Institute for Empirical Research in Economics
University of Zurich

Working Paper Series
ISSN 1424-0459

Working Paper No. 471

Do Wage Cuts Damage Work Morale? Evidence from a Natural Field Experiment

Sebastian Kube, Michel André Maréchal and Clemens Puppe

January 2010

Do Wage Cuts Damage Work Morale? Evidence from a Natural Field Experiment*

Sebastian Kube, Michel André Maréchal and Clemens Puppe

January, 2010

Abstract

Contractual incompleteness characterizes many employment relations. High work morale is therefore fundamental for sustaining voluntary cooperation within the firm. We conducted a natural field experiment testing to what extent wages affect work morale. The results provide clear-cut evidence showing that wage cuts have a detrimental impact on work morale. An equivalent wage increase, however, does not result in any productivity gains. Theses results highlight a strongly asymmetric response of work morale to wage variations.

JEL classification: C93, J30.
Keywords: morale, reciprocity, gift exchange, field experiment.

* We thank 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 Nicholas Bardsley, Monika Bütler, Alain Cohn, Stefano DellaVigna, 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 and UC Los Angeles, and UC Berkeley for very helpful comments. Kube: University of Bonn, Department of Economics, Adenauerallee 24-42, 53113 Bonn, Germany (email: kube(at)uni-bonn.de). Maréchal: University of Zurich, Institute for Empirical Research in Economics, Bluemlisalpstrasse 10, CH-8006 Zurich, Switzerland (email: marechal(at)iew.uzh.ch). Puppe: Karlsruhe Institute of Technology, Department of Economics and Business Engineering, 76128 Karlsruhe, Germany (email: clemens.puppe(at)kit.edu).

"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)

1 Introduction

Why are firms reluctant to cut wages during economic downturns? A prominent explanation for this puzzle is based on the psychology of work morale:[1] Work morale reflects the degree to which workers voluntarily cooperate and contribute to the employer's goals in the absence of reputation or pecuniary incentives. According to this view, work morale is sensitive to the relationship between the workers' actual wage and some 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 (see Slichter (1920, 1929), Solow (1979) or Akerlof (1982)), 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 naturally occurring employment relations. We conducted a controlled

[1] See Azariadis (1975), Lindbeck and Snower (1986) and Katz (1986) for alternative theories that explain downward wage rigidity.

field experiment and tested the extent to which workers reciprocate different hourly wages.[2] We hired job applicants to catalogue books for a *limited* time (i.e. excluding any possibility of reemployment) and announced a *projected* wage of €15 per hour. We actually paid this amount in our benchmark treatment, and it serves as an exogenous reference point for wage expectations. In our main treatment, we inform subjects immediately before they begin working that we will only pay them €10 per hour. In a second treatment, we do the opposite and communicate a pay raise from €15 to €20 per hour in order to explore asymmetries between the impact of wage cuts and pay raises on work morale.[3]

The results show that wage cuts have a severe impact on the effort workers provide. Productivity drops on average by more than 20 percent if workers experience a wage cut. This negative effect is remarkably persistent over time in both size and significance. Our results suggest that negative reciprocal behavior plays an important role in naturally occurring employment relations. In contrast, we find no evidence for positive reciprocal reactions to an equivalent pay raise. Together our results highlight a strongly asymmetric reaction of work morale to positive and negative deviations from the reference wage.

Our field experiment makes several contributions to the existing liter-

[2] By reciprocity we refer to the behavioral phenomenon of people responding towards (un)kind treatment likewise, even in the absence of reputational concerns. Economic theories formalize reciprocal behavior by incorporating the distribution of outcomes, the perceived kindness of intentions, or simply emotional states as arguments into individual utility function (see Charness and Rabin (2002), Falk and Fischbacher (2006), Rabin (1993), Dufwenberg and Kirchsteiger (2004), or Cox et al. (2007)).

[3] The second treatment is similar to Gneezy and List (2006), where workers had to catalogue books and wages were increased from 12 to 20 US Dollars.

ature. First, an impressive amount of laboratory evidence suggests that reciprocal behavior has important implications in *experimental* labor markets (e.g. see Fehr et al. (1993, 1997, 2007), Abeler et al. (forthcoming), Charness (2004) or Hannan et al. (2002)). However, laboratory experiments are generally characterized by a high level of experimenter scrutiny, which creates potential demand effects (see Zizzo (forthcoming)). Moreover, lab experiments generally do not involve the exertion of actual effort but simply consist of monetary transfers. The extent to which these results can be generalized to naturally occurring markets is thus not clear (see DellaVigna (2009), Falk and Heckmann (2009) or Levitt and List (2007)). We were able to observe subjects in a more natural – yet still controlled – working environment, because they performed a typical student helper's task and did not know that they were part of an experiment. Apart from the issue of generalizability, many of the existing experimental paradigms do not disentangle positive from negative reciprocal behavior. In the standard laboratory gift-exchange game (Fehr et al. (1993)), for example, a positive correlation between wages and effort could be driven by positive reciprocity towards high wages as well as retaliation for low wages.[4]

