SOCIAL COMPARISON AND EFFORT PROVISION: EVIDENCE FROM A FIELD EXPERIMENT

Alain Cohn
Department of Economics,
University of Zurich

Ernst Fehr
Department of Economics,
University of Zurich

Benedikt Herrmann
School of Economics,
University of Nottingham

Frédéric Schneider
Department of Economics,
University of Zurich

Abstract
Social comparison has potentially far reaching consequences in many economic domains. We conducted a field experiment to examine how social comparison affects workers' effort provision if their own wage or that of a co-worker is cut. Workers were assigned to groups of two, performed identical individual tasks, and received the same performance-independent hourly wage. Cutting both group members' wages caused a decrease in performance. But when only one group member's wage was cut, the affected workers decreased their performance more than twice as much as when both workers' wages were cut. This finding indicates that social comparison among workers affects effort provision because the only difference between the two wage-cut treatments is the *other* group member's wage level. In contrast, workers whose wage was not cut but who witnessed their group member's pay being cut displayed no change in performance relative to the baseline treatment in which both workers' wages remained unchanged. This indicates that social comparison exerts asymmetric effects on effort. (JEL: C93, J33, M53)

1. Introduction

The hypothesis that relative income affects individuals' welfare and behavior has a long history in economics (Duesenberry 1949; Veblen 1899). More recently, several studies have indicated that income comparisons affect happiness, job satisfaction and other reward experiences (e.g. Card et al. 2012; Fliessbach et al. 2007; Luttmer 2005).

The editor in charge of this paper was Stefano DellaVigna.

Acknowledgments: We would like to thank the participants at numerous conferences, workshops and seminars at which we presented this study, as well as the four anonymous referees for their valuable comments. We are especially grateful for the thoughtful input from the editor. We would like to acknowledge the generous support of the Swiss National Science Foundation (Grant No. 100018_140734/1 on "The distribution and determinants of social preferences").

E-mail: alain.cohn@econ.uzh.ch (Cohn); ernst.fehr@econ.uzh.ch (Fehr); Benedikt.herrmann@gmail.com (Herrmann); frederic.schneider@econ.uzh.ch (Schneider)

Journal of the European Economic Association xxx 2014 00(00):1–22

DOI: 10.1111/jeea.12079

Evidence for the behavioral effects of social comparison is still very scarce, however, although they have received attention throughout the history of economics.

Social comparison among workers may, for example, play an important role both within firms and in the labor market. It may affect firms' internal pay structure, employee morale, and thus firms' labor demand, firm boundaries, wage changes over time, and even the rate of unemployment.[1] The hypothesis that social comparison and associated fairness concerns affect the labor market dates back at least to Slichter (1920). Practitioners have also observed that unjust inequalities in wages within a firm "break down morale, create individual dissatisfaction, encourage absenteeism, increase labor turnover and hamper production" (in Kochan and Barocci 1985, p. 249).

We conducted a field experiment in collaboration with a firm that hired workers for a *one-time* sales promotion. The workers' task was to sell promotional cards that permitted entrance to specific nightclubs and bars on specific dates, yielding a clean measure of each worker's individual performance in terms of the number of cards sold. To handle the workers more efficiently, the firm formed groups of two workers, but both workers in each group performed the same individual tasks. The individual workers were also paid the same hourly wage during the first phase of the experiment. The formation of groups of two with identical tasks and identical pay made it very plausible that the two group members became natural comparison agents for each other, which set the stage for the second phase of the experiment.

The firm implemented three treatments in the second phase of the experiment. In the baseline treatment, the hourly wage was not changed relative to the first phase; the firm cut both group members' hourly wage by 25% in the general wage-cut treatment; and it cut only one group member's hourly wage by 25% in the unilateral wage-cut treatment. This design enables us to study social comparison effects by comparing the unilateral wage-cut treatment with the general wage-cut and the baseline treatments. In particular, if workers who received a wage cut behave differently in the unilateral wage-cut treatment compared to the general wage-cut treatment, social comparison plays a role in the disadvantaged workers' behavior because the only difference between those two treatments is the *other* group member's wage. The same is true for the comparison between the workers in the unilateral wage-cut treatment not subject to a wage cut and those in the baseline treatment: if the former behave differently than the latter, social comparison plays a role in the advantaged workers' behavior. Finally, our design also allows us to examine the impact of general wage cuts by comparing effort provision in the general wage-cut treatment and in the baseline treatment.

We observe that workers reduced their performance by 15% in the general wage-cut treatment relative to the baseline treatment, while those workers whose wages were lowered in the unilateral wage-cut treatment reduced their performance by 34%—that is, by more than twice as much. This result suggests that social comparison can have powerful and large effects on individual behavior—although all workers who suffered a wage cut earned exactly the same wage, their performance was very different.

1. Several authors have pointed out these implications of social comparison effects. Among them are Akerlof and Yellen (1990); Camerer and Malmendier (2007); Krueger and Mas (2004).

Interestingly, workers whose wages were not cut in the unilateral wage-cut treatment neither reduced nor increased their performance relative to the baseline treatment.

The strong negative reaction of the workers whose wages were cut in the unilateral wage-cut treatment is not only restricted to performance. We also find that these workers tended to reduce the quality of their work most, and suspect that they might have engaged in employee theft. Finally, we complement the insights from the behavioral responses with survey evidence on pay satisfaction. The survey analysis reveals that the workers who suffered the wage cut in the unilateral wage-cut treatment were the least satisfied with the level of their pay; they were also significantly less satisfied than the workers in the general wage-cut treatment.

In view of the one-time nature of the firm's promotion and the clear communication to the workers that they were only being hired for this campaign and that no future employment at the firm was possible, reputation or repeated game effects are unlikely to explain our results. It is thus plausible that social comparison effects drove the large performance reduction in the unilateral wage-cut treatment. The workers whose wage was cut in this treatment may have perceived their cut as much more unkind than a general wage cut. This interpretation is consistent with theories of intention-based reciprocity Dufwenberg and Kirchsteiger 2004; Falk and Fischbacher 2006; Rabin 1993 when the notion of kindness—which is key in these models—includes payoff comparisons across different workers. In addition, we show in this paper that the theory of inequality aversion (Fehr and Schmidt 1999), which assumes that payoff comparisons between agents enter their utility functions, predicts all three empirical regularities: the reduction of performance in the general wage-cut treatment, the stronger reduction of performance in the unilateral wage-cut treatment for the workers whose wage was cut, and the absence of any change in performance for the workers whose wage was not cut in the unilateral wage-cut treatment. Thus, inequality aversion may also have played a role in workers' behavior.

Our paper is related to observational studies that report a negative correlation between wage inequality within firms and workers' self-reported performance (Levine 1993; Pfeffer and Langton 1993). However, it is difficult to derive unambiguous causal conclusions from these data because the direction of causality could be the reverse—it is possible that performance differences caused wage differences. The results of a recent study based on a survey of university employees' self-reported job satisfaction is nicely related to our findings (Card et al. 2012). This study found the employees' job satisfaction to be asymmetrically related to information about their colleagues' wages. Employees whose wage was below the median reported a decrease in satisfaction with their pay and job, and a higher willingness to search for a new job, while satisfaction and the willingness to search remained unchanged for those whose wage was above the median. This observation is in line with our findings from the unilateral wage-cut treatment. We find that the unilateral wage cut not only influenced self-reported job satisfaction, but importantly, output asymmetrically: while the workers whose wage was cut responded with a strong decrease in performance, the workers whose wage was not cut did not respond at all.

