DO WAGE CUTS DAMAGE WORK MORALE? EVIDENCE FROM A NATURAL FIELD EXPERIMENT

Sebastian Kube
University Of Bonn

Michel André Maréchal
University Of Zurich

Clemens Puppe
Karlsruhe Institute of Technology

Abstract
Employment contracts are often incomplete, leaving many responsibilities subject to workers' discretion. High work morale is therefore essential for sustaining voluntary cooperation and high productivity in firms. We conducted a field experiment to test whether workers reciprocate wage cuts and raises with low or high work productivity. Wage cuts had a detrimental and persistent impact on productivity, reducing average output by more than 20%. An equivalent wage increase, however, did not result in any productivity gains. The results from an additional control experiment with high monetary performance incentives demonstrate that workers could still produce substantially more output, leaving enough room for positive reactions. Altogether, these results provide evidence consistent with a model of reciprocity, as opposed to inequality aversion. (JEL: C93, C30)

1. Introduction

> "Dissatisfaction of the workers with their treatment by the management is to be counted among the most important causes of low morale, for it is common knowledge that men tend to hold back and to do little as possible for those against whom they feel a grievance."
>
> Sumner H. Slichter (1920, p. 40)

Do wage cuts damage work morale? In the presence of incomplete contracts, work morale crucially determines the success of employment relationships. Work

The editor in charge of this paper was Stefano DellaVigna.

Acknowledgments: An earlier version of this paper circulated under the title "Putting Reciprocity to Work—Positive versus Negative Responses in the Field" and formed chapter three of Maréchal's Ph.D. thesis. We thank Veronica Block, Carsten Dietz, Philipp Stroehle, and especially Michael Weingärtner for providing excellent research assistance and Johannes Kaiser for programming the computer application. We are grateful to Stefano DellaVigna, several anonymous referees, Nicholas Bardsley, Monika Bütler, Alain Cohn, Simon Evenett, Armin Falk, Ernst Fehr, Uri Gneezy, Sally Gschwend, John List, Rupert Sausgruber, Christian Thöni, as well as the audiences at ESEM Vienna, ESEM Budapest, ESA Rome, IMEBE Malaga, University of St. Gallen, University of Innsbruck, University of Bonn, University of Karlsruhe, UC San Diego, UC Santa Cruz, UC Santa Barbara, UC Los Angeles, and UC Berkeley for very helpful comments.

E-mail: kube@uni-bonn.de (Kube); marechal@iew.uzh.ch (Maréchal); clemens.puppe@kit.edu (Puppe)

Journal of the European Economic Association August 2013 11(4):853–870

DOI: 10.1111/jeea.12022

morale reflects the degree to which workers voluntarily co-operate and contribute to the employer's goals. Scholars have argued that work morale is sensitive to the relationship between the workers' actual wage and a reference wage (e.g. see Bewley 1999). Positive and negative deviations from the reference wage are interpreted as kind or unkind; employees then reciprocate by exerting higher or lower effort, respectively. While this theoretical argument has a long tradition in economics (e.g. see Akerlof 1982 or Slichter 1920), corresponding field evidence is scarce—in particular with respect to the impact of wage cuts.

This paper sheds light on the interplay between wages and work morale in employment relations. We conducted a controlled field experiment and tested the extent to which workers react to wage cuts and corresponding pay raises. A university library hired workers to catalogue books for a *limited* time period (i.e. excluding any possibility of re-employment) and announced a *projected* wage of €15 per hour. We actually paid this amount in our baseline treatment, and it serves as an exogenous reference point for wage expectations. In our main treatment, we inform subjects immediately before work begins that we will only pay them €10 per hour. In a second treatment, we do the opposite and communicate a pay rise from €15 to €20 per hour in order to explore asymmetries between the impact of wage cuts and pay raises.

The results show that wage cuts have a severe impact on workers' productivity. On average output decreased by more than 20% when workers experienced a wage cut. Moreover, this negative effect was visible across the entire performance distribution and remained remarkably persistent over time, suggesting that negative reciprocal behavior plays an important role in the field. In contrast, we found no significant evidence for positive reciprocity in response to an equivalent pay rise. Average productivity was almost identical in the baseline and pay raise treatments, highlighting an asymmetric reaction of work morale to positive and negative deviations from the reference wage.

We show in a subsequent control experiment that the lack of positive reactions is unlikely due to a ceiling effect—that is, workers were not constrained by their physical limit in the baseline and pay raise treatments. For this purpose we hired more subjects for doing the same task, but incentivized their performance. Workers received a piece rate which increased stepwise up to €0.40 per book with the achievement of specific output targets. Average output was almost 25% higher than in the baseline treatment and the entire performance distribution was shifted towards higher outputs. This suggests that there was enough room for productivity to increase by a substantial amount and that our results are unlikely to be confounded by a ceiling effect.

Strikingly, we observe that in contrast to the quantity of output the wage cut did not hurt quality. In light of these results one could speculate that the wage cut is reciprocated by low work energy, which translates into low output quantity, rather than surreptitious sabotage of output quality.[1]

Our results underline that workers are not purely self-interested, because otherwise we should not observe any performance differences between treatments. In what follows we apply two prominent social-preference models of inequality aversion

1. We thank an anonymous referee for this suggestion.

and reciprocity, respectively, and discuss under which assumptions those models are reconcilable with our results.[2] We argue that action-based reciprocity models are better suited to explain our data than purely outcome-oriented models of inequity aversion. The reason is that due to other sources of income the firm's payoff is typically larger than that of the worker, no matter what effort level he or she chooses. Thus, the firm is always "ahead" of the worker, and considerations based on relative positions cannot explain the observed differences in workers' behavior. By contrast, in action-based reciprocity models (such as, e.g., Cox, Friedman, and Gjerstad 2007) individuals directly react to kind or unkind actions independently of their relative positions. Therefore, these models explain our results in a more straightforward manner.