Second, to the best of our knowledge, this is the first study providing controlled evidence for *negative* reciprocal behavior in a natural labor market situation.[5] The few existing field experiments focus on the economic

[4]See Offerman (2002), Engelmann and Ortmann (2009) for alternative lab experimental paradigms that allow a distinction between positive and negative reciprocity.

[5]An earlier experiment reported by Pritchard et al. (1972) comes close to our design. They found no significant treatment effects with respect to performance. Their experimental manipulation is arguably much weaker, however, because their subjects were only made to *believe* that they were accidentally over- or underpaid; their actual wages remained unchanged.

consequences of *positive* reciprocity, and their conclusions still remain ambiguous. Falk (2007), for example, shows that charitable donations increase substantially with the size of gifts included in solicitation letters, rendering gift giving profitable. Gneezy and List (2006), on the other hand, find that an increase in hourly wages has only a transient effect, which ultimately did not pay off for the employer. Other field experiments typically found only weak or moderate evidence for positive reciprocity (Hennig-Schmidt et al. (forthcoming), Cohn et al. (2009), Bellemare and Shearer (2009) or Al-Ubaydli et al. (2006)), with the exception of those studies analyzing non-monetary gifts (Maréchal and Thöni (2010), Kube et al. (2009)). In addition to testing for reciprocal reactions towards wage cuts, our design allows for a novel direct comparison between the influence of wage cuts and pay raises within the same framework, and highlights significant asymmetries in the field.

Third, Bewley (1999)) conducted interviews with compensation executives, exploring the reasons why firms are reluctant to cut wages or avoid hiring underbidders during economic downturns (see also Blinder and Choi (1990), Levine (1993), or Campbell and Kamlani (1997)). 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 policies. This line of research provides a valuable first indication on the role of work morale in labor markets. However, this methodology also has drawbacks. Social desirability effects are a well-known phenomenon in survey research; they question the extent to which we can take answers from interviews at face value (see Bertrand and Mullainathan (2001) or Krosnick (1999)). More important, while interviews provide some information with regard as to why

firms are reluctant to cut wages, ultimately they only reflect beliefs and do not measure the extent to which wages affect work morale.

Fourth, identifying the causal impact of wage cuts 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)). Consequently, there are only a few field studies and they rely on non-experimental data (see Mas (2006), Krueger and Mas (2004), and Lee and Rupp (2007)). These studies are embedded in an ongoing relationship between workers and employers, making it impossible to fully separate work morale from reputational motives.[6] There are at least two alternative pecuniary reasons why workers provide less effort after a wage cut in repeated interactions. First, workers could play a trigger strategy and punish the firm for cutting their wages by exerting lower effort (see Howitt (2002)). Second, lower wages reduce future rents and dampen the disciplining effect of getting fired (see Shapiro and Stiglitz (1984) or MacLeod and Malcomson (1989)). We took great care in making clear that we offer a one time job without any possibility of reemployment and can therefore rule out reputational motives. Furthermore, while effort often manifests itself in a multitude of dimensions, our simple data entry task allows us to measure work performance very accurately. In contrast, Lee and Rupp (2007), for example, have to rely on flight delays as the single proxy for the effort airline pilots provide. Flight delays, however, can serve only as a very crude proxy for effort and are strongly influenced by other forces beyond the pilots'

[6]Greenberg (1990) also uses quasi-experimental firm data to analyze the effect of a wage cut on employee theft. In addition to the fact that his experiment is not a one-shot situation, his analysis is unfortunately only based on three independent observations.

control.

The remainder of this paper is organized as follows: In the next section, we describe the experimental design. In Sections 3 and 4, the experimental results are presented and discussed. And finally Section 5 concludes the paper.

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 from all over the campus 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.[7] 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 picked 30 persons out of 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 front of a computer terminal and a table with a random selection of books. Their task was to enter the book's author(s), title, publisher, year of publication, and ISBN number into an electronic data base. This data entry task is well suited for our experiment, as it allows for a precise measurement of output and quality. Moreover, the task is relatively simple

[7] The announcement said "The hourly wage is projected to be €15," (the exact German wording was "Ihr Stundenlohn beträgt voraussichtlich €15"), in order to set expectations without cheating.

and can be done in isolation, allowing for more control than usually available in other field settings.[8] Participants were allowed to take a break whenever necessary. A research assistant explained the task to them, strictly following a fixed protocol. Then, subjects were told their actual hourly wage – which depended on the treatment assignment – and started working.