Our paper is also related to the literature on gift exchange in the labor market (Bellemare and Shearer 2009; Gneezy and List 2006; Kube, Maréchal, and Puppe 2012). These studies, however, do not examine the effect of wage comparisons among workers on effort because these studies focus on *bilateral* gift exchanges between workers and firms in the absence of *horizontal* social comparisons between workers.

A small number of recent laboratory studies have studied the role of social comparison information in effort provision. For example, Charness and Kuhn (2007) conducted a laboratory gift exchange experiment with one principal and two agents who differed in their effort productivity. They report that information about the other worker's wage did not influence effort. Gächter and Thöni (2010) implemented a similar treatment in which the laboratory workers did not differ in productivity. They find that wage comparison information reduces a subject's effort somewhat if the other worker earns considerably more (i.e., about twice as much). Finally, Clark, Masclet, and Villeval (2010) report results from a bilateral gift exchange experiment (with identical effort productivity across different firm–worker pairs) indicating that workers' ordinal wage rank in an experimentally constructed reference group has a positive impact on their effort.[2] Causal evidence on this topic in field settings and when subjects are unaware of being observed is very scarce, however. In the field, perhaps the most difficult problem is finding *exogenous* sources of wage variations that trigger social comparison. In view of the fact that some economists have raised skepticism regarding the generalizability of results from the laboratory to the field (for a discussion, see Falk and Heckman 2009), the demonstration of social comparison effects in the field seems important.[3]

Our results may help to explain several important economic phenomena. First, social comparison effects may cause a reduction in pay inequalities within the firm, which may be associated with smaller pecuniary performance incentives. Frank (1984), for example, shows that automobile and real estate salespeople were paid less than their marginal product—that is, their commission schemes were flatter than they should have been in the absence of social comparison effects. Second, transaction costs or property rights concerns alone may not drive firm boundaries (Holmström and Roberts 1998); social comparison may shape them as well. The success of a merger of different firms may strongly depend on whether workers compare their wages to those of the co-workers in the newly merged firm (Nickerson and Zenger 2008). This is because a merger may change the set of reference agents, implying that some of the incumbent workers may earn less than their new colleagues after the merger and thus reduce effort accordingly. As a final example, Akerlof and Yellen (1990) present a model in which

2. A possible explanation why the laboratory evidence indicates less wage comparison effects may be the lower saliency of the co-worker as a reference for comparison. There is typically no social interaction between workers in existing laboratory studies, except for the provision of the wage information. In contrast, the workers in our field experiment spent time together at group meetings before and after the shifts.

3. For a survey of field evidence on "nonstandard" economic preferences, see DellaVigna (2009).

social comparison may lead to involuntary unemployment. The key assumption in this model is that workers withdraw effort when their wage falls short of a "fair" wage, which depends on co-workers' wages. Social comparison thus provides a rationale for firms to pay efficiency wages to low-wage workers. Taken together, these examples suggest that social comparison among workers may have important implications in labor markets.

The remainder of this paper is organized as follows. The next section describes the economic environment in which the experiment took place. Section 3 presents our experimental design and Section 4 discusses various theories that predict social comparison effects. Section 5 describes our estimation strategy and presents the empirical results. Section 6 concludes the paper.

2. Experimental Environment

We collaborated with a small German firm that operates a nightlife web portal and sells a membership card permitting entrance to specific bars and nightclubs. Our field experiment is based on a short-term sales promotion aimed to make the firm's membership card more popular.[4] Because the firm had a relatively small budget and was not able to finance a large-scale promotion campaign from internal or external funds, it was willing to allow us to conduct the experiment in exchange for funding a large part of the campaign. We seized this opportunity and enabled the firm to run a short-term promotion campaign in two big cities and to hire almost one hundred workers exclusively for this campaign. The firm placed a job advertisement stipulating an hourly wage of about €10 on an online job search platform. To rule out potential influences of job applicants' past interactions with the firm's permanent staff, those applicants who knew a permanent employee were not hired. The firm's willingness to conduct an experiment involving wage cuts was also due to the fact that the promotion workers did not belong to the firm's permanent staff and that the firm had no intention of hiring any of the promotion workers after the campaign. The firm thus had no reason to fear effects on the morale of its own employees.

In each city, the sales promotion took place on two consecutive weekends. The promotion workers were hired to work two shifts on each of the two weekends—the first two shifts took place on Friday and Saturday evening of the first weekend, the third and fourth shift on Friday and Saturday evening of the next weekend. The workers were explicitly told that the promotion campaign would only take place on two consecutive weekends and that the firm was not hiring permanent employees. Thus, the promotion workers knew that none of them had a chance of becoming a permanent employee of the firm and that the campaign was strictly short-term. Another important aspect of the employment contract stipulated that workers had to compensate the firm with €30 if they missed a shift without giving a serious reason for their absence (e.g., medical

4. Sales promotions are an important industry: in the United States, about 8,000 companies generate an annual revenue of $12 billion in total (First Research 2012).

certificate for illness). The workers' task was to sell promotional cards at a price of €5 or in exchange for a customer's personal information that enabled the firm to contact him or her later. The personal information was recorded in a database and used to invite the customer to join the firm's nightlife web portal. Contacting the customers allowed the firm to verify ex post whether the customer information the workers provided was correct. The workers were, however, not informed that the validity of the customer information would be checked.

Before the first shift, the workers were invited to a training session during which they were made familiar with the objective and procedures of the promotion campaign, trained on how to approach potential customers, and provided with equipment and clothing. The workers were *randomly* allocated to groups of two in this training session and *randomly* assigned to be either "Worker 1" or "Worker 2" in a group. Group members had identical tasks and responsibilities and they stayed together for the entire employment period, making the co-worker in a group a natural reference agent.

Each group was assigned to a fixed location, which was either a shopping avenue or a large nightclub.[5] Because there are typically a large number of people (i.e. potential customers) at these locations, one worker alone would hardly have been able to fully exploit the available opportunities for approaching people. In fact, not even two workers were able to do this. This feature of the task provided a natural reason to form groups, while at the same time the two workers were assigned to different areas of the location in order to best "harvest" the potential customer pool. In this way we could form groups but still collect a meaningful performance measure—the number of sold cards—for each individual worker.

The workers were managed by group leaders who were permanent employees of the firm. Group leaders were responsible for two to three groups during a shift. Before the work shifts, they supplied the workers with promotional cards and other equipment at the particular locations.[6] During the shifts, they visited the locations in an unobtrusive way to check whether everything was going well and to assess various characteristics of the locations, such as the number of club visitors (i.e. potential customers).[7] They had a comprehensive set of instructions on how to communicate with the workers (see Online Appendix). In particular, they were instructed to treat all workers in the same manner and they were not allowed to motivate or rebuke particular workers. After the shifts, the group leaders met again with the workers to collect the revenues, the customer information workers collected, and the unsold cards. This provided a further occasion to establish the other worker in the group as the reference person, and also informed workers about the other's performance.