Our field experiment makes several contributions to the existing literature. First, a substantial amount of laboratory evidence demonstrates a positive relationship between wages and effort (e.g. see Fehr, Kirchsteiger, and Riedl 1993; Abeler et al. 2010; Charness 2004). However, the extent to which these results can be generalized to naturally occurring markets is not clear (see DellaVigna 2009; Falk and Heckmann 2009; Levitt and List 2007). Laboratory experiments are generally characterized by a high level of experimenter obtrusiveness, which could create demand effects. Moreover, lab experiments generally do not involve the exertion of actual effort but simply consist of monetary transfers. For these reasons, Gneezy and List (2006) conducted the first field experiment testing for positive reciprocity in the workplace. Mimicking the one-shot interactions in the lab they created short-term employment opportunities for data-entry and door-to-door fundraising. In addition to the fact that their workers were assigned a real effort task, subjects did not know that they were part of an experiment. In contrast to the evidence from the lab, Gneezy and List found that an increase in hourly wages had only a transient effect, which ultimately did not pay off for the employer.[3] Our field experiment builds on the basic design of Gneezy and List, but allows for a novel view on the impact of wage cuts.[4] In addition, the influence of wage cuts and pay raises on work morale can be studied within the same framework—highlighting significant asymmetries in the field.

Second, our findings are related to a series of interview studies on work morale (see Bewley 2005 for a review). For example, Bewley asked compensation executives for the reasons why firms are reluctant to cut wages or avoid hiring underbidders during economic downturns. The general insight from these interview studies is that the desire to maintain good work morale seems to be a key rationale employers provide for their

2. According to the taxonomy proposed by Card, DellaVigna, and Malmendier (2011) our study falls into the category of *Competing Models* field experiments.

3. Other field experiments typically found only weak or moderate evidence for positive reciprocity (Hennig-Schmidt, Rockenbach, and Sadrieh, 2010; Cohn, Fehr, and Goette, 2009; Bellemare and Shearer, 2009; Al-Ubaydli et al., 2011); an exception are those studies analyzing nonmonetary gifts (Maréchal and Thöni, 2010; Kube, Maréchal, and Puppe, 2012; Falk, 2007).

4. In an older experiment reported by Pritchard, Dunnette, and Jorgenson (1972), subjects were only made to *believe* that they were accidentally over- or underpaid; their actual wages remained unchanged. The results show no significant treatment effects, but their experimental manipulation is arguably much weaker. More recently, Cohn et al. (2011) conducted a field experiment showing that social comparison amplifies the impact of wage cuts.

policies. This valuable first indication on the role of work morale in labor markets is complemented by the causal evidence from our field experiments.

Third, identifying the causal impact of wage changes on work morale poses serious difficulties in the field. Changes in compensation generally reflect firms' choices and are therefore potentially endogenous due to unobservable confounds (see Shearer 2003). Moreover, employment contracts are frequently embedded in ongoing relationships between workers and employers. This implies that also pecuniary reasons might exist for workers' reaction to wage changes. In particular, (i) workers might provide less effort after a wage cut because they play a trigger strategy and punish the firm for cutting their wages (see Howitt 2002); or (ii) lower wages might dampen the disciplining effect of getting fired because they reduce future rents (Shapiro and Stiglitz, 1984; MacLeod and Malcomson, 1989). On the other hand, other effects could potentially mitigate the negative impact on work morale. Specifically, self-selection of workers might lead to a replacement of quitting workers by new ones who are willing to work at the lower wages. The experimental approach in our study controls for these issues and makes it possible to separate work morale from reputational and other confounding motives. In the experiment, wage changes were exogenous and reputation effects were minimized by design. We took great care in making clear that we offer a one time job without any possibility of reemployment. Consequently, our results nicely complement the few field studies that looked at workers' reactions to wage cuts using non-experimental data (see Greenberg 1990; Mas 2006; Lee and Rupp 2007).

The remainder of this paper is organized as follows. We describe the experimental design in Section 2. Sections 3 and 4 contain the results and robustness checks. Section 5 concludes the paper with a theoretical discussion of the results.

2. Experimental Design

In August 2006, the library of an economic chair at a German university had to be catalogued. We took this opportunity to run a field experiment and recruited workers with posters. The announcement said that it was a one-time job opportunity for one day (six hours), and that pay was *projected* to be €15 per hour.[5] The projected wage of €15 served as an exogenously set reference wage for the workers. About 200 persons applied during the two-month announcement phase. A research assistant randomly picked 30 persons from the list of applicants. They were invited via email and asked to confirm the starting date, reminding them that the job was projected to pay €15 per hour. Upon arrival, the subjects were seated in separate rooms in front of a computer terminal (with internet browser) and a table with a random selection of books. Their task was to enter each book's author(s), title, publisher, year of publication, and ISBN number into an electronic data base. The computer application (see Figure 1) in which

5. The announcement said "The hourly wage is projected to be €15" (the exact German wording was "Ihr Stundenlohn beträgt voraussichtlich €15"), in order to leave room for later wage changes without cheating.

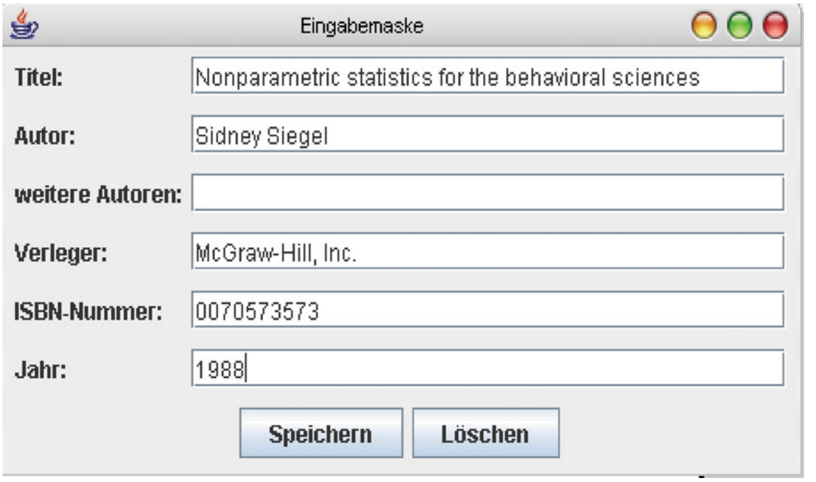

FIGURE 1. Screenshot: Computer application.

they entered the details of the books recorded the exact time of each log, allowing us to reconstruct the number of books each person entered over time without having to monitor work performance explicitly. This data entry task is well suited for our experiment and is commonly used in field experiments[6] because it allows for a precise measurement of output and quality. Moreover, the task is relatively simple and can be done in isolation, allowing for more control than usually available in other field settings. Participants were allowed to take a break whenever necessary. A research assistant explained them the task by strictly following a fixed protocol. Right before workers started their task, they were told their actual hourly wage—which depended on the treatment assignment.