We conducted three different treatments. The hourly wage paid in our benchmark treatment was €15 ("Baseline"), €20 in "PayRaise" and €10 in "PayCut".[9] 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 (2002)). We opted for a relatively neutral framing of wage changes and gave subjects no reason why they were paid more or less than the projected €15.[10] In our first wave of experiments, we had 10 subjects each in the benchmark and in the wage cut treatments, and 9 subjects in the pay raise treatment, because one subject did not show up for work.

We invited three subjects per day – one in each treatment. In order to avoid any treatment contaminations through social interaction, subjects

[8] Data entry tasks are thus frequently used in field experiments (see Gneezy and List (2006), Kube et al. (2009), Kosfeld and Neckermann (2009) and Hennig-Schmidt et al. (forthcoming) for some recent examples).

[9] €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.

[10] None of the subjects 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".

showed up sequentially at different times and were separated from each other, in different rooms at an online computer terminal. Furthermore, all subjects interacted with the same research assistant, circumventing any confounding experimenter effects.[11] The computer application in which 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.[12] After 6 hours of work, all subjects completed a brief questionnaire. In order to observe their behavior in a natural environment, subjects 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. We have data from 68 workers in total: 25 in Baseline, 21 in PayCut and 22 in PayRaise.

3 Results

Randomization Check

Table 3 reports summary statistics and tests whether observable covariates are balanced across treatments using Pearson's χ^2 or Kruskal-Wallis tests. With the exception of Room A, which was used less frequently in treatment PayCut, we cannot reject the null hypothesis that observable worker characteristics and the environmental conditions are balanced across treatments.

[11] The research assistant knew neither the purpose of the study nor the reason for the differing wages.

[12] See Figure 2 in the Appendix for a screen shot.

In summary, the randomization resulted in a fairly well balanced set of workers and environmental conditions. We include room fixed effects as well as starting-time fixed effects in our regression models.

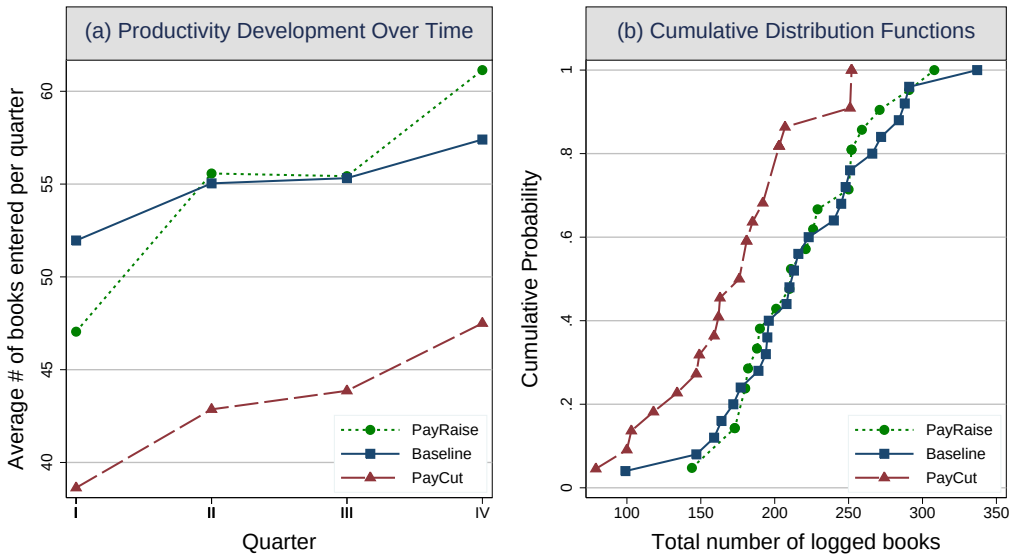
Wages and Work Morale

Panel (a) in Figure 1 illustrates average worker productivity (measured by the number of books logged) per 90 minute time interval, or quarter, for each of the three different treatments. Table 1 contains the average treatment effects – i.e. 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 treatments. This effect is highly significant from both a statistical and economical point of view (see columns three and four in Table 1). On average, output was 21 percent (or 47 books) lower in treatment PayCut than in Baseline. Moreover, as can be inferred from Figure 1, 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 two of Table 1). Overall, we find no evidence for positive reciprocal behavior. Average output is virtually identical in the Baseline and PayRaise treatments, with 219.3

Figure 1: Work Morale as a Function of Wages



Notes: Panel (a) depicts the average number of books logged 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).

and 218.6 books, respectively.

