in the pair treatment: thus, peer effects raise productivity. Finally, low-productivity workers are the most sensitive to the behavior of peers.

I. Introduction

Scholars in many disciplines have long tried to estimate empirically the extent to which individual behavior is modified by peer effects. The reason why doing this is difficult despite the apparent wealth of evidence from daily experience is that observational data do not allow us to easily separate the pure effect of peer behavior from the effect of confounding

We would like to thank Urs Fischbacher, Simon Gächter, Lorenz Goette, Sally Gschwend, Soren Johansen, Michael Kosfeld, Ed Lazear, Alan Manning, Imran Rasul, Gianfranco Pieretti, Eric Verhoogen, and seminar participants at University of Toulouse, CESifo, European University Institute, European Summer Symposium in Labour Economics, London School of Economics, University of Siena, and University of Pisa for insightful comments and suggestions. Christoph Braun, Jasmin Gülden, and Barbara Nabold provided excellent editorial and research assistantship. Contact the corresponding author, Armin Falk, at falk@iza.org.

[*Journal of Labor Economics,* 2006, vol. 24, no. 1]

0734-306X/2006/2401-0002$10.00

factors. Using data from a controlled field experiment in which randomly selected subjects were paid independently of their work output, we show in this article 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 effects on work behavior.

To understand the nature of our experiment, consider two individuals working on separate tasks, where one is in sight of the other. Suppose that we observe them behaving in a similar way, which we suspect could be generated by peer effects. To be precise, we say that peer effects exist if the output of individual i increases when the output of j increases 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 us to fully control for these local attributes, we could observe the behavior of i and j changing simultaneously, even in the absence of true peer effects, simply because some unobserved local attributes have changed. Second, it is possible that the two individuals have similar characteristics that would make them behave similarly even if they were not working in sight of one another. With respect to both of these possibilities, it could also happen that i and j decide to work near each other 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 peers would instead be the result of sorting according to local or personal attributes.

The most recent generation of studies that try to measure peer effects with observational data has made several important steps toward solving these problems.[1] However, even if the setting offers an almost perfect opportunity to identify peer effects in many of these studies, the impossibility of controlling for all 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 Katz et al. (2001) and Sacerdote (2001), who use data based on randomized assignments of individuals to peer groups. However, both of these papers are confronted with 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

[1] See, among others, Wilson (1987), Case and Katz (1991), Crane (1991), Glaeser, Sacerdote, and Scheinkman (1996), Aaronson (1998), Encinosa, Gaynor, and Rebitzer (1998), Van den Berg, Van der Klaaw, and Van Ours (1998), Bertrand, Luttmer, and Mullainathan (2000), Ichino and Maggi (2000), Katz, Kling, and Liebman (2001), Sacerdote (2001), and Topa (2001). See also the literature based on the classic Hawthorne experiments (e.g., Whitehead [1938]; and, more recently, Jones [1990]).

cannot completely exclude the possibility that these effects might be due to local time-varying shocks. This is less of a problem in Katz et al. (2001), which analyzes the consequences of randomly changing the residential neighborhood of families residing in high-poverty public housing projects and therefore is not primarily interested in isolating pure peer effects from local effects. A further important difference with respect to our setting is that neither of these papers focuses on a work environment.

In contrast, we focus explicitly on a real work environment in our study, and we aim to assess the existence of peer effects in a fully controlled setting where no possible confounding factor can hinder this assessment. As in any other controlled experiment, the possibility of obtaining clean evidence complements the evidence generated by observational studies in an informative way.[2]

Our subjects were recruited randomly and asked to perform a typical short-term job, one that would be paid independently of individual or team output. The work task was to stuff letters into envelopes. We studied two treatments. In the "pair" treatment, our main treatment, two subjects work simultaneously in the same room. This setting allows for the possibility that the behavior of a subject is affected by the behavior of the other member of the pair. Given two subjects i and j in a pair, we speak of positive peer effects if the output of i systematically raises the output of j and vice versa, leading to similar output levels within the pair. A formal characterization of this definition will be given in Section III. In the second treatment (the "single" treatment), which serves as our control, peer effects are ruled out by design because subjects work alone in a room. Output in this treatment reveals the level of productivity in the absence of any peer influence. The comparison of the outputs arising in the pair treatment with those from the single treatment permits the assessment of the effects of peers on individual productivity.

Our main results are as follows. First, we find strong and unambiguous evidence for the existence of positive peer effects in the pair treatment. This can be inferred from the fact that output within pairs is very similar, while it differs substantially between pairs. This difference is particularly striking when compared to what happens in random allocations of subjects from the pair and the single treatment in simulated pairs. By comparing the standard deviation of output within and between true and simulated pairs, we show that peer effects are large and highly significant. Second, even though economic incentives are identical, average output in the pair treatment is higher than that in the single treatment. Thus, peer effects

[2] For related literature on laboratory experiments aimed at measuring peer effects, see Falk and Fischbacher (2002) and Falk, Fischbacher, and Gächter (2003). Nagin et al. (2002) provide, instead, an example of controlled experimentation in a real labor setting, although their focus is on a different issue.

significantly increase output. Third, we show that peer influence affects subjects differently. In particular, we find that it mainly improves the output of less productive subjects. Finally we derive an implicit estimate for the strength of peer effects. Interestingly, the estimated coefficient is very similar to a comparable estimate derived by Ichino and Maggi (2000) with observational data.