We conducted three different treatments. The hourly wage paid was €15 in the "Baseline" treatment, €20 in "PayRaise" and €10 in "PayCut".[7] Because the experiment was set up as a one-shot situation, our manipulation represents a cut with respect to an exogenous wage expectation—and not with respect to the past wage which serves as a reference point in ongoing employment relations. We thus capture what is arguably a key aspect of wage cuts, namely the induced disappointment and the break of a trust relation between workers and the firm (see Bewley 2005). In order to keep communication constant across treatments we opted for a neutral framing of wage changes and gave workers no reason why they were paid more or less than the projected €15.[8] In our first wave of experiments, we had ten workers each in the

6. See Gneezy and List 2006; Kube, Maréchal, and Puppe 2012; Kosfeld and Neckermann 2011; Hennig-Schmidt, Rockenbach, and Sadrieh 2010 for recent examples.

7. €10 still exceed the hourly wages usually paid to a student helper at German universities, which is about €8. We paid slightly higher wages in order to avoid selection problems arising from workers quitting due to higher outside options. None of the invited workers refused to work for €10.

8. None of the workers actually asked for an explanation. The exact wording was, "We pay you an hourly wage of €20 (€10). Your hourly wage is thus €20 (€10) instead of €15".

benchmark and in the wage cut treatments, and nine workers in the pay raise treatment, because one worker did not show up for work.

We invited three workers per day—one in each treatment. The assignment of workers to the treatment groups was randomized. In order to avoid any treatment contaminations through social interaction, workers showed up sequentially at different times and were separated from each other, in different rooms. We did not explicitly tell them that we had employed other workers. Furthermore, all workers interacted with the same research assistant, circumventing any confounding experimenter effects.[9] After six hours of work, all workers completed a brief questionnaire. In order to observe their behavior in a natural environment, workers were not told that they were taking part in an experiment.

In October 2008, we increased our sample size and ran a second wave of identical treatments. None of the workers from the second wave had participated in the first wave. For our main treatments, we have data from 68 workers in total: 25 in Baseline, 21 in PayCut and 22 in PayRaise.

In April 2011, we conducted an additional control experiment "PieceRate" with 18 new subjects. This provides a benchmark for assessing workers' physical limits on the data-entry task.[10] Their task was the same: cataloguing books for six hours. Instead of a flat wage of €15 per hour, they faced a strong performance incentive. Every worker i received a base salary of €10 and a piece rate which increased with the total number of books Y_i entered. The total payment Π_i for a subject was given by the following formula:

$$\Pi_i(Y_i) = €10 + \begin{cases} €0.15 \cdot Y_i, & \text{if } Y_i < 240, \\ €0.30 \cdot Y_i & \text{if } 240 \leq Y_i < 260, \\ €0.40 \cdot Y_i & \text{if } 260 \leq Y_i. \end{cases} \tag{1}$$

3. Results

Panel (a) in Figure 2 illustrates average worker productivity (measured by the number of book entries) per 90 minute time interval, or quarter, for each of the three different treatments. Table 1 contains the average treatment effects—that is, the difference in average number of books logged—and the p-values from the corresponding nonparametric Wilcoxon rank-sum tests for the null hypothesis of equal output between treatments.

The results show a substantial difference in productivity between the Baseline and PayCut treatment. This effect is highly significant from a statistical and economical point of view (see columns (3) and (4) in Table 1). On average, output was 21% (or 47

9. The research assistant neither knew the purpose of the study nor the reason for the differing wages.

10. We refrained from re-running our main treatments as controls because the results from this control experiment are only used to demonstrate that workers can potentially go beyond the performance observed in the above main treatments. We thus implicitly assume that the distribution of workers' skill levels remained stable over time.

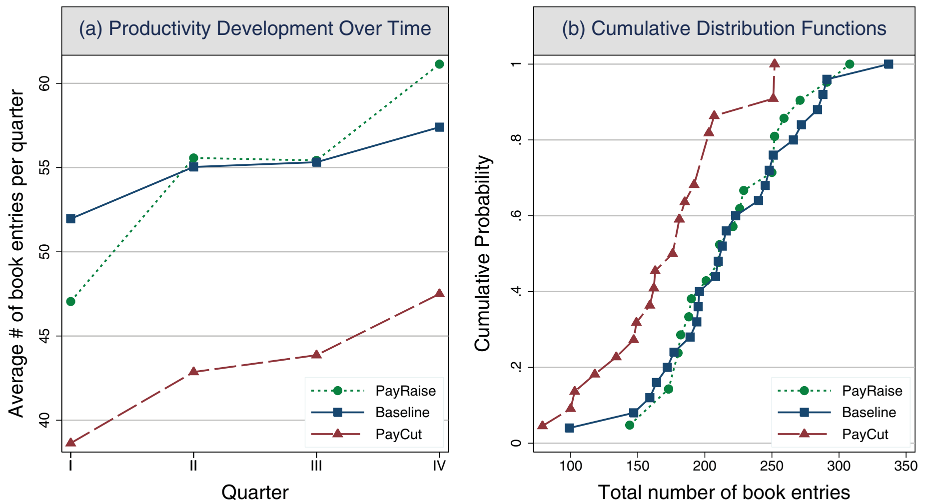

FIGURE 2. Productivity as a function of wages. Panel (a) depicts the average number of books entered per quarter (90 minutes) for the three treatments PayRaise, PayCut, and Baseline. The corresponding cumulative distribution functions for total work performance are illustrated in panel (b).