The cumulative distribution functions in Panel (b) of Figure 1 show that our results are not driven by one or two individual workers; instead they reflect a broad 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 logged 200 or fewer books is only around 40 percent in the Baseline treatment, it amounts to 80 percent in the PayCut treatment.

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

Table 1: Average Treatment Effects by Time Intervals: # Books Logged

Time interval	(1) **PayRaise-Baseline**	(2) $p > \|z\|$	(3) **PayCut-Baseline**	(4) $p > \|z\|$
Quarter I	-4.9	0.247	-13.3	0.001
Quarter II	0.5	0.757	-12.2	0.012
Quarter III	0.1	0.991	-11.5	0.013
Quarter IV	3.7	0.508	-9.9	0.026
All quarters	-0.7	0.991	-46.6	0.005
Observations		N=46		N=47

Notes: Columns 1 and 3 report average treatment effects for the treatments PayRaise and PayCut in comparison with Baseline by 90 minutes time intervals, or quarters. The outcome variable is the number of books logged. 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.

follows:

$$Y_{it} = \alpha + \beta_1 PR_i + \beta_2 PC_i + \beta_3 PR_i * Q_{it} + \beta_4 PC_i * Q_{it} + \gamma Q_{it} + \theta_i + \omega_i + \epsilon_{it}, \quad (1)$$

where Y_{it} represents the number of books logged 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 Baseline treatment 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, vectors containing room (ω_i) and starting time (θ_i) fixed effects are included in our set of control variables. We estimated our model using Ordinary Least Squares (OLS). Standard errors are corrected for clustering, 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 (column 1), whereas the coefficient for PayRaise does not reach statistical significance. Moreover, all of the PayCut and Quarter interaction terms 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 hence tended to strengthen with the elapse of time. A further interesting result - which is also clearly visible in Figure 1 - is that the number of books logged per quarter increased substantially over time, which we interpret as a learning effect.

Robustness Checks

We performed several robustness checks. First, we control for socioeconomic characteristics in column (2) of Table 2 by expanding the set of control variables with the workers' age, gender, and subject of studies. The results remain unchanged. Second we include the hourly wage earned at the most recent job prior to the experiment as a proxy for human capital.[13] As demonstrated in columns (3) and (4) of Table 2, controlling for previous wages does not affect the key results.

Third, as an alternative to using OLS with clustered standard errors, we estimated a random effects model with Generalized Least Squares. The main

[13] The information about previous wages is missing for 21 workers. These subjects are therefore excluded from the sample when we control for previous wages (columns 3 to 5 of Table 2).

results remained unchanged with respect to this alternative specification.[14]

Fourth, in addition to the effect on the quantity of output, we also investigated the impact of our treatments on output quality. We measured output quality by the ratio of faultless logs to the total number of books entered (see Hennig-Schmidt et al. (forthcoming) for a similar approach).[15] The average quality ratio amounts to 84.4 percent in treatment Baseline. Interestingly, we find that quality is with 90.4 percent 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 percent in PayRaise, and was also slightly higher than in the Baseline treatment. Nevertheless the difference does not reach statistical significance ($p = 0.800$). Overall we find evidence for a quantity-error trade-off: The number of errors is positively and significantly correlated with the number of books logged (Spearman's $\rho = 0.531$; $p < 0.0001$). We therefore use the number of correct logs as a composite measure of work performance, taking into account of both the quantity and the quality dimension of effort. The results are displayed in column (5) 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 logs as the dependent variable and the number of typing errors as an additional control variable. The results are robust to this alternative specification.

[14] The results are available upon request.