The remainder of this article is organized as follows. The next section, Section II, describes the design of our experiment. Section III discusses the behavioral hypotheses, while Section IV presents our results. Section V concludes.

II. Design of the Field Experiment

The goal of this article was to study potential peer 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 two different treatments. Before discussing our treatments in detail, we describe the recruitment process, the work task, and the procedures.

A. Recruitment

All of 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). Students were asked in announcements posted on blackboards if they would like to do a simple short-term job requiring no previous knowledge. In the announcement, it was stated that the job was a one-time 4-hour job and it paid SFr 90 (1 SFr [Swiss franc] ≈ .70 US$ ≈ €.70). The payment was obviously attractive, as we were able to recruit the number of subjects we had planned to recruit within 24 hours.

Students applied by e-mail. After receiving their applications, we informed them of the precise date and location where they were expected to carry out the job. The experiment took place during the spring 2002 vacation, a period covering 2 weeks. The work was performed in a high school building in Winterthur.

B. Procedure and Task

Upon arrival, subjects were welcomed and informed about the task and the procedural details. In particular, they were told that they had to work for 4 hours without a break and that at the end of this time, they would receive their payment.

We chose a work task that was simple, required no previous knowledge, and was easy to measure. In particular, students had to prepare the mailing of a questionnaire study for the University of Zurich. This job basically involved stuffing letters into envelopes. First, subjects had to fold two

FIG. 1.—One of the desks used for the experiments

sheets of paper (one sheet contained the description of the questionnaire; the other was to be filled out by the recipients of the study). After placing the two sheets into the envelope, subjects had to seal the envelope and put an "A-priority" sticker on it. When a set of 25 envelopes had been completed, the set had to be bundled with a rubber band and put in a box. The work environment was exactly the same for each subject, including, for example, the same type of desk and chair and the same large number of envelopes and sheets (fig. 1 displays a picture of a subject's desk). Payment was independent of output and was paid in cash. The procedure was exactly the same in both treatments.

Note that we determined a subject's output only after he or she had been paid. This implies that we have only one output observation per subject. An interesting question that we cannot address with this procedure concerns the importance of peer effects over time. For example, peer effects may be stronger at the beginning than at the end of the experiment. To address differences over time it would have been necessary to document output not only at the end of the experiment but also during the experiment. This would have required, for example, installing a camera and videotaping subjects' behavior or having an experimenter count output during the experiment. It is likely that this type of intervention would have been experienced as quite "unnatural" by our subjects. Since one of our primary goals was to implement a rather natural work environment, we decided against this option.

C. Treatments

We studied two treatments, the pair treatment and the single treatment. In the pair treatment, two subjects did the task described above at the same time in the same room. The two desks were situated in such a way that a subject could easily realize the output of the other subject (the position of the second desk can be seen in the background of fig. 1). Subjects were free to communicate, but they were instructed that they had to perform independently the task described above. Hence, they were not allowed to engage in teamwork or division of labor. We invited only students from different high schools to participate in this treatment in order to minimize the possibility that two subjects in the pair treatment would know each other.

In the pair treatment, peer effects were possible. In contrast, peer effects were ruled out by design in the single treatment. In this control treatment, everything was exactly the same as in the pair treatment, except that in this case each subject worked alone in a room. Since subjects did not have any contact with another subject and were not informed about other subjects' output in this treatment, the single treatment rules out any potential peer effect stemming from a coworker. Therefore, a comparison of output arising in the single treatment with that of the pair treatment indicates the potential effects of peers on productivity.

A total of 24 subjects participated in our study, eight in the single treatment and 16 (eight pairs) in the pair treatment. Treatment assignment was random. No subject participated in more than one treatment.

From a methodological point of view, some aspects of 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, that is, instead of choosing real effort, subjects choose between numbers knowing that each choice has a different cost. This procedure has been used in tournament experiments (e.g., Bull, Schotter, and Weigelt 1987) and in efficiency wage experiments (e.g., Fehr, Kirchsteiger, and Riedl 1993). Some authors have recently conducted so-called real effort experiments to study incentive mechanisms and efficiency wages. In Fahr and Irlenbusch (2000), subjects had to crack walnuts; in Van Dijk, Sonnemans, and Van Winden (2001), subjects performed cognitively demanding tasks on the computer (two-variable optimization problems); and in Gneezy (2003) and Gneezy, Niederle, and Rustichini (2003), subjects had to solve mazes at the computer. While these real effort tasks resemble regular work more than just choosing a number, subjects probably did not perceive these tasks as economically valuable. This means that an important dimension of regular work was missing. In contrast, subjects in our study performed a regular, economically valuable job.