TABLE 1. Average treatment effects by time intervals: no. of book entries.

Time interval	(1) PayRaise-Baseline	(2) $p > \|z\|$	(3) PayCut-Baseline	(4) $p > \|z\|$
Quarter I	−4.9 (−9.5%)	0.247	−13.3 (−25.6%)	0.001
Quarter II	0.5 (1.0%)	0.757	−12.2 (−22.1%)	0.012
Quarter III	0.1 (0.2%)	0.991	−11.5 (−20.7%)	0.013
Quarter IV	3.7 (6.5%)	0.508	−9.9 (−17.2%)	0.026
All quarters	−0.8 (−0.3%)	0.991	−46.7 (−21.3%)	0.005
Observations		$N = 46$		$N = 47$

Notes: Columns (1) and (3) report average treatment effects (percentages in parentheses) for the treatments PayRaise and PayCut in comparison with Baseline by 90 minutes time intervals, or quarters. The outcome variable is the number of book entries. Columns (2) and (4) report the corresponding p-values from a nonparametric (two-sided) Wilcoxon rank-sum test for the null hypothesis of equal output between treatments.

books) lower in treatment PayCut than in Baseline. Moreover, as can be inferred from Figure 2, the productivity gap is stable over time. It remains large and significant for all four quarters.

On the other hand, the average treatment effect for the pay raise is slightly negative (although insignificant: $p = 0.247$) during the first quarter. Interestingly, the effect tends to become positive over the course of time, but does not reach statistical significance in any quarter (see column (2) of Table 1). Overall, we find no evidence for positive reciprocal behavior. Average output is almost identical in the Baseline and PayRaise treatments, with 219.4 and 218.6 books, respectively.

The cumulative distribution functions in panel (b) of Figure 2 show that our results are not driven by one or two individual workers; instead they reflect a broad

TABLE 2. Regression analysis.

	(1)	(2)	(3)
	Total entries		Correct entries
PayRaise	1.537	−3.243	−1.594
	(3.261)	(3.186)	(3.058)
PayCut	−13.967***	−15.577***	−10.172**
	(3.767)	(3.871)	(3.940)
Quarter II	5.132***	3.080*	1.800
	(0.850)	(1.569)	(1.718)
Quarter III	5.515***	3.360*	2.320
	(0.998)	(1.743)	(2.000)
Quarter IV	9.221***	5.440**	5.680**
	(1.328)	(2.660)	(2.268)
PayRaise ∗ Quarter II		5.444***	7.105***
		(1.979)	(2.144)
PayRaise ∗ Quarter III		5.021**	6.347**
		(2.365)	(2.525)
PayRaise ∗ Quarter IV		8.655**	6.225**
		(3.339)	(2.867)
PayCut ∗ Quarter II		1.147	1.473
		(2.048)	(2.249)
PayCut ∗ Quarter III		1.867	2.453
		(2.421)	(2.585)
PayCut ∗ Quarter IV		3.424	1.684
		(3.086)	(2.669)
Constant	99.508***	101.505***	103.073***
	(28.714)	(29.072)	(27.821)
Controls:			
Socioeconomic?	YES	YES	YES
Room FE?	YES	YES	YES
Starting time?	YES	YES	YES
Observations		272	
No. of workers		68	

Notes: This table reports OLS coefficient estimates (standard errors adjusted for clustering at the individual level are reported in parentheses). The dependent variables are the number of book entries per quarter, respectively the number of correct book entries in column (3). The treatment dummies PayCut and PayRaise are interacted with the Quarter dummies II to IV. The dummy for treatment Baseline is omitted from the regression model and serves as the reference category. Definitions and summary statistics for the control variables are reported in the additional Online Appendix.
*Significant at 10%; **significant at 5%; ***significant at 1%.

behavioral phenomenon. While the distribution functions for PayRaise and Baseline are closely intertwined, the distribution function for PayCut is clearly shifted towards lower performance. For example, while the fraction of workers who entered 200 or fewer books is only around 40% in the Baseline treatment, it amounts to 80% in the PayCut treatment.

The panel regression results in Table 2 are in line with the preceding nonparametric analysis. Our main regression model is specified as follows:

$$Y_{it} = \alpha + \beta_1 PR_i + \beta_2 PC_i + \beta_3 PR_i \cdot Q_{it} + \beta_4 PC_i \cdot Q_{it} + \gamma Q_{it} + \delta X_i + \epsilon_{it}, \tag{2}$$

where Y_{it} represents the number of books entered by worker i in quarter t. Q_{it} is a vector consisting of dummy variables indicating the corresponding quarter and PC_i and PR_i, respectively, indicate whether a worker was in the PayCut or PayRaise treatment. The dummy for the Baseline is omitted from the model and serves as the reference category. We explore how treatment effects evolve over time, and interact both treatment indicators with the quarter dummy variables. Furthermore, room fixed effects, starting time, and socioeconomic background (age, gender, and subject of studies) are included in our set of control variables X_i. We estimated our model using ordinary least squares (OLS). Standard errors are corrected for clustering on the individual level, accounting for individual dependency of the error term ϵ_{it} over time.

The coefficient estimate for PayCut is highly significant and has the expected sign in the benchmark model without interaction effects (column (1)), whereas the coefficient for PayRaise is close to zero and does not reach statistical significance. Moreover, all of the PayCut and Quarter interaction terms in column (2) are relatively small and insignificant, highlighting temporal stability of the treatment effects during the observed time span. On the other hand, the estimated PayRaise and Quarter interaction terms indicate that the effect of the pay raise is significantly higher after quarter one. Positive reciprocal reactions thus tend to strengthen with the elapse of time. A further interesting result—which is also visible in Figure 2—is that the number of entries per quarter increases substantially over time, which we interpret as a learning effect.