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

Table 2: Panel Regressions

	(1)	(2)	(3)	(4)	(5)
	total logs				**correct logs**
PayRaise	-3.833	-4.583	-0.217	-2.121	2.371
	(3.378)	(3.519)	(3.451)	(3.757)	(3.568)
PayCut	-14.097***	-15.826***	-16.270***	-16.551***	-11.485***
	(3.463)	(3.358)	(4.303)	(4.389)	(3.974)
Quarter II	3.080*	3.080*	4.353**	4.353**	3.765**
	(1.562)	(1.578)	(1.860)	(1.888)	(1.727)
Quarter III	3.360*	3.360*	5.235***	5.235***	4.471**
	(1.736)	(1.753)	(1.722)	(1.748)	(1.832)
Quarter IV	5.440**	5.440**	5.765	5.765	6.176**
	(2.650)	(2.676)	(3.677)	(3.733)	(2.950)
PayRaise∗Quarter II	5.444***	5.444***	4.118*	4.118*	4.412*
	(1.971)	(1.991)	(2.364)	(2.401)	(2.224)
PayRaise∗Quarter III	5.021**	5.021**	3.824	3.824	4.706*
	(2.356)	(2.380)	(2.507)	(2.546)	(2.621)
PayRaise∗Quarter IV	8.655**	8.655**	8.706*	8.706*	5.176
	(3.326)	(3.359)	(4.438)	(4.506)	(3.617)
PayCut∗Quarter II	1.147	1.147	-0.891	-0.891	-1.380
	(2.039)	(2.060)	(2.310)	(2.345)	(2.258)
PayCut∗Quarter III	1.867	1.867	0.534	0.534	-0.009
	(2.411)	(2.436)	(2.978)	(3.024)	(2.715)
PayCut∗Quarter IV	3.424	3.424	3.389	3.389	1.131
	(3.074)	(3.105)	(4.160)	(4.224)	(3.397)
Constant	58.529***	80.643***	65.912***	77.181***	60.446***
	(3.384)	(9.829)	(6.980)	(11.211)	(10.091)
Controls:					
Socioeconomic?	NO	YES	NO	YES	YES
Previous wage?	NO	NO	YES	YES	YES
Room FE?	YES	YES	YES	YES	YES
Starting time FE?	YES	YES	YES	YES	YES
Obs.	272	272	188	188	188

Notes: This table reports OLS coefficient estimates (standard errors adjusted for clustering are reported in parentheses). The dependent variable is the number of books logged per quarter, respectively the number of *correctly* logged books in column (5). The treatment dummies PayCut and PayRaise are interacted with the quarter dummies II to IV. Definitions and summary statistics for the additional control variables are reported in Tables 4 and 3. Due to item non-response the sample size is lower in columns (3) to (5) where we control for previously earned hourly wages. Significance levels are denoted as follows: * p<0.1, ** p<0.05, *** p<0.01.

4 Discussion

The results show a striking asymmetry between the effect of wage cuts and pay raises. We discuss two potential explanations for this finding. The first explanation concerns the parameterization of the experiment. Because we intended to leave room for wage cuts, our baseline wage is with €15 already quite generous. Our subjects earned on average a bit more than €10.5 (see Table 3) in previous employment relations. If there is only a positive correlation between effort and wages when wages are below what workers perceive as fair wage, our design would favor finding treatment effects for wage cuts.[16] A recent field experiment conducted by Cohn et al. (2009) provides evidence supportive for this view. They find that only workers who felt dissatisfied with their baseline wage reciprocated after a pay raise. Kube et al. (2009) conducted additional field experiments using the same paradigm as in the present study, but a lower baseline wage. Their subjects were recruited for €12, which is much closer to the €10.5 our subjects were accustomed to earning in the past. The implemented 20 percent wage increase, however, did not result in any significant productivity gains. Interestingly, the results show at the same time that an equivalent non-monetary gift resulted in substantially higher output, suggesting that there was still room for positive reciprocal behavior.

Second, substantial experimental evidence demonstrates that losses loom larger than gains of equal size (e.g. Kahneman et al. (1991, 1986), or Gächter et al. (2007)). More generally, the psychological literature suggests that

[16] Akerlof and Yellen (1990) for example assume such a discontinuity in their "Fair Effort-Wage Hypothesis".

negative or bad events have greater influence than do good ones in a great variety of contexts (see Rozin and Royzman (2001)). Baumeister et al. (2001) conclude in their extensive literature survey that the predominance of "bad over good" may be considered as a "general principle or law of psychological phenomena (p. 323)". This negativity bias manifests itself, for example, in higher physiological arousal or attention to negative events. In the context of our experiment, this would imply that wage cuts get more attention and are evaluated more negatively than a corresponding wage gift is appreciated

5 Concluding Remarks

Maintaining high work morale is of paramount importance for firms whenever workers' effort is not fully contractible. A longstanding explanation for downward wage rigidity presumes that wage cuts damage work morale (see Slichter (1920, 1929), Solow (1979) or Akerlof (1982)). However, corresponding evidence from the field is scarce because compensation schemes usually do not vary exogenously but reflect firms' decisions (see Shearer (2003)). Apart from these identification problems, ongoing relations between workers and firms make it hard to disentangle work morale from alternative pecuniary or reputational motives in the field.