III. Behavioral Hypotheses

To illustrate what we expect to happen in our experiment in the presence of peer effects, 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 effects by modeling 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 effects $L_{12} = 0$, while if instead $L_{12} \leq 0$, there are peer effects in the sense that the cost of producing X_i is lower when average production is higher. Note that our goal is not to identify the determinants of peer effects but just to understand what we should see in the data generated by our experiment if peer effects are at work. 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.[3] Note that the way we model peer effects differs from the way peer effects are modeled, for example, in Kandel and Lazear (1992) or Huck et al. (2002). The latter models focus on the question of whether peer pressure improves performance in situations where payments are based on team incentives, that is, where a subject's effort has a positive externality on other team members' profits.

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 positive peer effects exist and $\partial X_i / \partial \theta_i > 0$. Note that equation (1) is a structural condition because $\bar{X}_i^e$ is endogenous.

[3] For discussions of possible determinants of peer effects see, among others, Kandel and Lazear (1992), Akerlof (1997), Spagnolo (1999), and Huck, Kübler, and Weibull (2002).

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, note that, in the pair treatment, $\bar{X}_i^e = X_j$, where i and j denote the two subjects in a pair. Moreover, since the environment is controlled by the experimenter and kept constant for all subjects and treatments, $Y_i^e = Y$. Thus, linearizing, for simplicity, equation (1), we obtain

$$X_i = Y + \beta X_j + \theta_i \tag{3}$$

and symmetrically for j. In this equation, β measures how the output of i depends on the output of j when they work in a pair. Within this context, we say that peer effects exist and are positive if the output of i increases with the output of j, which formally means the following:

Definition 1. If $\beta > 0$, positive peer effects exist in a pair. $\beta = 0$ implies absence of peer effects, while these effects are negative if $\beta < 0$.

In the equilibrium of the pair treatment, the output of subject i is given by

$$X_i^p = \frac{Y}{(1-\beta)} + \frac{\theta_i + \beta\theta_j}{(1-\beta^2)}, \tag{4}$$

while the same subject in the single treatment would produce

$$X_i^s = Y + \theta_i, \tag{5}$$

since in this treatment no other subject exercises any pressure on i. Symmetrically, we can derive analogous expressions for j. It is important to note that random assignment ensures that types θ are randomly distributed in the two treatments.

Points P and S in figure 2 describe the respective equilibria of the pair and single treatments. The figure also plots the reaction curves described by equation (3) for the pair treatment, which cross at P, and by equation (5) for the single treatment, which cross at S.

It is immediately obvious that the difference between the output levels of the two subjects within each pair is equal to

$$|X_i^p - X_j^p| = \frac{|\theta_i - \theta_j|}{1+\beta}. \tag{6}$$

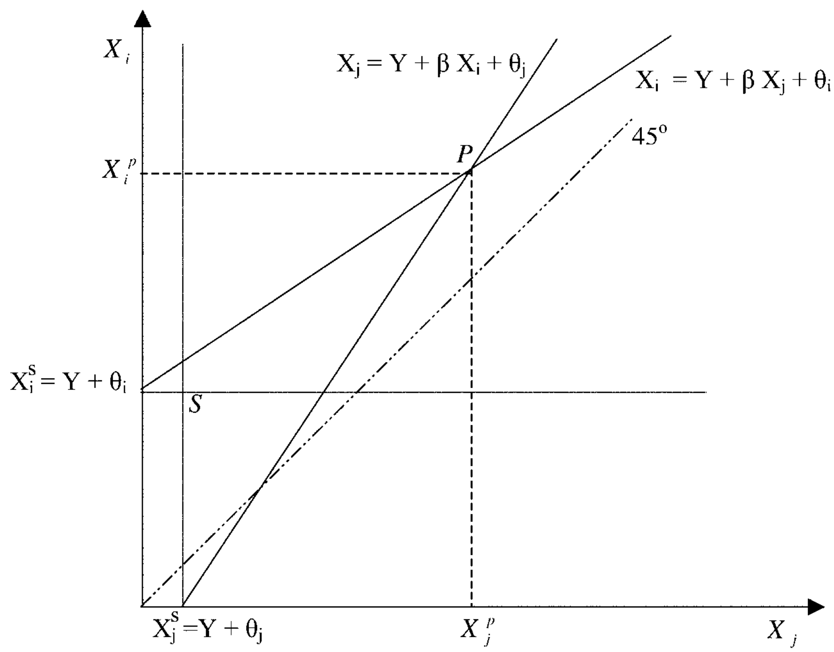

FIG. 2.—Reaction curves and equilibria in the pair and single treatments

As a result, positive peer effects can be detected in the pair treatment according to the following proposition, which will be tested in Section IV.