4. Robustness Checks

We performed several robustness checks. First, in addition to the quantity of output, we also investigated the impact of our treatments on output quality. We measured output quality by the ratio of faultless entries to the total number of books entered (for a similar approach see Hennig-Schmidt, Rockenbach, and Sadrieh 2010).[11] The average quality ratio amounts to 84.4% in treatment Baseline. Interestingly, we find that quality is with 90.4% significantly higher in the PayCut treatment (Wilcoxon rank-sum test: $p = 0.030$), suggesting that the lower typing speed resulted in fewer mistakes. Quality measured 87.7% in PayRaise, and was also slightly higher than in the Baseline treatment. Nevertheless the difference does not reach statistical significance ($p = 0.800$). In order to account for both, the quantity and the quality dimension of effort, we used the number of correct entries as a composite measure of work performance. The results are displayed in column (3) of Table 2 and show that the coefficient estimate for PayCut remains large and statistically significant. We also experimented with an alternative specification using the total number of entries as the

11. Two research assistants searched for incorrectly entered ISBN numbers and spelling mistakes in the book titles (using an automatic spell check program).

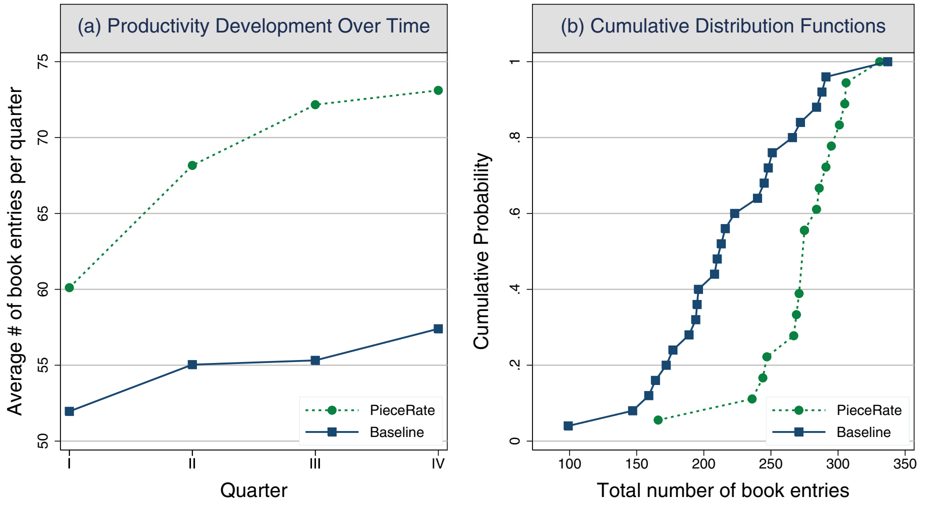

FIGURE 3. Incentive effects. Panel (a) depicts the average number of books entered per quarter (90 minutes) for the treatments PieceRate and Baseline. The corresponding cumulative distribution functions for total work performance are illustrated in panel (b).

dependent variable and the number of typing errors as an additional control variable. The results are robust using this specification.

Second, if the employees already worked at their physical limit in treatment Baseline we would observe that higher wages are ineffective even if workers wanted to provide more effort. We tested whether workers reached their physical limit in Baseline with the additional control experiment "PieceRate". Instead of a fixed hourly wage, workers were paid a piece rate which increased stepwise up to €0.40 per book with the achievement of specific output targets (see equation (1) for the exact formula).

The results in Figure 3 show that workers produced substantially more output when they were paid by performance rather than a fixed wage. Average output was with 274 books 24.7% higher in treatment PieceRate than in Baseline. This difference is statistically significant according to a Wilcoxon rank-sum test ($p < 0.001$). Moreover, panel (b) in Figure 3 highlights that the entire performance distribution is shifted towards a higher output. Only one subject was below the average or median output of treatment Baseline. In contrast to these output differences, quality of output was largely unaffected by the piece rate. The average quality in the PieceRate treatment was 83.2% and did not significantly differ from Baseline (Wilcoxon rank-sum test: $p = 0.483$). The regression results in column (1) and (2) of Table C.1 in Appendix C show that the results remain unchanged if we control for other potential influences. In column (3) we use the number of correct entries as the dependent variable. The results remain qualitatively unchanged if we take into account both, the quantity and quality dimension of output. Altogether the results from treatment

PieceRate suggest that workers neither reached their physical limit in Baseline nor in PayRaise.

Finally, as alternatives to using OLS with clustered standard errors, we conducted all regressions using (i) bootstrapped standard errors or (ii) a random effects model with generalized least squares. The main results remained unchanged with respect to these alternative specifications. We also experimented by adding the hourly wage earned at their most recent job as a proxy for human capital to our empirical models. We found that controlling for previous wages does not affect the results.[12]

5. Discussion

The results show that wage cuts have a severe impact on productivity. Moreover, this negative effect remained large and significant over the course of the entire working period. While these results are supportive for the idea that wage cuts damage work morale, we found no significant evidence that an equivalent pay raise fostered productivity. Assuming that workers are purely self-interested we should not observe any performance differences between treatments, because wages were not tied to performance and employment relations were one-shot. As such, our findings can be considered as a reduced-form step forward in understanding the role of social preferences in labor markets. The question remains which model of social preferences is able to explain our results. Two prominent approaches in the literature are models of inequality aversion and models of reciprocity. Both approaches provide potential explanations for the behavior observed in laboratory gift-exchange games. In the following we briefly discuss these models in light of our field experiment.