This study fills this gap and provides clear-cut evidence on the impact of wage cuts on work morale using a labor market field experiment. In addition, the paper provides a novel direct comparison between the impact of wage cuts and corresponding wage increases within the same framework. In summary, our results show that wage cuts have a severe impact on produc-

tivity. Moreover, this negative effect remains large and significant over the course of the entire working period. While these results are supportive for the notion that wage cuts damage work morale, we find no evidence that pay raises foster work morale. An equivalent pay raise resulted in virtually no change in productivity levels. Together, our results provide new evidence stressing the importance of work morale and highlight a strongly asymmetric performance response to wage variations.

References

Abeler, J., S. Altmann, S. Kube, and M. Wibral (forthcoming): "Gift Exchange and Workers' Fairness Concerns: When Equality is Unfair", *Journal of the European Economic Association*.

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

Akerlof, G. A. and J. L. Yellen (1990): "The Fair Wage-Effort Hypothesis and Unemployment", *Quarterly Journal of Economics*, Vol. 105, pp. 255–283.

Al-Ubaydli, O., S. Andersen, U. Gneezy, and J. List (2006): "For Love or Money? Testing Non-Pecuniary and Pecuniary Incentive Schemes in a Field Experiment", Working Paper.

Azariadis, C. (1975): "Implicit Contracts and Unemployment Equilibria", *Journal of Political Economy*, Vol. 83, pp. 1183–1202.

Baumeister, R.F., E. Bratslavsky, C. Finkenauer, and K. D. Vohs (2001): "Bad is stronger than good", *Review of General Psychology*, Vol. 5, pp. 323–370.

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

Bertrand, M. and S. Mullainathan (2001): "Do People Mean What They Say? Implications for Subjective Survey Data", *American Economic Review*, Vol. 91, pp. 67–72.

Bewley, T. F. (2002): "Fairness, Reciprocity, and Wage Rigidity", *Cowles Foundation Discussion Paper No. 1383*.

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

Blinder, A. S. and D. H. Choi (1990): "A Shred of Evidence on Theories of Wage Stickiness", *Quarterly Journal of Economics*, Vol. 105, pp. 1003–1015.

Campbell, C. M. and K. S. Kamlani (1997): "The Reasons for Wage Rigidity: Evidence From a Survey of Firms", *Quarterly Journal of Economics*, Vol. 112, pp. 759–789.

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

Charness, G. and M. Rabin (2002): "Understanding Social Preferences with Simple Tests", *Quarterly Journal of Economics*, Vol. 117, pp. 817–869.

Cohn, A., E. Fehr, and L. Goette (2009): "Fairness and Effort: Evidence from a Field Experiment", Working Paper.

Cox, J. C., D. Friedman, and S. Gjerstad (2007): "A Tractable Model of Reciprocity and Fairness", *Games and Economic Behavior*, Vol. 59, pp. 17–45.

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

Dufwenberg, M. and G. Kirchsteiger (2004): "A Theory of Sequential Reciprocity", *Games and Economic Behavior*, Vol. 47, pp. 268–298.

Engelmann, D. and A. Ortmann (2009): "The Robustness of Laboratory Gift Exchange: A Reconsideration", Working Paper.

Falk, A. (2007): "Gift-Exchange in the Field", *Econometrica*, Vol. 75, pp. 1501–1511.

Falk, A. and U. Fischbacher (2006): "A Theory of Reciprocity", *Games and Economic Behavior*, Vol. 54, pp. 293–315.

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

Fehr, E., S. Gächter, and G. Kirchsteiger (1997): "Reciprocity as a Contract Enforcement Device", *Econometrica*, Vol. 65, pp. 833–860.

Fehr, E., G. Kirchsteiger, and A. Riedl (1993): "Does Fairness Prevent Market Clearing? An Experimental Investigation", *Quarterly Journal of Economics*, Vol. 108, pp. 437–460.

Fehr, E., A. Klein, and K. Schmidt (2007): "Fairness and Contract Design", *Econometrica*, Vol. 75, pp. 121–154.

Gächter, S., A. Hermann, and E. J. Johnson (2007): "Individual-Level Loss Aversion in Riskless and Risky Choices", CeDEx Discussion Paper No. 2007-02.

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

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

Hannan, R. L., J. Kagel, and D. V. Moser (2002): "Partial Gift Exchange in an Experimental Labor Market: Impact of Subject Population Differences, Productivity Differences and Effort Requests on Behavior", *Journal of Labor Economics*, Vol. 20, pp. 923–951.