PROPOSITION 1. If positive peer effects exist, that is, $\beta > 0$, the absolute value of the difference between output levels within pairs should be smaller than if there were no peer effects.

An illustration of proposition 3 is given with the help of figure 2, where P shows an equilibrium with $\beta > 0$ and S shows an equilibrium with $\beta = 0$. Since P is closer to the 45° line than S, output levels are more similar in P in comparison to S. Moreover, it is obvious that a higher β implies output levels that are increasingly similar in the P equilibrium.

The setting of our experiment offers the possibility for testing further implications of peer effects. In the absence of these effects, the distributions of output should be the same in the pair treatment and in the single treatment. This is so because the economic incentives are identical in both conditions. Of course, there might be individual differences because some subjects are, for example, more talented than others or feel more obliged to perform well than others do. However, since subjects are randomly allocated to the treatment conditions, individual differences should cancel out.

On the contrary, if peer effects do exist, it is easy to show that the average output in the two treatments should differ. Using equation (3), the average output of i and j when they work in a pair is

$$\frac{X_i^P + X_j^P}{2} = \frac{Y}{(1-\beta)} + \frac{(\theta_i + \theta_j)/2}{1-\beta}, \tag{7}$$

while the average output of the same two subjects working alone in the single treatment would be

$$\frac{X_i^S + X_j^S}{2} = Y + \frac{\theta_i + \theta_j}{2}. \tag{8}$$

A comparison of equations (7) and (8) shows that, in the presence of positive peer effects such that $0 < \beta < 1$, average output is higher in the pair treatment than in the single treatment. This can also be inferred from figure 2, where output in the P equilibrium is clearly higher compared to output in the S equilibrium. If instead $\beta > 1$, the output level of the two subjects would still be higher in the pair treatment but it would be equal to infinity. On the contrary, in the case of negative effects ($\beta < 0$), the output of a subject reduces the output of the other, and then the output of the pair treatment would be lower than the output of the single treatment. Our model therefore suggests a second proposition, which will be tested in Section IV.

Proposition 2. In the presence of positive peer effects, the average output of the pair treatment exceeds that of the single treatment.

Note that proposition 2 states a behavioral consequence of peer effects, which is similar to the so-called social facilitation paradigm in social psychology. According to this paradigm, even the mere presence of another person improves one's performance. Numerous studies have supported evidence for this type of behavior.[4]

Our final proposition derives immediately from the first two given the setting, but it is worth stating it explicitly. Let us focus on a pair of subjects i and j and consider the differences $\Delta X_i = X_i^P - X_i^S$ and $\Delta X_j = X_j^P - X_j^S$, which measure the change between the two potential output levels in the pair and in the single treatment, respectively, for i and j. Using (4) and (5), note that $\Delta X_j - \Delta X_i = (X_i^S - X_j^S) - [(X_i^S - X_j^S)/(1+\beta)]$. It is easy to see that, if $\beta > 0$, this quantity is strictly positive only when

[4] See, e.g., Zajonc (1965), Cottrell et al. (1968), and Hunt and Hillery (1973). In Allport (1920), the 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-yard mark from a standing start at traffic lights. He reports that, if there are two cars at the traffic light, the time to travel the 100 yards is significantly shorter than if there is just one car.

$X_i^S > X_j^S$, that is, only when i is the member of the pair who would produce more if both subjects were working alone.[5] Thus, the following proposition holds.

PROPOSITION 3. Given a pair of subjects working together, the subject who would be less productive working alone is the one whose output increases by more when joining the pair as opposed to working alone.

Hence, our simple model suggests three propositions that describe the implications of peer effects in our treatments. We test these propositions in the next section, where we also show how our data can be used, in light of the model described above, to derive an implicit estimate of β.

IV. Results

In this section, we present our results and test our behavioral predictions. Our main interest concerns the existence of positive peer effects, which are revealed by the observation that output levels within pairs are similar in the pair treatment.

In order to test proposition 1, consider the standard deviation of output within and between pairs.[6] In the absence of peer effects (i.e., $\beta = 0$), working in a particular pair has no effect on individual behavior. In this case, therefore, the standard deviations of output within pairs should be identical to those generated by any simulated configuration of pairs constructed from the subjects who worked alone in the single treatment. Given our data, there are 105 possible configurations of four pairs with eight individuals.[7] Only one of these 105 configurations originates a hypothetical within standard deviation lower than that obtained with the true pairs of the pair treatment. The likelihood that this finding is just pure coincidence in the absence of peer effects is below 1%.

A second test of the same proposition is possible by comparing the true pairs and simulated pairs using only the subjects involved in the pair treatment. In the absence of peer effects (i.e., if $\beta = 0$), the standard deviation of output within the true pairs should be identical to the within standard deviation generated by any simulated 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 obtained with the true pairs should differ in any specific direction. These comparisons are shown in figures 3, 4, and 5.

[5] This is also graphically evident in fig. 2.