Models of inequality aversion assume that people face a trade-off between maximizing their own income and equalizing income distributions. In a standard laboratory gift-exchange game with symmetric outside options the worker's payoff exceeds the employer's payoff if the employer paid a high fixed wage upfront. By exerting costly effort workers increase the employer's payoff and therefore reduce payoff inequalities. Hence, if workers are sufficiently inequality averse, higher wage payments lead to higher effort levels (see Fehr and Schmidt 1999, p. 848ff). While these models (see also Bolton and Ockenfels 2000) provide a plausible interpretation of the laboratory data, it is unclear to what extent they explain our field data.[13] In contrast to the lab, a worker's payoff is usually well below the firm's payoff in the field. The firm (i.e. here the university) typically has substantial income from other projects. Consequently, the firm does not "fall behind" by paying a high wage.[14] Costly effort reduces the worker's payoff and widens the payoff gap between the firm and the worker. An inequality-averse worker would thus choose minimal effort irrespective of the actual wages. This hypothesis is clearly rejected by our data. Our results could

12. The results from these robustness checks are available upon request.

13. We thank the editor for pointing this out.

14. See Appendix A for a formalization of this argument using Fehr and Schmidt (1999).

still be made consistent with a model of inequality aversion if we additionally assume that workers neglect the firms' other sources of income and focus solely on current bilateral rents—that is, workers are "narrow bracketing" (see Read, Loewenstein, and Rabin 1999, p. 186). Evidently, in that case appropriate additional assumptions about the magnitude of (the worker's perception of) these rents are necessary (see Appendix A for details).

As pointed out by Card, DellaVigna, and Malmendier (2011), it might be easier to reconcile gift-exchange in the field with an action-based reciprocity model.[15] Consider, for example, the following simple model based on Cox, Friedman, and Gjerstad (2007). Let the utility of a worker be given by $u(x_w, x_f) = x_w + \theta x_f$.[16] The firm's payoff $x_f = v(e) - w$ equals the value $v(e)$ of the output that is produced by the worker minus the unconditional wage payment w. The worker's payoff $x_w = w - c(e)$ equals the wage payment minus effort costs $c(e)$. If the firm acts unkind the "emotional state function" θ has a negative sign. The worker will then choose minimal effort, because costly effort reduces the worker's payoff and increases the firm's payoff, both of which causes a drop in the worker's utility. By contrast, if the firm acts kind θ is strictly positive. In this case, the (θ-weighted) value of the produced output $v(e)$ enters positively into the workers utility via the firm's payoff. Optimal effort e^* then needs to satisfy $\theta v'(e^*) = c'(e^*)$. For a large class of reasonable cost- and production-functions, there exists a unique solution that exceeds the minimal effort level (see Appendix B for details).

The model's predictions are in line with our field data under the following assumptions. First, an arbitrary wage cut is considered an unkind action ($\theta \leq 0$) and a pay raise a kind action ($\theta > 0$). Second, θ is assumed to be positive in treatment Baseline, too. Otherwise workers would choose minimal effort in Baseline. This assumption seems plausible for our setting, because we chose a quite generous wage of €15 in Baseline, leaving room for wage cuts. Our workers earned on average a bit more than €10.50 in previous employment relations (see Table 1 in the Online Appendix). Third, one has to assume that either the increase in θ between Baseline and PayRaise is small, or that marginal costs of effort are increasing in effort. Both of these conditions seem plausible; the latter corresponds to standard convexity assumptions on cost functions. Under either assumption, effort would react less to a pay rise than a pay cut, in line with our data. In contrast to the inequality-aversion models discussed previously, no additional assumption about narrow bracketing is needed. In this action-based reciprocity model workers will also engage in gift-exchange with firms that are "ahead" in payoffs, as long as the wage payment is not perceived as unkind—that is, as long as $\theta > 0$ holds. The key feature of the underlying explanation is the asymmetry of the worker's optimal effort response function: for all $\theta > 0$ there is an interior solution strictly above the minimal effort level, while for all $\theta \leq 0$ the

15. See also Levine (1998) or Rabin (1993) for type-based and intention-based models of reciprocity, respectively.

16. This corresponds to the utility function in Cox, Friedman, and Gjerstad (2007) with $\alpha = 1$.

optimal response is constant at the minimal level. This can easily lead to an asymmetric reaction of optimal effort to wage cuts versus wage increases.

While our study provides suggestive evidence that action-based reciprocity models can explain gift-exchange in the field, a rigorous test of the different models needs additional information. Specifically, it would require precise knowledge of workers' effort costs and workers' perceptions of output value. The latter might simply be induced or elicited. Actual effort costs could be estimated from data on workers' output under different piece rates; the piece rates being randomly generated to minimize the impact of intentions. Additional treatments with wage changes of different magnitudes would further help to pin down the curvature of the wage–output function.

Future studies might also explore the determinants of the emotional state θ in more detail. For example, workers' evaluation of wage cuts and pay raises—and consequently their behavioral reaction to them—might depend on the explanations for the wage changes (e.g., Greenberg 1990; Chen and Horton 2009). Wage cuts might be perceived differently during a recession if workers understand that the wage cuts are necessary for the company to stay afloat (or to avoid layoffs).

In ongoing relations, workers might face a trade-off between giving in to their negative reactions and increasing the risk of being fired because of poor work performance. Furthermore, self-selection might mitigate the negative effects of wage cuts. Workers who are dissatisfied with a wage cut could quit the job and might be replaced by other workers who are willing to work for those wages. These issues promise to be interesting topics for future research.

Appendix A: Application of Fehr and Schmidt (1999)

In this appendix, we explore the possibility to explain our data with the concept of inequality aversion. In the model of Fehr and Schmidt (1999), for instance, the utility function of the worker is given by

$$u_w(x_w, x_f) = x_w - \alpha \cdot \max\{0; x_f - x_w\} - \beta \cdot \max\{0; x_w - x_f\},$$

where x_w is the worker's payoff and x_f the firm's payoff. In most specifications of the model one has $0 < \beta < \alpha$ (so that u_w has a kink at the point where $x_f = x_w$) and $\beta < 1$. The worker's payoff x_w is composed of the unconditional wage payment w minus cost of effort $c(e)$ plus income from other resources m. The effort level e is assumed to be non-negative and above some minimal acceptable level e_{min}. The firm's payoff x_f is given by the value $v(e)$ of the output produced by the worker minus the wage payment w plus income from other resources M. The difference $x_f - x_w$ is given by $M - m - 2w + v(e) + c(e)$. Under usual circumstances, the firm's income M will be much larger than the single worker's income m also after correcting for the term $2w$ so that the payoff difference $x_f - x_w$ is generally positive. In that case, the worker's utility function reduces to

$$u_w(x_w, x_f) = x_w - \alpha \cdot [x_f - x_w].$$

Assuming that $v(\cdot)$ and $c(\cdot)$ are strictly increasing in effort, this implies that the worker will always choose the minimal admissible effort level, independently of the specific wage payment. But this is rejected by our data.