Hennig-Schmidt, H., B. Rockenbach, and A. Sadrieh (forthcoming): "In Search of Workers' Real Effort Reciprocity - A Field and a Laboratory Experiment", *Journal of the European Economic Association.*

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

Kahneman, D., J. L. Knetsch, and R. Thaler (1986): "Fairness as a Constraint on Profit Seeking: Entitlements in the Market", *American Economic Review*, Vol. 76, pp. 728–41.

Kahneman, D., J. L. Knetsch, and R. H. Thaler (1991): "Anomalies: The Endowment Effect, Loss Aversion, and Status Quo Bias", *Journal of Economic Perspectives*, Vol. 5, pp. 193–206.

Katz, L. F. (1986): "Efficiency Wage Theories: A Partial Evaluation", *NBER Macroeconomics Annual Review of Psychology*, Vol. 1, pp. 235–276.

Kosfeld, M. and S. Neckermann (2009): "Working for Nothing? The Effect of Non-material Awards on Employee Performance", Working Paper.

Krosnick, J. A. (1999): "Survey Research", *Annual Review of Psychology*, Vol. 50, pp. 537–67.

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

Kube, S., M. A. Maréchal, and C. Puppe (2009): "The Currency of Reciprocity - Gift-Exchange in the Workplace", Working Paper.

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

Levine, D. I. (1993): "Fairness, Markets, and Ability to Pay: Evidence from Compensation Executives", *American Economic Review*, Vol. 83, pp. 1241–1259.

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

Lindbeck, A. and D. J. Snower (1986): "Wage Setting, Unemployment, and Insider-Outsider Relations", *American Economic Review, Papers and Proceedings*, Vol. 76, pp. 235–239.

MacLeod, B. W. and J. M. Malcomson (1989): "Implicit Contracts, Incentive Compatibility, and Involuntary Unemployment", *Econometrica*, Vol. 57, pp. 447–480.

Maréchal, M. A. and C. Thöni (2010): "Hidden Persuaders: Do Small Gifts Lubricate Business?", Working Paper.

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

Offerman, T. (2002): "Hurting Hurts More Than Helping Helps", *European Economic Review*, Vol. 46, pp. 1423–1437.

Pritchard, R. D., M. D. Dunnette, and D. O. Jorgenson (1972): "Effects of Perceptions of Equity and Inequity on Worker Performance and Satisfaction", *Journal of Applied Psychology*, Vol. 56, pp. 75–94.

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

Rozin, P. and E. B. Royzman (2001): "Negativity Bias, Negativity Dominance, and Contagion", *Personality and Social Psychology Review*, Vol. 5, pp. 296–320.

Shapiro, C. and J. E. Stiglitz (1984): "Equilibrium Unemployment as a Worker Discipline Device", *American Economic Review*, Vol. 74, pp. 433–444.

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

Slichter, S. (1920): "Industrial Morale", *Quarterly Journal of Economics*, Vol. 35, pp. 36–60.

Slichter, S. H. (1929): "The Current Labor Policies of American Industries", *Quarterly Journal of Economics*, Vol. 43, pp. 393–435.

Solow, R. M. (1979): "Another Possible Source of Wage Stickiness", *Journal of Macroeconomics*, Vol. 1, pp. 79–82.

Zizzo, D. J. (forthcoming): "Experimenter Demand Effects in Economic Experiments", *Experimental Economics*.