5. Working hours for shopping avenues were from 5 pm to 8 pm and for nightclubs from 11 pm to 2 am.

6. The meetings with the different groups were temporally and spatially separated to prevent them from meeting each other.

7. Group leaders had to ensure, for example, that the workers did not run out of cards to sell.

TABLE 1. Hourly wages in €.

Treatment	HH		LL		HL	
Worker	1	2	1	2	1	2
Pre-intervention	12	12	12	12	12	12
Post-intervention	12	12	9	9	12	9

Note: After the intervention, treatment HL is split into two treatment conditions, HLnocut (for worker 1) and HLcut (for worker 2).

3. Experimental Design

We implemented a difference-in-difference setup. The first week—which we call the pre-intervention phase—allows us to capture the workers' baseline performance and therefore to control for individual heterogeneity. This is important because many field experiments report substantial heterogeneity in unobserved worker characteristics and emphasize the importance of controlling for it (Shearer 2004; Fehr and Goette 2007). During the first week, all workers were paid €12 per hour. The group leader informed workers about the relevant wage during the joint meeting with their co-worker shortly before the beginning of the first shift (see Online Appendix).

In the second week—which we call the post-intervention phase—we randomly assigned groups to one of three treatments. In the unilateral wage-cut treatment, denoted by HL, Worker 2's hourly wage was cut by €3, while Worker 1 continued to earn the previous hourly wage of €12. Treatment HL thus consists of two groups, that of low-wage workers (HLcut) and that of high-wage workers (HLnocut).[8] In the general wage-cut treatment, denoted by LL, both Worker 1 and Worker 2 suffered a wage cut of €3. In contrast, both Worker 1 and Worker 2 continued to earn the previous hourly wage of €12 in the baseline treatment, denoted by HH (see Table 1 for a treatment overview). The group leader informed workers of the wage change during the joint meeting with their co-worker shortly before the beginning of the third shift. In the HL treatment, for example, group leaders used the following phrase: "All Worker 1's continue to earn €12 per hour while all Worker 2's receive €9 instead of €12 per hour. This was the manager's decision" (see Online Appendix).

We stratified treatment assignment by city, location type, and gender to maximize the statistical power of the experimental design.[9] To avoid contamination between treatments, we arranged the work schedule so that workers from different treatments did not meet; furthermore, a few workers who knew each other were assigned to the same treatment but to different groups within the treatment. We also exclusively formed

8. Note that even the low-wage workers earned an hourly wage of more than €10 on average. We thus did not violate the employment contract.

9. We conducted the experiment twice in two consecutive months, with the same number of workers and groups per treatment in each month. This also allowed us to stratify treatments across locations. If we had previously run treatments HH or LL at a location, we subsequently ran treatment HL, and vice versa.

same-gender groups to avoid gender-related confounds. Importantly, the workers were unaware that they were participating in an experiment to rule out experimenter demand effects.[10]

Following the experiment, the firm mailed the workers a short feedback survey. The survey asked a variety of questions related to the sales promotion. The first part included questions pertaining to the work clothes, training day, and feedback of the customers. The second part covered several aspects of the work environment. The question of key interest to us asked the survey participants to report their pay satisfaction on this job. Specifically, they had to indicate the extent to which they agreed with the statement "I am satisfied with my wage" with answer categories on a seven-point Likert scale ranging from "I do not agree at all" to "I completely agree". Further questions aimed to collect information on workers' intrinsic motivation, and wage expectations prior to starting the job. In an effort to achieve a high response rate, we rewarded workers' participation in the survey with €5. Two-thirds of the work force ultimately participated.

4. Hypotheses

How do workers respond to wage cuts? And how does their response depend on the wages paid to their co-workers? We designed our experiment so that—under the standard assumption that workers only have self-regarding preferences—they should not respond to a change in the hourly wage. In particular, we implemented the wage cut in such a way that we can minimize the possibility that low-paid workers perceive that they have a lower re-employment probability due to lower past performance or other personal characteristics. As mentioned above, it was made transparent at the beginning that there was no prospect of employment at the firm beyond the promotion campaign. In addition, the workers were randomly allocated to the position of Worker 1 and Worker 2 in the group upon entering the employment relationship. They thus perceived their allocation to the position of Worker 1 and Worker 2 as unrelated to any personal characteristics. The exogenous allocation to the position of Worker 1 or Worker 2 in combination with the announcement at the beginning of the post-intervention phase that *all* Workers 2 would receive a wage cut made it clear that the wage cut was unrelated to previous performance. Therefore, effort variations in response to the wage cut cannot be interpreted as a response to a change in pecuniary incentives.

These features of the experiment also imply that efficiency wage incentives in the spirit of Shapiro and Stiglitz (1984) cannot play a role. According to this model, workers may respond to a wage cut with an effort reduction because workers lose less rent at the lower wage if they were fired. This model does not predict, however, that wage cuts reduce effort more in treatment HL than in treatment LL because the model

10. While two workers contacted the manager to complain about the wage cut, none of the workers asked whether they were participating in an experiment.

does not rely on any sort of social comparison or interdependence between workers. Moreover, because the employment contract lasted for exactly two weekends, there was no threat of firing.

The independence of current wages, future wages, and future employment prospects from the workers' effort level thus implies that self-regarding workers have no reason to respond to the wage cut with a change in the effort level. If we assume, however, that workers are also motivated by fairness or reciprocity concerns, the predictions are different. Many social preference theories assume that workers compare their outcomes to fair reference outcomes in order to define measures of fairness or reciprocity (e.g., Dufwenberg and Kirchsteiger 2004; Falk and Fischbacher 2006; Rabin 1993). In other words, these theories typically rely in some form on assumptions about social comparison. According to intention-based reciprocity approaches, for example, workers may view a wage cut as a hostile act by the firm, and may consequently reduce their effort. The workers may view cutting only one worker's wage in a group as even more hostile, and this may further amplify the effort reduction.

Likewise, outcome-based considerations of fairness that are based on the comparison between the actual wage and a fair reference wage may also affect workers' fairness assessments and, therefore, their response to a wage cut. We show in the Online Appendix, for example, that the Fehr and Schmidt model of inequality aversion makes specific, testable predictions in our setting. We would like to mention, however, that other social preferences approaches may make similar predictions. Thus, we do not view evidence that is consistent with these predictions as evidence in favor of a particular social preference model.[11]

The Fehr–Schmidt model assumes that, in addition to their own material payoff, workers also want to avoid payoff inequalities between themselves and other workers, and the aversion against disadvantageous inequality ("envy") is stronger than the aversion against advantageous inequality ("compassion"). We present two versions of this model in the Online Appendix, one with linear cost of effort, and one with quadratic cost of effort. Because the linear model is highly tractable and intuitive, we base our exposition of the resulting predictions on this version. It generates the following hypotheses about performance changes from the pre- to post-intervention phase if workers are sufficiently inequality averse.