Some laboratory experiments suggest that subjects are "narrow bracketing"—that is, they neglect income and rents from other sources (e.g. see Read, Loewenstein, and Rabin 1999). While this assumption seems less plausible in the field it can yield a possible explanation of our data under additional assumptions. If the worker neglects income from other sources (M and m), the optimal effort choice will depend on the sign of $x_f - x_w$. First, if $[x_f(e_{min}) - x_w(e_{min})] = v(e_{min}) + c(e_{min}) - 2w > 0$, then the worker will always choose the minimal admissible effort level e_{min} since the difference $x_f - x_w$ is increasing in e (as already argued above). If on the other hand, the difference $x_f - x_w$ is negative for all admissible levels of effort and all *relevant* wage payments, then the worker's utility reduces to

$$u_w(x_w, x_f) = x_w - \beta \cdot \left[x_w - x_f\right].$$

This implies that the optimal effort level e^* is either given by the first-order condition

$$\frac{\beta}{1-\beta} \cdot v'(e^*) = c'(e^*),$$

or by the corner solution e_{min}, depending on the specification of the functions $c(\cdot)$ and $v(\cdot)$. In any case, optimal effort would be independent of the wage payment which is inconsistent with our data. Thus, to explain our data one would have to assume that, at least for some of the actual wages of our treatments, the "equalizing" effort $\tilde{e}$—that is, the effort $\tilde{e}$ for which $x_f = x_w$, is admissible. Indeed, as is easily verified, $\tilde{e}$ is the only other candidate for an interior solution. Note that of the three candidate solutions to the optimal effort response problem in fact only $\tilde{e}$ depends on the wage payment (because satisfaction of the condition $x_f = x_w$ depends on the wage payment). Whether $\tilde{e}$ is the actual solution depends on further specifications of the functions $c(\cdot)$ and $v(\cdot)$.

Appendix B: Application of Cox et al. (2007)

The following simple model is based on Cox, Friedman, and Gjerstad (2007). The utility function of the worker is given by $u_w(x_w, x_f) = x_w + \theta x_f$, where x_w is the worker's payoff, x_f the firms payoff, and θ is an "emotional state function". The functional form corresponds to the Cox et al. model with $\alpha = 1$. The worker's payoff x_w is composed of the unconditional wage payment w plus rents m from other sources minus cost of effort $c(e)$; the cost function $c(\cdot)$ is assumed to be increasing and convex. The firm's payoff x_f is given by the value $v(e)$ of the output produced by the worker plus rents M from other projects minus the wage payment w; the production function $v(\cdot)$ is assumed to be increasing and concave. The exact size and the sign of the "emotional state function" depends on reciprocity. In line with Cox et al., θ is assumed to be positive, except when the firm acts unkindly in which case θ becomes negative. For our purposes it is natural to assume that θ is positive also in the "neutral" benchmark case when the worker is paid the announced wage. This is justified by the fact that the the wage of €15 in our benchmark treatment is already quite generous.[17]

The worker's optimal effort choice is given by the solution to

$$\max_{e \geq e_{min}} \quad (m + w - c(e) + \theta \cdot v(e) - \theta \cdot w + \theta \cdot M),$$

where e_{min} is the minimal effort level, as above. If $\theta < 0$, the solution to this problem is $e^* = e_{min}$ since the objective function is decreasing. On the other hand, if $\theta > 0$ and $\theta \cdot v'(e_{min}) > c'(e_{min})$ then the solution e^* is strictly above e_{min}. If in addition, $c'(\cdot)$ is unbounded as e gets large, then the solution e^* is given by the first-order condition[18]

$$\theta \cdot v'(e^*) = c'(e^*).$$

Thus, if $\theta > 0$ then optimal effort e^* will (in general) be above the minimal level. Moreover, e^* varies continuously with the state function θ. In particular, depending on the specific shape of $v(\cdot)$ and $c(\cdot)$, optimal effort may be quite insensitive to variations of θ as long as the latter remains positive. Thus, optimal effort can also be insensitive to an increase in the wage payment w. This occurs either if θ itself is relatively insensitive to an increase of w, or if c is sufficiently convex so that a rise of θ induces only a small increase in effort because effort is too costly. By contrast, if θ changes its sign and becomes negative, optimal effort suddenly drops to the minimal level. Given appropriate specifications of $v(\cdot)$, $c(\cdot)$, and θ, the present model can thus naturally explain the observed strong negative reaction to wage cuts as opposed to the mild (or absent) positive reaction to wage increases. Note in particular, that the optimal effort response naturally depends on the magnitude of the wage payment, but not on income from unrelated sources. Unlike inequality aversion models, an action-based reciprocity model does not need any assumption about "narrow-bracketing".

The assumption that θ becomes negative after a wage cut, induces an asymmetry in workers behavior because optimal effort is always at the minimal acceptable level in case of a wage cut. While our data are consistent with this asymmetry, other assumptions on θ would make similar predictions. For instance, one could assume that θ is a positive function of wages and is kinked at a "neutral" reference wage. But even without such a kink, an asymmetry could occur due to increasing marginal costs of effort. In general, one would then obtain interior solutions both for the wage increase and the wage cut.

17. Our workers earned on average a bit more than €10.50 in previous employment relations.

18. The second-order condition is satisfied due to convexity of $c(\cdot)$ and concavity of $v(\cdot)$.

Appendix C: Additional Regressions

TABLE C.1. Regression analysis: PieceRate versus Baseline.