[6] We use standard deviations instead of differences to facilitate the computation and the comparison of within and between statistics. This, however, does not change the substance of our results because, in our specific case, the standard deviation within a pair is equal to the absolute value of the difference between the output levels of the pair divided by the square root of 2.

[7] 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., eight in our case.

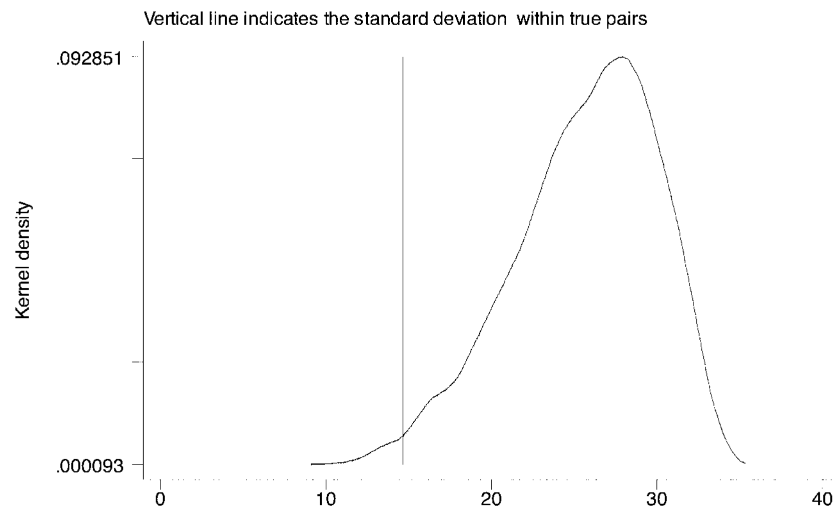

FIG. 3.—Standard deviation within true and hypothetical pairs in pair sample

Figure 3 plots the kernel density of the simulated within-pairs standard deviations computed for 20,271 randomly chosen different configurations of pairs of the 16 individuals involved in the pair treatment. To be more precise, we generated all 2,027,025 possible configurations of eight pairs with these 16 individuals,[8] and for one out of every 100 of these configurations we computed the within-pairs standard deviation.[9]

The variation of these simulated within standard deviations ranges from 9.6 to 34.8 letters. The vertical line in figure 3 identifies the standard deviation within true pairs, that is, that 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 simulated configurations originated a lower value. This 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. In other words, in the absence of any peer effect, the probability of observing a within-pairs deviation as low as 14.6 is, on average, less than

[8] See n. 7.

[9] We would have liked to have computed the within-pairs standard deviations for all 2,027,025 configurations, but this calculation would have required a substantial amount of computer time without any major gain from the viewpoint of the reliability of our results.

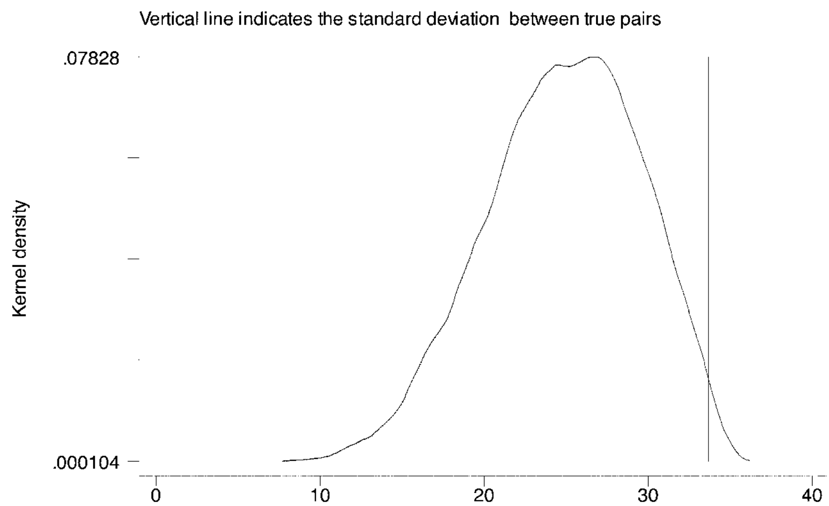

FIG. 4.—Standard deviation between true and hypothetical pairs in pair sample

1.17%.[10] Hence, we can reject the hypothesis of the absence of peer effects with a high level of confidence.

In line with figure 3, we find in figure 4 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 simulated configurations of pairs. The chance that such a high between standard deviation could be generated in the absence of peer effects is extremely low (in particular, smaller than 1.15%). Moreover, figure 5 plots the kernel density of the between minus within difference for each hypothetical configuration of pairs. 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 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 simulated 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

[10] Note that the standard deviations computed for the simulated configurations are identically but not independently distributed random variables. Because of stochastic variation, the true probability of observing a within standard deviation smaller than 14.6 in a simulated configuration might be larger or smaller than 1.17%. However, it will be equal to this value on average, since the random variables are identically distributed.