HYPOTHESIS 1. *Workers in treatment group HLcut decrease their performance after the wage cut more than those in treatment LL.*

HYPOTHESIS 2. *Workers in treatment group HLnocut do not change their performance relative to treatment HH after their co-workers suffer the wage cut.*

HYPOTHESIS 3. *Workers in treatment LL decrease their performance after the wage cut relative to workers in treatment HH.*

11. Therefore, our paper falls into the category of "Single Model" studies (Card, DellaVigna, and Malmendier 2011).

The intuition for these hypotheses is best developed if we start explaining Hypothesis 3 first. Payoff equality between the two group members exists if both workers put forward the same effort and receive the same wage. Payoff equality between the firm and a worker exists if the workers' material payoff (measured by wage minus effort cost) is equal to the firm's profit from that worker (measured by revenue minus wage). In the equilibrium of treatment HH, both group members choose an identical effort level that equalizes the payoffs across all three parties. Since the wage is lower in treatment LL than in HH, the firm would have a payoff advantage if workers put forward the same effort in LL as in HH. After the wage cut, the two group members will therefore re-establish equality between themselves and the firm by reducing effort in treatment LL.

Next, let us explain Hypothesis 1. Why does the Worker 2 with low pay in HL provide even less effort than workers in LL? Suppose that Worker HLnocut—whose wage is not cut—provides the same effort as in HH, and Worker HLcut puts forward the same effort as in LL. Recall that the same effort as in LL renders Worker HLcut's material payoff equal to the firm's profit from this worker, but Worker HLcut now suffers from envy with respect to Worker HLnocut. By further reducing the effort (cost) below the level of treatment LL, Worker HLcut can reduce the payoff disadvantage relative to Worker HLnocut, but this also generates a payoff advantage relative to the firm. However, because envy looms larger than compassion, it is better for Worker HLcut to reduce envy with regard to Worker HLnocut and accept the small compassion cost with regard to the firm. Worker 2's effort in treatment HL will thus be lower than in LL.

Why does Worker HLnocut provide the same effort as workers in HH? The intuition for Hypothesis 2 goes as follows. By increasing the effort above the HH level (i.e., by reducing her own material payoff), Worker HLnocut would reduce the payoff advantage with regard to Worker HLcut, but this would also cause a payoff disadvantage with regard to the firm. Because envy looms larger than compassion, a higher effort than in HH would therefore imply a large envy cost with regard to the firm and a small compassion gain with regard to the worker. As a consequence, Worker HLnocut prefers providing the same effort as workers in treatment HH.

While the linear effort cost term is practical, in many respects it is also unrealistic because it is likely that workers in our experiment find it more and more difficult to generate additional sales from a limited number of passers-by. In addition, because the marginal cost of effort is always lower than the marginal revenue, the efficient outcome would be infinite effort, or an arbitrary exogenous maximum effort level. We therefore also discuss what happens to our hypotheses if we exchange the linear effort cost term with a quadratic one in the Online Appendix. While the solutions to this model become more difficult to interpret, we can state the same qualitative predictions as in the linear case.

Intention-based fairness concerns can also predict our results. As we show in the Online Appendix, if a simple model of reciprocity (Cox, Friedman, and Gjerstad 2007) is extended to include social comparison between workers, this model generates similar hypotheses to the outcome-based model, except for the unaffected worker in treatment

TABLE 2. Descriptive statistics.

Treatment Group	HH	LL	HL: HLnocut	HL: HLcut	Total
Workers (#)	24	24	24	24	96
Female (#)	18	18	19	19	74
Age, mean (s.d.)	20.5	21.2	20.2	21.1	20.7
	(2.1)	(2.5)	(1.9)	(2.7)	(2.3)
Reported sick (#)	1	1	0	1	3
Performance, mean (s.d.)					
Pre-intervention	20.8	22.4	24.3	22.0	22.4
	(6.0)	(4.7)	(9.1)	(9.9)	(7.7)
Post-intervention	22.8	21.4	26.5	18.4	22.3
	(9.4)	(4.6)	(10.2)	(11.3)	(9.6)
% workers with $y_{post} > y_{pre}$	47.8	30.4	58.3	17.4	38.7

HL, who is predicted to work harder as a response to the other worker's reduced wage. Taken together, both types of concerns—reciprocity or inequality aversion—predict that (i) workers will respond to the general wage cut with an effort reduction and (ii) that the effort reduction will be even stronger in the event of a unilateral wage cut. Considerations of fairness and reciprocity are, of course, also one possible reason why workers provide positive effort in the absence of pecuniary performance incentives. Another possibility is that they may be intrinsically motivated to perform the task.

5. Estimation Strategy and Results

Table 2 shows the descriptive statistics of our experiment. The sample consisted of 96 workers in 48 groups. None of the workers quit after the wage cut.[12] This is not surprising considering that workers were contractually obliged to compensate the firm with €30 if they missed a shift. About three-quarters of the workers were women, and the age ranged from 18 to 28 years. All but one were German citizens. On average, the workers exchanged 22.3 cards per shift and the customers provided personal information instead of paying €5 for the card in 97.9% of these exchanges. Since the firm verified the correctness of the customer information, we know that this information was wrong only in 2.4% of the cases.

12. Three workers fell ill before any wage cut was announced and missed both post-intervention shifts. These workers were substituted with replacement workers who came in at the time when the treatments were implemented and were treated in exactly the same way as the replaced workers would have been treated. They were also informed about the wage changes in their respective work group. We exclude replacement workers from the analysis, but our results do not change if we include them.

We balanced the treatments over cities, location types (i.e., street versus nightclub), and gender. The pre-intervention phase enables us to control for worker heterogeneity by including individual fixed effects into our regression. Because workers were always assigned to the same location, individual fixed effects not only capture time-invariant differences across workers but also location-specific factors. In addition, we are able to include a proxy for the demand potential for cards as a control variable. Recall that the group leaders visited "their" groups several times during a shift. These visits allowed them to assess the number of potential customers for each group and shift on a five-point scale (–2 = very low, –1 = moderately low, 0 = normal, 1 = moderately high, 2 = very high). We use this assessment to control for potential variations in demand across groups, shifts, and treatments. We therefore also report a specification in which we control for demand potential.[13]

Using the balanced panel data structure, we estimate the following difference-in-difference regression model with each worker as a panel unit and each group as an independent observation:

$$\log\left(y_{ikt}\right) = \alpha + \varphi_i + \delta D_{kt} + \beta_1 t + \beta_2 \left(LL_i \times t\right) + \beta_3 \left(\text{HLnocut}_i \times t\right) + \beta_4 \left(\text{HLcut}_i \times t\right) + \varepsilon_{ikt}.$$

$\log\left(y_{ikt}\right)$ denotes the logarithm of the average performance of worker i in group k and weekend t. We use the total number of "correct" exchanges as a measure of performance—that is, we exclude those sales associated with wrong customer information from our performance measure. The constant α captures average pre-intervention performance, φ_i represents individual fixed effects, D_{kt} controls for differences in demand potential, and t is a dummy variable that equals 1 if the observation is from the post-intervention phase and zero otherwise. The coefficient of the variable t thus measures the change in performance in the post-intervention phase relative to the average pre-intervention performance. The coefficients of the three interaction terms $LL_i \times t$, $\text{HLnocut}_i \times t$, and $\text{HLcut}_i \times t$ are our difference-in-difference estimators of interest, as they capture the differences across treatments (relative to control treatment HH) in the performance change between pre- and post-intervention phase. Finally, ε_{ikt} is the idiosyncratic error term, which we allow to be clustered over groups. As expected, the individual fixed effects are significant ($p < 0.01$, F-test).

5.1. Main Results

Based on the estimation strategy described above, we can report the following result with regard to Hypothesis 1, which involves a comparison between the unilateral and the general wage-cut treatment.