	(1)	(2)	(3)
	Total entries		Correct entries
PieceRate	16.170***	10.862*	11.730**
	(5.556)	(5.822)	(5.810)
Quarter II	5.163***	3.080*	1.800
	(1.142)	(1.600)	(1.752)
Quarter III	7.000***	3.360*	2.320
	(1.569)	(1.778)	(2.039)
Quarter IV	8.605***	5.440*	5.680**
	(2.342)	(2.713)	(2.313)
PieceRate ∗ Quarter II		4.976**	6.033**
		(2.074)	(2.348)
PieceRate ∗ Quarter III		8.696***	8.569***
		(2.930)	(2.861)
PieceRate ∗ Quarter IV		7.560	6.598
		(4.792)	(4.193)
Constant	169.078**	171.300**	219.352***
	(69.545)	(70.159)	(62.674)
Controls:			
Socioeconomic?	YES	YES	YES
Room FE?	YES	YES	YES
Starting time?	YES	YES	YES
Observations		172	
No. of workers		43	

Notes: This table reports OLS coefficient estimates (standard errors adjusted for clustering at the individual level are reported in parentheses). The dependent variables are the number of book entries per quarter, respectively the number of correct book entries in column (3). The treatment dummy PieceRate is interacted with the quarter dummies II to IV. The dummy for treatment Baseline is omitted from the regression model and serves as the reference category. Definitions and summary statistics for the control variables are reported in the additional Online Appendix.
*Significant at 10%; **significant at 5%; ***significant at 1%.

References

Abeler, Johannes, Steffen Altmann, Sebastian Kube, and Matthias Wibral (2010). "Gift Exchange and Workers' Fairness Concerns: When Equality is Unfair." *Journal of the European Economic Association*, 8, 747–772.

Akerlof, George A. (1982). "Labor Contracts as Partial Gift Exchange." *Quarterly Journal of Economics*, 97, 543–569.

Al-Ubaydli, Omar, Steffen Andersen, Uri Gneezy, and John List (2011). "Carrots that Look Like Sticks: Toward an Understanding of Multitasking Incentive Schemes." Working paper.

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

Bewley, Truman F. (1999). *Why Wages Don't Fall During a Recession*. Harvard University Press.

Bewley, Truman F. (2005). "Fairness, Reciprocity, and Wage Rigidity." In *Moral Sentiments and Material Interests: The Foundations of Cooperation in Economic Life*, edited by H. Gintis, S. Bowles, R. T. Boyd, and E. Fehr. MIT Press.

Bolton, Gary E. and Axel Ockenfels (2000). "A Theory of Equity, Reciprocity and Competition." *American Economic Review*, 90, 166–193.

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

Charness, Gary (2004). "Attribution and Reciprocity in an Experimental Labor Market." *Journal of Labor Economics*, 22, 665–688.

Chen, Daniel and John G. Horton (2009). "The Wages of Pay Cuts: Evidence from a Field Experiment." Working paper.

Cohn, Alain, Ernst Fehr, and Lorenz Goette (2009). "Fairness and Effort: Evidence from a Field Experiment." Working paper.

Cohn, Alain, Ernst Fehr, Benedikt Herrmann, and Frederik Schneider (2011). "Social Comparison in the Workplace: Evidence from a Field Experiment." IZA Working Paper No. 5550.

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.

Falk, Armin (2007). "Gift-Exchange in the Field." *Econometrica*, 75, 1501–1511.

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

Fehr, Ernst, Georg Kirchsteiger, and Arno Riedl (1993)."Does Fairness Prevent Market Clearing? An Experimental Investigation." *Quarterly Journal of Economics*, 108, 437–460.

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

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.

Greenberg, Jerald (1990). "Employee Theft as a Reaction to Underpayment Inequity: The Hidden Cost of Pay Cuts." *Journal of Applied Psychology*, 75, 561–568.

Hennig-Schmidt, Heike, Bettina Rockenbach, and Abdolkarim Sadrieh (2010). "In Search of Workers' Real Effort Reciprocity—A Field and a Laboratory Experiment." *Journal of the European Economic Association*, 8, 817–837.

Howitt, Peter (2002). "Looking Inside the Labor Market: A Review Article." *Journal of Economic Literature*, 40, 125–138.

Kosfeld, Michael, and Susanne Neckermann (2011). "Getting More Work for Nothing? Symbolic Awards and Worker Performance." *American Economic Journal: Microeconomics*, 3, 86–99.

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

Lee, Darin and Nicholas G. Rupp (2007). "Retracting a Gift: How Does Employee Effort Respond to Wage Reductions?" *Journal of Labor Economics*, 25, 725–761.

Levine, David K. (1998). "Modeling Altruism and Spitefulness in Experiments." *Review of Economic Dynamics*, 1, 593–622.

Levitt, Steven D. and John A. List (2007). "What do Laboratory Experiments Measuring Social Preferences tell us about the Real World." *Journal of Economic Perspectives*, 21, 153–174.

MacLeod, Bentley W. and James M. Malcomson (1989). "Implicit Contracts, Incentive Compatibility, and Involuntary Unemployment." *Econometrica*, 57, 447–480.

Maréchal, Michel A. and Christian Thöni (2010). "Hidden Persuaders: Do Small Gifts Lubricate Business?" Working paper.

Mas, Alexandre (2006). "Pay, Reference Points, and Police Performance." *Quarterly Journal of Economics*, 121, 783–821.

Pritchard, Robert D., Marvin D. Dunnette, and Dale O. Jorgenson (1972). "Effects of Perceptions of Equity and Inequity on Worker Performance and Satisfaction." *Journal of Applied Psychology*, 56, 75–94.

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

Read, Daniel, George Loewenstein, and Matthew Rabin (1999). "Choice Bracketing." *Journal of Risk and Uncertainty*, 19, 171–197.

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

Shearer, Bruce S. (2003). "Compensation Policy and Worker Performance: Identifying Incentive Effects from Field Experiments." *Journal of the European Economic Association*, 1, 503–511.

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

Supporting Information

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

Online Appendix.