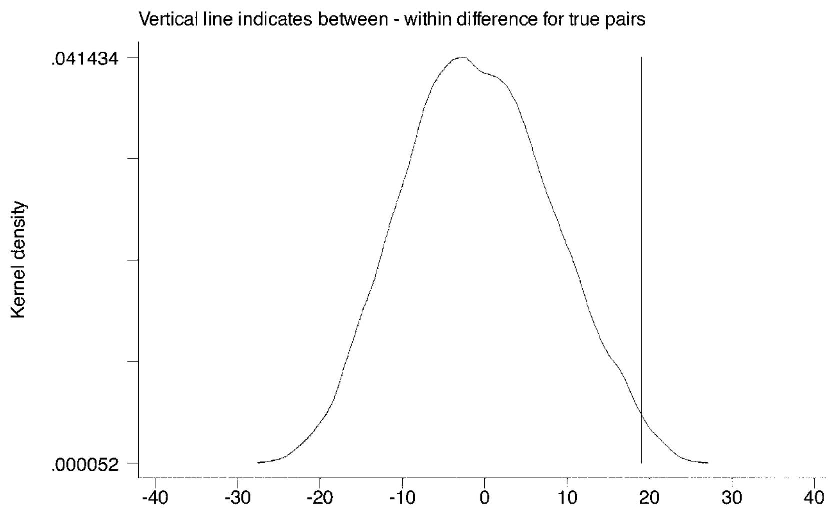

Fig. 5.—Between minus within standard deviation for true and hypothetical pairs

standard deviation to be smaller than the between standard deviation or vice versa, figure 5 suggests that, when individuals are paired in the same room, the between-pairs deviation is significantly larger than the within-pairs deviation. This implies that, ceteris paribus, working in pairs induces more similar output levels than working separately.

We now turn to our second proposition. Remember that, according to standard economic theory, average output levels in the pair treatment and in the single treatment should be similar, because incentives are identical in both treatments. In the presence of peer effects, however, output should be higher in the pair treatment compared to the single treatment. This is in fact what we find. The average output in the single treatment is 190 envelopes, while the average output in the pair treatment is 221 envelopes. The difference is not only sizable in percentage terms (16.3%) but also statistically significant despite the small sample size. To show this, we regress outputs in both treatments on a treatment dummy for the pair treatment. The respective p-value of this dummy is 0.068. This is confirmed by the nonparametric Wilcoxon rank sum test ($p = 0.049$, one sided). Thus peer effects lead to higher average output, as hypothesized in proposition 2.

Finally we provide evidence in favor of our third proposition, which suggests that, in the presence of positive peer effects, the subject of a pair who would be less productive in the single treatment is the one who would increase the productivity by more when joining the pair as opposed to working alone. With the data at our disposal, we cannot perform a direct test of proposition 3, because for each subject we observe the output

Table 1
Quantiles of the Output Distribution in Each Treatment

Quantile	Single Treatment	Pair Treatment	Difference
10th	133	175	42
25th	173	207	34
50th	194	212	18
75th	213	236	23
90th	256	265	9

Note.—Columns 1 and 2 of the table report the quantiles of the output distribution for the single and the pair treatments, estimated using a quantile regression of output on a dummy for the pair treatment plus a constant. Column 3 reports the absolute value of the difference between the quantiles estimated for the two treatments.

level only in the treatment he or she is assigned to but not the output level in the counterfactual treatment. However, the comparison of the quantiles of the distributions for the single and the pair treatment provide evidence that is consistent with this proposition. These quantiles are shown in columns 1 and 2 of table 1. The output of the 10th quantile in the single treatment is 133, while the output of the same quantile in the pair treatment is 175, which implies a difference of 42 letters. For the 90th quantile, instead, the output levels of the two treatments are, respectively, 256 and 265, with a difference of only nine letters. Thus, at low productivity levels, peer effects determine large differences in output between the pair treatment and the single treatment, while at high productivity levels the differences are small. In fact, the Spearman rank correlation between the differences and the corresponding productivity levels is negative and significant (Spearman's $\rho = -0.900$, $p = .018$, one sided). This is what one would expect to see in the data given proposition 3.

We conclude this section by showing how, in the light of our simple model of Section III, the data generated by our experiment can be used to estimate β. Remember that this parameter measures how the output of i influences the output of j in a pair and vice versa. Equations (4) and (5) say that a subject i's outputs in the pair treatment and in the single treatment are given, respectively, by $X_i^p = [Y/(1-\beta)] + [(\theta_i + \beta\theta_j)/(1-\beta^2)]$ and $X_i^s = Y + \theta_i$. Substituting the sample averages $\bar{X}^p$ for X_i^p and $\bar{X}^s$ for θ_i and θ_j, after simplifying we can compute the average β solving $\bar{X}^p = (\bar{X}^s + \beta\bar{X}^s)/(1-\beta^2)$ or $221 = (190 + \beta 190)/(1-\beta^2)$. This gives an implicit estimate of $\beta = 0.14$, which implies that, when the output of j increases by one unit, the output of i increases by 0.14 units, on average. Of course, we do not claim that 0.14 is a universal number. Yet, it is interesting and reassuring to see that Ichino and Maggi (2000), who derive a comparable estimate of β with observational data, get very similar numbers. Their study exploits personnel longitudinal data on workers in different branches of a large Italian firm to analyze peer effects in absenteeism behavior. On the basis of a first-difference version of equation (3), they look at movers between branches and estimate how much their absenteeism changes as a function of the difference in the absenteeism of stayers of the departure and arrival branch. This procedure allows them to control for individual and branch fixed effects as well as for several individual and local time-varying covariates. Depending on the specification, their estimates of β are $\beta = 0.14$, $\beta = 0.18$, and $\beta = 0.15$.