13. Note that we cannot reject the null hypothesis that demand was equally distributed across treatments ($p = 0.43$, Kruskal–Wallis test). However, controlling for demand potential enables a more precise estimate of the treatment effects. All p-values reported in this paper are two-sided.

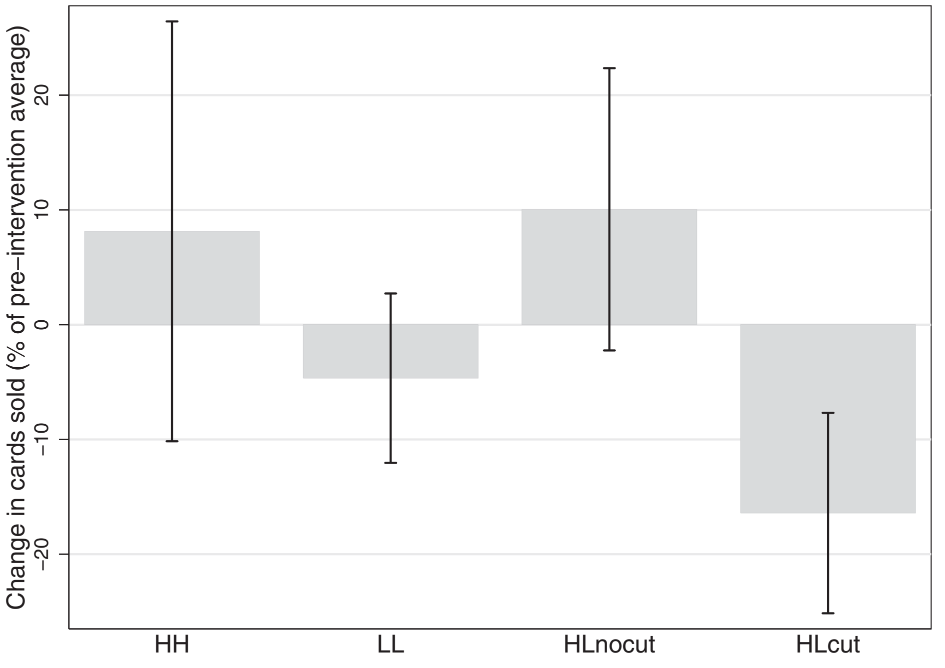


FIGURE 1. Average treatment effects on individual work performance. Absolute change in performance (number of cards), from pre- to post-intervention, scaled as a fraction of the average pre-intervention performance (= 22.4 cards per shift), in treatments HH (baseline), LL (general wage cut), HLnocut (spared workers in the unilateral wage-cut treatment), and HLcut (unilateral wage cut). Error bars show 95% confidence intervals; for HH and LL, we bootstrap with group clusters, 1,000 repetitions. Replacement workers are excluded.

RESULT 1. A *unilateral* wage cut causes a much stronger decrease in performance than a general wage cut.

We illustrate this result in Figure 1, which presents the average performance changes from the pre- to post-intervention phase by treatment in percent of the pre-intervention average. A comparison between treatment LL and HLcut shows that while performance decreased by 5% between the two phases in treatment LL, workers reduced their performance by 16% in treatment HLcut. Thus, although workers in treatment HLcut suffered the same monetary loss as workers in treatment LL, their drop in performance was more than three times as great. This difference is already significant in the absence of further controls ($p < 0.05$, rank-sum test). In addition, workers in HLcut also reduced their performance compared to their co-workers in HLnocut who did not suffer the pay cut ($p < 0.01$, signed-rank test). Finally, the regressions in Table 3 further corroborate this finding. The first column in Table 3 represents the previous model without controlling for demand potential, while we control for

TABLE 3. Treatment effects on individual work performance.

Dependent variable: log(cards sold)	(1)	(2)
HLcut × Post-intervention	–0.306***	–0.342***
	(0.103)	(0.101)
LL × Post-intervention	–0.106	–0.145*
	(0.090)	(0.079)
HLnocut × Post-intervention	0.034	0.015
	(0.095)	(0.089)
Post-intervention	0.058	0.090
	(0.080)	(0.074)
Demand potential		0.117***
		(0.034)
Constant	3.057***	3.048***
	(0.015)	(0.015)
Worker fixed effects	Yes	Yes
N	189	179
Adj. R^2	0.202	0.312
Hypothesis tests		
H1 (HLcut × post = LL × post)	$p = 0.01$	$p < 0.01$
H2 (HLnocut × post = 0)	$p = 0.72$	$p = 0.87$
H3 (LL × post = 0)	$p = 0.24$	$p = 0.07$

Notes: OLS estimates. Robust standard errors corrected for clusters on groups in parentheses. The constant term captures the average performance in the pre-intervention phase. Hypothesis tests are Wald tests.
*Significant at 10%; ***significant at 1%.

demand in the second column.[14] Regardless of whether we control for demand potential, the coefficients of "HLcut × Post-intervention" in the two regressions are highly significant; they indicate that the unilateral wage cut reduced HLcut workers' performance by 31% to 34% compared to control treatment HH ($p < 0.01$, Wald test). Moreover, a comparison between the coefficient of "LL × Post-intervention", which measures the impact of the wage cut in LL, and the coefficient of "HLcut × Post-intervention", which measures the impact of the wage cut in HLcut, shows that the reduction in HLcut is about 20 percentage points larger than that in treatment LL—a difference that is significant ($p = 0.01$ and $p < 0.01$ for models (1) and (2), respectively, Wald tests).

Workers were assigned the role of Worker 1 and Worker 2 at the very beginning. However, some workers in treatment HL may have erroneously believed that the wage cut was a consequence of their low previous performance relative to that of their co-worker. If this were the case, workers should have reacted differently to the unilateral wage cut depending on whether they had performed better or worse than their co-worker before the intervention. To test this, we divided workers in treatment group HLcut into

14. When the demand potential is included in the regression, the number of observations decreases by 10 because the demand measure is missing for some shifts. The group leaders forgot to collect this information in these cases.

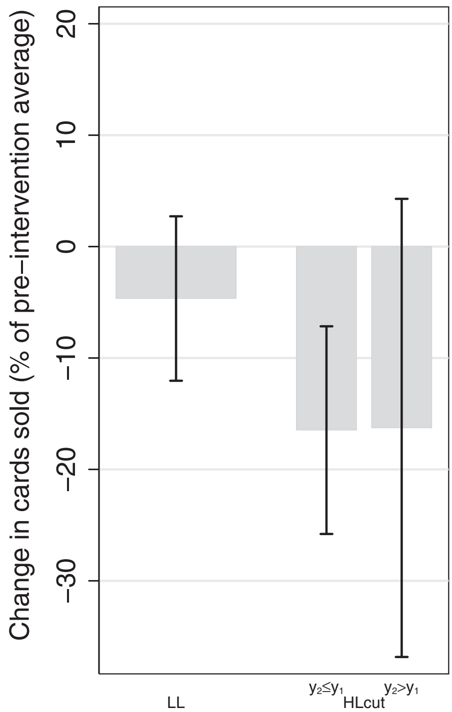


FIGURE 2. Treatment effect on HLcut workers, split by pre-intervention performance relative to co-worker. The left bar shows the treatment effect on worker performance in the LL group, for comparison. The two smaller bars on the right show how workers in the HLcut condition reacted to the intervention. The left small bar, labeled "$y_2 \leq y_1$" shows how workers reacted who, in the pre-intervention shifts, had a *lower or equal* output than that of their co-worker; the right small bar, labeled "$y_2 > y_1$" shows how workers' performance changed in the HLcut condition if they had a *higher* output than that of their co-worker in the pre-intervention shifts. Vertical axis: change in number of cards from pre- to post-intervention, scaled as a percentage of average pre-intervention performance = 22.4 cards per shift. Error bars show 95% confidence intervals; for LL, we bootstrap with group clusters, 1,000 repetitions. Replacement workers are excluded.

those whose pre-intervention performance was higher than or equal to their co-worker's performance and those whose performance was lower (see Figure 2). We find that both groups significantly reduced their performance in the post-intervention phase ($p =$ 0.07 and 0.02, respectively, signed-rank tests). In addition, the performance reduction is not significantly different between the two groups ($p = 0.49$, rank-sum test).