6 Appendix

Figure 2: Screenshot: Computer Application

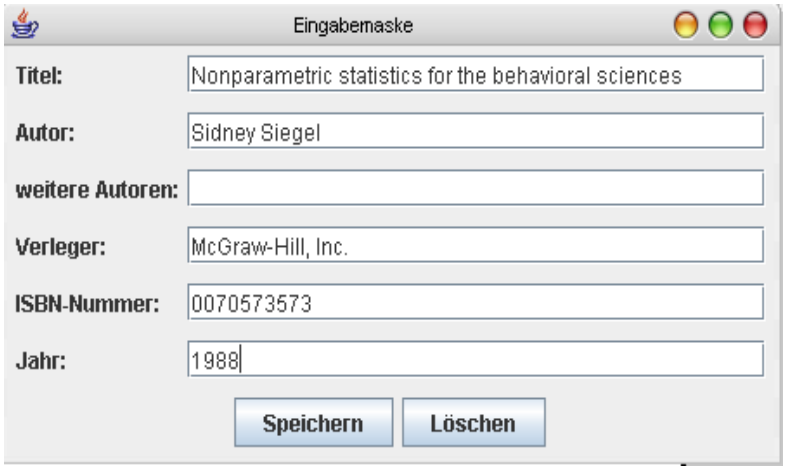


Table 3: Summary Statistics and Randomization Check

Variable	PayCut (N=21) Mean	PayCut (N=21) Std. Dev.	Baseline (N=25) Mean	Baseline (N=25) Std. Dev.	PayRaise (N=22) Mean	PayRaise (N=22) Std. Dev.	Full Sample (N=68) Mean	Full Sample (N=68) Std. Dev.	Kruskal-Wallis/ χ^2 p-value
Age	23.571	2.767	24.360	3.340	22.955	3.552	23.662	3.254	0.118
Male	0.571	0.507	0.480	0.510	0.500	0.512	0.515	0.503	0.815
Math and Physics	0.143	0.359	0.080	0.277	0.045	0.213	0.088	0.286	0.522
Engineering and IT	0.286	0.463	0.160	0.374	0.364	0.492	0.265	0.444	0.278
Arts and Social Science	0.238	0.436	0.360	0.490	0.227	0.429	0.279	0.452	0.527
Economics	0.286	0.463	0.360	0.490	0.364	0.492	0.338	0.477	0.829
Previous wage (Euro/h)	10.530	2.677	11.188	5.331	9.742	3.071	10.550	3.895	0.664
Room A	0.048	0.218	0.320	0.476	0.273	0.456	0.221	0.418	0.066
Room B	0.238	0.436	0.200	0.408	0.091	0.294	0.176	0.384	0.416
Room C	0.381	0.498	0.240	0.436	0.364	0.492	0.324	0.471	0.529
Room D	0.190	0.402	0.080	0.277	0.136	0.351	0.132	0.341	0.544
Room E	0.143	0.359	0.160	0.374	0.136	0.351	0.147	0.357	0.972
Start 9:00am	0.190	0.402	0.440	0.507	0.455	0.510	0.368	0.486	0.128
Start 9:30am	0.333	0.483	0.160	0.374	0.091	0.294	0.191	0.396	0.115
Start 9:45am	0.143	0.359	0.200	0.408	0.136	0.351	0.162	0.371	0.807
Start 10am	0.190	0.402	0.080	0.277	0.136	0.351	0.132	0.341	0.544
Start 10:30am	0.143	0.359	0.120	0.332	0.182	0.395	0.147	0.357	0.835

Notes: The last column of this table contains p-values from Pearson's χ^2 tests for binary and Kruskal-Wallis tests for non-binary controls. Due to item non-response concerning previous wage levels the corresponding sample sizes are lower than for the other variables: PayCut (N=17), Baseline (N=17) and PayRaise (N=13).

Table 4: Control Variables: Wording and Coding (Translated from German to English)

Variable	**Definition**	**Question wording [Possible answers in brackets]**
Socioeconomic		
Age	years	Age? [free form]
Male	1=yes; 0=no	Gender? [free form]
Math and Physics	1=yes ; 0=no	Subject of studies? [free form]
Engineering and Computer Science	1=yes ; 0=no	
Arts and Social Science	1=yes ; 0=no	
Economics	1=yes ; 0=no	
Previous wage		
Previous wage	Euro per hour	What was your hourly wage on your last job? [free form]

Table 5: Summary Data: Total Number of Books Logged and Quality Ratio

PayCut			Baseline			PayRaise		
Participant	**# books**	**Quality ratio**	**Participant**	**# books**	**Quality ratio**	**Participant**	**# books**	**Quality ratio**
1	149	.966	23	288	.885	48	221	.945
2	192	.869	24	210	.847	49	226	.964
3	203	.945	25	147	.727	50	211	.791
4	252	.936	26	164	.914	51	180	.844
5	118	.915	27	223	.618	52	188	.835
6	203	.896	28	159	.924	53	144	.930
7	103	.961	29	189	.888	54	229	.829
8	79	.911	30	272	.886	55	252	.829
9	100	.920	31	196	.913	56	210	.871
10	162	.913	32	291	.666	57	259	.911
11	251	.940	33	195	.979	58	173	.797
12	163	.754	34	248	.899	59	271	.863
13	252	.928	35	99	.909	60	180	.961
14	176	.829	36	251	.928	61	252	.837
15	147	.972	37	266	.751	62	308	.883
16	181	.900	38	284	.838	63	201	.890
17	159	.911	39	194	.932	64	190	.910
18	181	.883	40	213	.896	65	291	.920
19	203	.960	41	208	.875	66	250	.816
20	207	.859	42	172	.517	67	182	.939
21	134	.888	43	245	.734	68	173	.849
22	185	.832	44	216	.884			
			45	337	.875			
			46	177	.875			
			47	240	.925			
Average	*172.7*	*.904*		*219.36*	*.844*		*218.6*	*.877*