V. Concluding Remarks

In this article, we have presented clean evidence in favor of the existence of peer effects. We have shown, using a controlled field experiment, that the behavior of subjects working in pairs is significantly different than the behavior of subjects working alone. The standard deviations within pairs are significantly smaller than those between pairs. As a second result, peer effects work in the direction of raising the overall average productivity significantly. We also show that the less productive workers are more affected by peer effects than the high-productivity workers. In other words, "bad apples," far from damaging "good apples," seem instead to gain in quality when paired with the latter. The presence of peer effects is robust and quantitatively important, even though subjects interacted only once and did not know each other. This suggests that the effects measured in our study are a lower boundary for the effects that prevail in actual labor relations.

In our experiment, we implemented a fixed pay regime; that is, payments were not conditioned on output. This was done in order to keep things as simple as possible. In many firms, however, more sophisticated pay regimes are common. Often pay is based on individual performance, as in piece rate schemes. Alternatively, pay often depends on group performance, as in profit sharing, team-based incentives, or stock option plans. It would be interesting to study the interaction of these incentive schemes and peer effects in our real effort setup.

If pay is based on group incentives, it may well be the case that peer effects play a decisive role in enhancing performance. This holds, in particular, if group members can directly exhibit peer pressure in the form of social exclusion or sanctions or by pushing poor performers out of the team. This type of peer pressure is not modeled in our experiment, where peer effects are operative through comparisons only. There is, however, evidence from public goods experiments that peer pressure enhances team performance. In these experiments, subjects are given the opportunity to sanction other players after observing their contributions to the team. In the absence of such sanction possibilities, cooperation and team performance are typically rapidly declining over time. In contrast, if sanctions are possible, free riders are severely sanctioned and contribution levels are high and stable (Fehr and Gächter 2000). The driving forces behind this type of sanctions are studied in much detail in Falk, Fehr, and Fischbacher (2005).

In the case of individual pay-for-performance, peer effects may play a less important role simply because the situation is much more structured and individual output produces no externality on the earnings of other employees. However, even if pay is based on individual output, employees may care about their own output and pay relative to the output and pay of their coworkers. If this holds, peer effects of the kind studied in this article may affect effort even in pay-for-performance schemes.

The existence of peer effects raises an interesting question concerning the efficient design of the workplace. Should employees work in groups or alone? If they work in groups, how should low-productivity and high-productivity workers be optimally grouped? In light of our results, the output maximizing strategy is to let people work in groups rather than alone. In addition, since high-productivity workers seem to elevate low-productivity workers, it may be optimal to group low- and high-productivity workers instead of grouping workers of similar productivity. These conclusions, however, are based on peer effects observed in a fixed pay regime and for simple job tasks. Whether they remain valid in a more general setting and, specifically, if employees are paid according to group incentives or individual pay-for-performance remains an empirical question to be considered in future research.

References

Aaronson, Daniel. 1998. Using sibling data to estimate the impact of neighborhoods on children's educational outcomes. *Journal of Human Resources* 33, no. 4:915–46.

Akerlof, George. 1997. Social distance and social decisions. *Econometrica* 65, no. 4:1005–27.

Allport, Floyd H. 1920. The influence of the group upon association and thought. *Journal of Experimental Psychology* 3, no. 3:159–82.

Bertrand, Marianne, Erzo F. P. Luttmer, and Sendhil Mullainathan. 2000. Network effects and welfare cultures. *Quarterly Journal of Economics* 115, no. 3:1019–55.

Bull, Clive, Andrew Schotter, and Keith Weigelt. 1987. Tournaments and piece rates: An experimental study. *Journal of Political Economy* 95, no. 1:1–33.

Case, Anne, and Lawrence Katz. 1991. The company you keep: The effect of family and neighborhood on disadvantaged youth. NBER Working Paper no. 3705, National Bureau of Economic Research, Cambridge, MA.

Cottrell, Nicholas B., Dennis L. Wack, Gary J. Sekerak, and Robert H. Rittle. 1968. Social facilitation of dominant responses by the presence of an audience and the mere presence of others. *Journal of Personality and Social Psychology* 9, no. 3:245–50.