Next we turn to the test of Hypothesis 2, which involves a comparison between the workers whose wage was not cut in the unilateral wage-cut treatment and the workers in the baseline treatment.

RESULT 2. A *unilateral* wage cut has no effect on the group members whose wage is not cut.

First support for Result 2 comes from Figure 1. The comparison between HH and HLnocut shows that, on average, workers in treatment HH displayed a performance increase between the pre- and the post-intervention phase, possibly due to learning effects. The workers in treatment HLnocut showed almost the same average performance increase as workers in treatment HH ($p = 0.37$, rank-sum test). In addition, both regression estimates of "HLnocut × Post-intervention" in Table 3 confirm the finding that workers in treatment HLnocut did not perform differently than those in treatment HH ($p = 0.72$ and $p = 0.87$, Wald tests).[15] Thus, performing the same task but earning more than one's co-worker was not associated with higher effort.

Finally, we turn to Hypothesis 3, which involves a comparison between the general wage cut and the baseline treatment.

RESULT 3. A *general* wage cut causes a decrease in performance.

We illustrate this result again in Figure 1. Performance in control treatment HH *increased* by 8% relative to the pre-intervention average. In contrast, performance in treatment LL *decreased* by 5% after the wage cut, but without further controls, this difference is not significant ($p = 0.37$, rank-sum test). We therefore turn to the regression in Table 3 which controls for demand potential. The coefficient of demand is highly significant ($p < 0.01$, Wald test) and has the expected sign: the more potential customers present at a location, the more cards the workers were able to sell. Moreover, the adjusted R^2 increases substantially if we include the demand variable in the regression. Overall, these facts indicate that the number of potential customers is an important control variable that needs to be included in the regression.

The coefficient on "LL × Post-intervention" has a negative sign in both specifications reported in Table 3. If we control for demand potential, performance dropped by roughly 15% compared to control treatment HH. This constitutes a marginally significant performance decrease ($p = 0.07$, Wald test).

5.1.1. Cumulative Distributions of Performance Changes. Figure 3 shows the cumulative distribution functions of the performance changes from pre- to post-intervention phase by treatment in percent of the pre-intervention average. The upper panel shows that in treatment HLcut the entire distribution function of the performance change is shifted towards a lower performance compared to treatment LL.

The lower panel of Figure 3 shows that the cumulative distribution functions for the HH and HLnocut treatment are close together, which suggests that by and large workers performed similarly in these two treatments.[16]

5.1.2. Effect on Firm Profit. The adverse behavioral reactions to the wage cuts raise the question whether the cuts were profitable for the firm. We therefore calculate group-level profits as revenue minus wage costs.[17] For the revenue, we assume that

15. Note that the lowest detectable difference between HH and HLnocut workers on the 10% significance level is a performance difference of about 11 percentage points.

16. We find two outliers. One HLnocut worker exhibits a large performance decrease when the co-worker suffered a wage cut, and one HH worker displays a large performance increase between the two phases.

17. We neglect additional costs, such as fixed costs (e.g., office rent) and other variable costs (e.g., hiring and training costs).

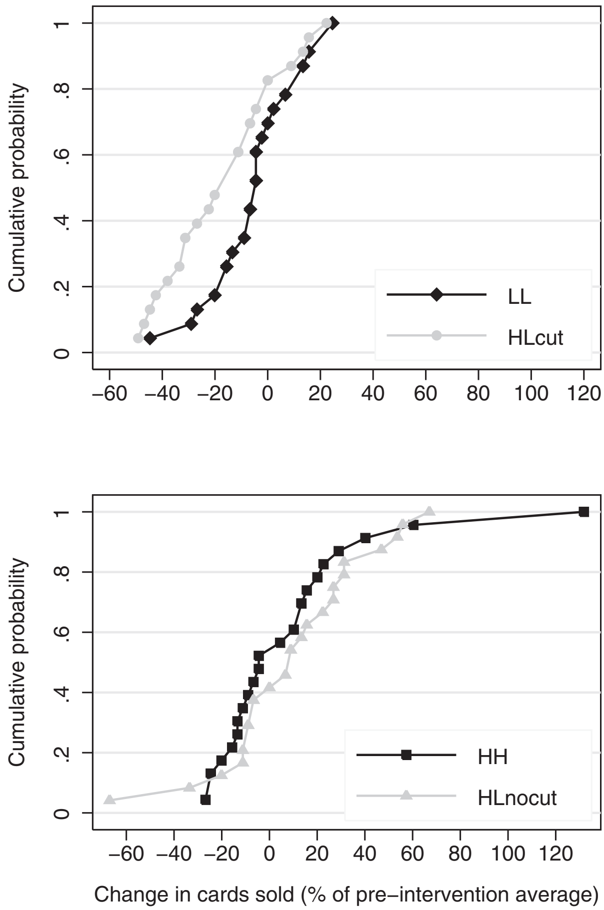
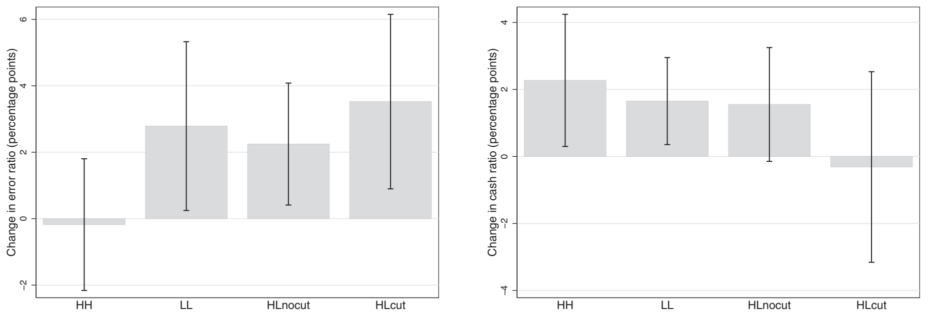


FIGURE 3. Cumulative distributions of the change in individual work performance. Cumulative distributions of the absolute change in performance (number of cards), from pre- to post-intervention, scaled as a fraction of average pre-intervention performance (= 22.4 cards per shift). Upper panel: LL (general wage cut) and HLcut (unilateral wage cut). Lower panel: HH (baseline) and HLnocut (spared workers in the unilateral wage-cut treatment).

exchanges for customer information were equally valuable for the firm as cash sales—that is, selling the cards at a price of €5. Accordingly, in the pre-intervention phase, where workers exchanged an average of 22.4 cards per shift, they generated a revenue of $22.4 \times €5 = €112$ at wage costs of $3 \times €12 = €36$. This corresponds to a group profit of €152 per shift.