Crane, Jonathan. 1991. The epidemic theory of ghettos and neighborhood effects on dropping out and teenage childbearing. *American Journal of Sociology* 96, no. 5:1226–59.

Encinosa, William E., Martin Gaynor, and James B. Rebitzer. 1998. The sociology of groups and the economics of incentives: Theory and evidence on compensation systems. Working paper, Department of Economics, Case Western Reserve University.

Fahr, René, and Bernd Irlenbusch. 2000. Fairness as a constraint on trust in reciprocity: Earned property rights in a reciprocal exchange experiment. *Economics Letters* 66, no. 3:275–82.

Falk, Armin, Ernst Fehr, and Urs Fischbacher. 2005. Driving forces behind informal sanctions. *Econometrica* 73, no. 6:2017–30.

Falk, Armin, and Urs Fischbacher. 2002. "Crime" in the lab-detecting social interaction. *European Economic Review* 46, nos. 4–5:859–69.

Falk, Armin, Urs Fischbacher, and Simon Gächter. 2003. Two neighborhoods: Social interactions in the lab. Working Paper no. 150, Institute for Empirical Research in Economics, University of Zurich.

Fehr, Ernst, and Simon Gächter. 2000. Cooperation and punishment in public goods experiments. *American Economic Review* 90, no. 4: 980–94.

Fehr, Ernst, George Kirchsteiger, and Arno Riedl. 1993. Does fairness prevent market clearing? An experimental investigation. *Quarterly Journal of Economics* 108, no. 2:437–59.

Glaeser, Edward L., Bruce Sacerdote, and Jose A. Scheinkman. 1996. Crime and social interactions. *Quarterly Journal of Economics* 111, no. 2:507–48.

Gneezy, Uri. 2003. Do high wages lead to high profits? An experimental study of reciprocity using real effort. Working paper, Graduate School of Business, University of Chicago.

Gneezy, Uri, Muriel Niederle, and Aldo Rustichini. 2003. Performance in competitive environments: Gender differences. *Quarterly Journal of Economics*, 118, no. 3:1049–74.

Huck, Steffen, Dorothea Kübler, and Jörgen Weibull. 2002. Social norms and incentives in firms. Working Paper no. 565, Research Institute of Industrial Economics, Stockholm.

Hunt, Peter J., and Joseph M. Hillery. 1973. Social facilitation in a coaction setting: An examination of the effects over learning trials. *Journal of Experimental Social Psychology* 9, no. 6:563–71.

Ichino Andrea, and Giovanni Maggi. 2000. Work environment and individual background: Explaining regional shirking differentials in a large Italian firm. *Quarterly Journal of Economics* 115, no. 3:1057–90.

Jones, Stephen R. G. 1990. Worker interdependence and output: The

Hawthorne Studies reevaluated. *American Sociological Review* 55, no. 2:176–90.

Kandel, Eugene, and Edward P. Lazear. 1992. Peer pressure and partnerships. *Journal of Political Economy* 100, no. 4:801–17.

Katz, Lawrence F., Jeffrey R. Kling, and Jeffrey B. Liebman. 2001. Moving to opportunities in Boston: Early results of a randomized mobility experiment. *Quarterly Journal of Economics* 116, no. 2:607–54.

Manski, Charles F. 1993. Identification of endogenous social effects: The reflection problem. *Review of Economic Studies* 60, no. 3:531–42.

Nagin, Daniel S., James B. Rebitzer, Seth Sanders, and Lowell J. Taylor. 2002. Monitoring, motivation and management: The determinants of opportunistic behavior in a field experiment. *American Economic Review* 92, no. 4:850–73.

Sacerdote, Bruce. 2001. Peer effects with random assignment: Results for Dartmouth roommates. *Quarterly Journal of Economics* 116, no. 2: 681–704.

Spagnolo, Giancarlo. 1999. Social relations and cooperation in organizations. *Journal of Economic Behavior and Organization* 38, no. 1:1–25.

Topa, Giorgio. 2001. Social interactions, local spill-overs, and unemployment. *Review of Economic Studies* 68, no. 2:261–95.

Towler, Gary. 1986. From zero to one hundred: Coaction in a natural setting. *Perceptual and Motor Skills* 62:377–78.

Van den Berg, Gerard J., Bas Van der Klaauw, and Jan C. Van Ours. 1998. The effect of neighborhood characteristics on the labor supply of welfare recipients. Working paper, CentER, Tilburg.

Van Dijk, Frans, Joep Sonnemans, and Frans Van Winden. 2001. Incentive systems in a real effort experiment. *European Economic Review* 45, no. 2:187–214.

Whitehead, Thomas N. 1938. *The industrial worker: A statistical study of human relations in a group of manual workers.* Cambridge, MA: Harvard University Press.

Wilson, William J. 1987. *The truly disadvantaged.* Chicago: University of Chicago Press.

Zajonc, Robert B. 1965. Social facilitation: A solution is suggested for an old unresolved social psychological problem. *Science* 149, no. 3681: 269–74.