To calculate the revenue in the post-intervention phase, we use the estimates from column (2) in Table 3. Accordingly, the group revenue in treatment HH increased to $109\% \times 2 \times €112 = €244$, resulting in a group profit of €172 per shift. The corresponding group profit in treatment LL is $(109\% - 14.5\%) \times 2 \times 22.4 \times €5 - 2\times 3 \times €9 = €158$. In treatment HL, an HLnocut worker generated $(109\% + 1.5\%) \times$



(a) Change in error rate (b) Change in cash ratio

FIGURE 4. Treatment effect on fraction of erroneous addresses and fraction of cash sales of total exchanges. Change in the ratio of exchanges with incorrect customer information (panel (a)) and the ratio of cash sales to the total amount of exchanges (panel (b)) from pre- to post-intervention phase by treatment in percentage points. Treatment HH is the baseline condition, treatment LL is the general wage-cut condition, treatment group HLnocut are the spared workers in the unilateral wage-cut condition, and treatment group HLcut are the workers who suffered the wage cut in the unilateral wage-cut condition. Error bars show 95% confidence intervals; we bootstrap with group clusters, 1,000 repetitions for HH and LL. Replacement workers are excluded.

22.4 × €5 – 3 × €9 = €97, and an HLcut worker (109% – 34.2%) × 22.4 × €5 − 3 × €9 = €57, yielding a group profit of €154. Based on this calculation, we find that the unilateral wage-cut treatment generated the lowest profit, also lower than the general wage-cut treatment. Thus, the small compensating effect on the workers whose wage was not cut in the unilateral wage-cut treatment did not offset the large performance decrease of the workers who suffered the pay cut in this condition.

5.2. Work Quality and Employee Theft

In addition to the amount of "correct" exchanges as a measure of performance, we also examine the impact of the wage cuts on the ratio of exchanges that were associated with "incorrect" customer information. While we cannot discriminate between customers who gave incorrect information and workers who faked customer information in order to harm the firm, this measure may still serve as a proxy for the quality of work. An example of lower work quality caused by disgruntled workers is provided by Krueger and Mas (2004). They show that the quality of Bridgestone/Firestone tires decreased significantly after the management announced a decrease in new hires' wages. In contrast, Kube, Maréchal, and Puppe (2013) provide evidence in the context of a book entry task of higher quality after a wage cut. They attribute the lower amount of faulty book entries to the lower typing speed. Hence, the existing evidence on the impact of wage cuts on quality of work is inconclusive.

Figure 4(a) displays the changes in the error ratio from pre- to post-intervention phase by treatment in percentage points. The figure reveals a similarly large increase in the error ratio for all workers who either suffered or witnessed a wage cut. However, when we consider the relevant treatment comparisons, only the comparison between treatment HH and treatment LL achieves weak statistical significance (HH versus LL: $p = 0.09$, HH versus HLnocut: $p = 0.20$, LL versus HLcut: $p = 0.78$, rank-sum tests).

In addition to reducing the quality of work, workers could also harm the firm by pocketing the cash sales and faking customer information in such a way that cash sales would look like exchanges for customer information. We thus further investigate the impact of the wage cuts on employee theft. Figure 4b presents the changes in the ratio of cash sales to the total amount of exchanges from pre- to post-intervention phase by treatment in percentage points. In contrast to LL workers, HLcut workers exhibited a decrease in the cash ratio ($p = 0.09$, rank-sum test). This finding provides suggestive evidence that, in addition to a lower performance, HLcut workers also negatively reciprocated the wage cut by stealing money from the firm. For the remaining treatment comparisons, we cannot reject the hypothesis that theft was statistically the same (HH vs. LL: $p = 0.48$, HH vs. HLnocut: $p = 0.30$, rank-sum tests).

5.3. Pay Satisfaction

We complement the behavioral insights from our experiment with survey data on pay satisfaction. Table 4 shows in a regression framework how workers' pay satisfaction varies across treatments. Because some workers did not participate in the survey, we additionally estimate a model where we control for differences in age, gender, wage expectations prior to the job, as well as intrinsic motivation (see column (2) of Table 4).[18] The results on pay satisfaction are qualitatively equivalent to the behavioral effects. Compared to the general wage-cut treatment, the unilateral wage-cut treatment causes lower pay satisfaction for the affected workers ($p = 0.05$, Wald test); there is no difference in pay satisfaction between those whose wage was not cut in treatment HL relative to the workers in the baseline treatment HH ($p = 0.98$, Wald test), and compared to the baseline treatment HH the general wage cut reduced pay satisfaction a little, though not significantly ($p = 0.32$, Wald test). These findings point to a fairness-based explanation as to why there was such a drop in performance in the treatment group HLcut.

6. Conclusions

This paper reports evidence from a field experiment on the effects of wage cuts and whether these effects depend on the wages paid to co-workers. We collaborated with a firm that hired workers for a one-time sales promotion. The firm formed groups of two, and the workers in a group had identical individual tasks. Group members were paid the same hourly wage in the first phase of the experiment. We subsequently implemented

18. 13 workers in treatment group HH, 19 workers in LL, 17 workers in HLnocut, and 12 workers in HLcut participated in the follow-up survey.

TABLE 4. Treatment effects on individual pay satisfaction.

Dependent variable: pay satisfaction	(1)	(2)
HLcut	–1.051**	–1.523***
	(0.413)	(0.421)
LL	–0.753*	–0.500
	(0.433)	(0.443)
HLnocut	–0.149	–0.008
	(0.332)	(0.410)
Constant	6.385***	5.658**
	(0.179)	(2.604)
Controls		
Age	No	Yes
Gender	No	Yes
Wage expectations	No	Yes
Intrinsic motivation (three items)	No	Yes
N	61	58
Adj. R^2	0.047	0.130

Notes: OLS estimates. Robust standard errors in parentheses. Pay satisfaction is measured as the extent to which workers agreed to the statement "I am satisfied with my wage" with answer categories on a seven-point Likert scale ranging from "not at all agree" to "completely agree". The constant term captures the average pay satisfaction in treatment HH. The number of observations in model 2 is smaller because of item nonresponse from three workers. Intrinsic motivation is measured using the following three regressors: the extent to which workers agreed to the statements "I felt co-responsible for the success of the sales promotion" and "I am only satisfied with myself if I do a good job" with answer categories on a seven-point Likert scale ranging from "not at all agree" to "completely agree", and pre-intervention performance.
*Significant at 10%; **significant at 5%; ***significant at 1%.

three treatments that enabled us to examine social comparison effects and the effects of wage cuts on performance. We find that cutting both group members' wages by 25% caused a decrease in performance by 15% relative to the baseline treatment. In contrast, cutting only one group member's wage by 25% caused a decrease in the performance of the affected workers by 34%, i.e., by more than twice as much as in the general wage-cut treatment, even though pay levels and the size of the pay cuts were identical in both treatments. These workers were also less satisfied with their wage and may have engaged in more employee theft than the other workers, including the groups where both workers received a wage cut. The unilateral wage cut had an asymmetric effect on group members, that is, those workers whose wages were not cut neither decreased nor increased their performance relative to the baseline treatment.

Our results provide causal evidence for the role of social comparison among workers in effort provision. Such social comparison effects may have important consequences for a wide range of economic issues. They may induce employers to compress the wage distribution across workers, they may also be an important factor in merger and outsourcing decisions, and, as demonstrated in Akerlof and Yellen (1990), they may even lead to involuntary unemployment.

Although we provide evidence that social comparison *can* have large effects on worker behavior in a field setting, there are some limitations to the generalizability of our results. For example, our study was conducted within a limited time span. It could be that workers' decline in performance might recover after some time, meaning that if we had followed workers for more than two days, the treatment effect may have decreased. Another limitation of our design is that we did not provide an explanation for the wage cut. A sensible explanation may attenuate the detrimental effect on performance. Finally, we paid workers a fixed hourly wage; this means we cannot immediately translate our results to a piece-rate environment. With such a schedule, workers would still have an incentive to produce output. However, as suggested by our findings on work quality and stealing, workers may react in other, unincentivized dimensions of their work.

References

Akerlof, George A. and Janet L. Yellen (1990). "The Fair Wage–Effort Hypothesis and Unemployment." *Quarterly Journal of Economics*, 105, 255–283.

Bellemare, Charles and Bruce Shearer (2009). "Gift Giving and Worker Productivity: Evidence from a Firm-Level Experiment." *Games and Economic Behavior*, 67, 233–244.

Camerer, Colin F. and Ulrike Malmendier (2007). "Behavioral Economics of Organizations." In *Behavioral Economics and Its Applications*, edited by Peter Diamond and Hannu Vartiainen. Princeton University Press, pp. 233–290.

Card, David, Stefano DellaVigna, and Ulrike Malmendier (2011). "The Role of Theory in Field Experiments." *Journal of Economic Perspectives*, 25(3), 39–62.

Card, David, Alexandre Mas, Enrico Moretti, and Emmanuel Saez (2012). "Inequality at Work: The Effect of Peer Salaries on Job Satisfaction." *American Economic Review*, 102(6), 2981–3003.

Charness, Gary and Peter Kuhn (2007). "Does Pay Inequality Affect Worker Effort? Experimental Evidence." *Journal of Labor Economics*, 25, 693–723.

Clark, Andrew E., David Masclet, and Marie Claire Villeval (2010). "Effort and Comparison Income: Experimental and Survey Evidence." *Industrial and Labor Relations Review*, 63, 407.

Cox, James C., Daniel Friedman, and Steven Gjerstad (2007). "A Tractable Model of Reciprocity and Fairness." *Games and Economic Behavior*, 59, 17–45.

DellaVigna, Stefano (2009). "Psychology and Economics: Evidence from the Field." *Journal of Economic Literature*, 47, 315–372.

Duesenberry, James S. (1949). *Income, Saving and the Theory of Consumer Behavior*. Harvard University Press.

Dufwenberg, Martin and Georg Kirchsteiger (2004). "A Theory of Sequential Reciprocity." *Games and Economic Behavior*, 47, 268–298.

Falk, Armin and Urs Fischbacher (2006). "A Theory of Reciprocity." *Games and Economic Behavior*, 54, 293–315.

Falk, Armin and James J. Heckman (2009). "Lab Experiments Are a Major Source of Knowledge in the Social Sciences." *Science*, 326, 535–538.

Fehr, Ernst and Lorenz Goette (2007). "Do Workers Work More if Wages Are High? Evidence from a Randomized Field Experiment." *American Economic Review*, 97(1), 298–317.

Fehr, Ernst and Klaus M. Schmidt (1999). "A Theory of Fairness, Competition, and Cooperation." *Quarterly Journal of Economics*, 114, 817–868.

First Research (2012). "Sales Promotion Services Industry Profile." www.firstresearch.com/industry-research/Sales-Promotion-Services.html, retrieved 17 June 2012.

Fliessbach, K., B. Weber, P. Trautner, T. Dohmen, U. Sunde, C. E. Elger, and A. Falk (2007). "Social Comparison Affects Reward-Related Brain Activity in the Human Ventral Striatum." *Science*, 318, 1305–1308.

Frank, Robert H. (1984). "Are Workers Paid their Marginal Products?" *American Economic Review*, 74(4), 549–571.

Gächter, Simon and Christian Thöni (2010). "Social Comparison and Performance: Experimental Evidence on the Fair Wage-Effort Hypothesis." *Journal of Economic Behavior and Organization*, 76, 531–543.

Gneezy, Uri and John A. List (2006). "Putting Behavioral Economics to Work: Testing for Gift Exchange in Labor Markets Using Field Experiments." *Econometrica*, 74, 1365–1384.

Holmström, Bengt and John Roberts (1998). "The Boundaries of the Firm Revisited." *Journal of Economic Perspectives*, 12(4), 73–94.

Kochan, Thomas A. and Thomas A. Barocci (1985). *Human Resource Management and Industrial Relations: Text, Readings, and Cases*. Little, Brown.

Krueger, Alan B. and Alexandre Mas (2004). "Strikes, Scabs, and Tread Separations: Labor Strife and the Production of Defective Bridgestone/Firestone Tires." *Journal of Political Economy*, 112, 253–289.

Kube, Sebastian, Michel André Maréchal, and Clemens Puppe (2012). "The Currency of Reciprocity: Gift Exchange in the Workplace." *American Economic Review*, 102(4), 1644–1662.

Kube, Sebastian, Michel André Maréchal, and Clemens Puppe (2013). "Do Wage Cuts Damage Work Morale? Evidence from a Natural Field Experiment." *Journal of the European Economic Association*, 11, 853–870.

Levine, David I. (1993). "What Do Wages Buy?" *Administrative Science Quarterly*, 38, 462–483.

Luttmer, Erzo F. P. (2005). "Neighbors as Negatives: Relative Earnings and Well-Being." *Quarterly Journal of Economics*, 120, 963–1002.

Nickerson, Jack A. and Todd R. Zenger (2008). "Envy, Comparison Costs, and the Economic Theory of the Firm." *Strategic Management Journal*, 29(13), 1429–1449.

Pfeffer, Jeffrey and Nancy Langton (1993). "The Effect of Wage Dispersion on Satisfaction, Productivity, and Working Collaboratively: Evidence from College and University Faculty." *Administrative Science Quarterly*, 38, 382–407.

Rabin, Matthew (1993). "Incorporating Fairness into Game Theory and Economics." *American Economic Review*, 83(5), 1281–1302.

Shapiro, Carl and Joseph E. Stiglitz (1984). "Equilibrium Unemployment as a Worker Discipline Device." *American Economic Review*, 74(3), 433–444.

Shearer, Bruce (2004). "Piece Rates, Fixed Wages and Incentives: Evidence from a Field Experiment." *Review of Economic Studies*, 71, 513–534.

Slichter, Sumner H. (1920). "Industrial Morale." *Quarterly Journal of Economics*, 35, 36–60.

Veblen, Thorstein (1899). *The Theory of the Leisure Class: An Economic Study of Institutions*. Macmillan.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's website:

Supporting Material for Social Comparison and Effort Provision