Putting Reciprocity to Work -- Positive versus Negative Responses in the Field

Sebastian Kube, Michel André Maréchal and Clemens Puppe

November 2006 Discussion Paper no. 2006-27

Department of Economics University of St. Gallen

Editor: Prof. Jörg Baumberger
University of St. Gallen
Department of Economics
Bodanstr. 1
CH-9000 St. Gallen
Phone +41 71 224 22 41
Fax +41 71 224 28 85
Email joerg.baumberger@unisg.ch

Publisher: Department of Economics
University of St. Gallen
Bodanstrasse 8
CH-9000 St. Gallen
Phone +41 71 224 23 25
Fax +41 71 224 22 98

Electronic Publication: http://www.vwa.unisg.ch

Putting Reciprocity to Work -- Positive versus Negative Responses in the Field[1]

Sebastian Kube. Michel André Maréchal and Clemens Puppe

Authors' addresses:

Sebastian Kube and Clemens Puppe
University of Karlsruhe
Department of Economics
76128 Karlsruhe
Tel. +49 721 608 30 77
Fax +49 721 608 30 82
Email lastname[at]wior.uni-karlsruhe.de
Website www.wior.uni-karlsruhe.de/LS_Puppe

Michel André Maréchal
FEW-HSG
Varnbüelstrasse 14
9000 St. Gallen
Tel. +41 71 224 23 12
Fax +41 71 224 23 02
Email MichelAndre.Marechal@unisg.ch
Website www.few.unisg.ch

[1] We thank Carsten Dietz, Philipp Stroehle, and especially Michael Weingärtner for providing excellent research assistance and Johannes Kaiser for programming the computer application. *Further acknowledgements to be added.*

Abstract

We study the role of reciprocity in a labor market field experiment. In a recent paper, Gneezy and List (2006) investigate the impact of gift exchange in this context and find that it has only a transient effect on long run outcomes. Extending their work to examine both positive and negative reciprocity, we find consonant evidence in the positive reciprocity condition: the gift does not work well in the long run (if at all). Yet, in the negative reciprocity treatment we observe much stronger effects: a wage reduction has a significant and lasting negative impact on efforts. Together, these results highlight the asymmetry of positive and negative reciprocity that exists in the field, and provide an indication of the relative importance of each in the long run.

Keywords: reciprocity, gift exchange, field experiment

JEL Classification: C93, J30.

1 Introduction

Reciprocal motivations are probably the most prominent explanation for the hypothesis of efficiency wages and the associated phenomenon of involuntary unemployment in equilibrium (see Akerlof (1982), Akerlof and Yellen (1988, 1990)). An impressive amount of lab experimental evidence underpins the role of reciprocity[1] in labor market relationships. In Fehr et al. (1993)'s classic gift exchange experiment, a principal can pay a wage to an agent, who in turn exerts costly effort that generates a payoff for the principal. They observe a positive wage-effort relationship: paying above-minimum wages (i.e. sharing part of the profits) is perceived as kind and rewarded by a high effort of the worker. Hence, reciprocity seems to be able to mitigate enforcement problems of incomplete contracts (see Fehr et al. (1997)). Moreover, trust contracts relying on gift exchange can outperform explicit contracts as Fehr et al. (Forthcoming) show in their laboratory experiments.

However more recently, critical voices have emerged questioning to what extent this lab evidence can be generalized to actual labor markets.[2] Gneezy and List (2006) (henceforth GL) investigate the robustness of positive reciprocity, specifically with respect to one crucial aspect that has previously not been considered in lab experiments: the duration of the task. In their natural field experiment the recruited students had to perform a data entry or a door-to-door fund raising task. During the first few hours GL observe a significant difference in performance between their *Gift* treatment, where the subjects received an unexpected increase in hourly wages, and the control treatment. These treatment effects did, however, not persist and quickly vanished. GL attribute this decrease to an adaptation after the immediate 'hot' phase of decision making.

[1] By reciprocity we refer to the psychological phenomenon of people's (un)kind reactions towards (un)kind actions in the absence of reputational concerns. For formal models of reciprocity, cf. Rabin (1993), Dufwenberg and Kirchsteiger (2004), Falk and Fischbacher (2006) or Cox et al. (2005).

[2] Note that this is not about external validity of laboratory results in general. The question is rather whether the usually implemented gift exchange design captures all relevant features of a natural occurring work environment, see Harrison and List (2004), List (2006), Levitt and List (2006).

The present study is motivated by a number of recent experimental studies and surveys suggesting that reciprocal reactions are of asymmetric nature. Specifically, it is frequently observed that negative reciprocity is a much stronger and more robust phenomenon than positive reciprocity.[3] We therefore hypothesize that in contrast to a wage increase, cutting wages can have a persistent detrimental impact on work morale and efforts. In order to test this hypothesis we conducted an experiment in a natural occurring work environment. The small library of an economic chair at a German University had to be catalogued, and we thus had the opportunity to implement a design which closely followed GL's library task. We hired students to catalogue the books of the library for a limited time duration of six hours at an announced hourly wage of *presumably* 15 Euros - the amount actually paid out in our benchmark treatment. In our main treatment, we triggered negative reciprocity by telling them, right before they started to work, that we would pay them only 10 Euro per hour. While this treatment complements the work of GL[4], we additionally ran a positive reciprocity treatment (by paying subjects 20 instead of 15 Euros per hour) for the purpose of replicating GL's results. Since we were especially interested in the duration of the effects, after six hours of work each subject was offered to work an additional hour for 15 Euro (irrespective of the initial treatment).

Our results suggest that wage-cuts can have severe implications on workers' effort. Moreover, during the observed period there is no significant indication that workers adapted to the lower wage over time and increased their effort relative to the benchmark treatment. On the contrary, the detrimental effect is so strong that it cannot be compensated for by setting the wages back to their original level. Our results suggest that negative reciprocity plays an

[3] In the experiments conducted by Engelmann and Ortmann (2001), Pereira et al. (2006) and Offerman (2002) negative reciprocity is stronger than positive reciprocity. Surveying firms, Bewley (1999), Campbell and Kamlani (1997) as well as Blinder and Choi (1990) report that respondents believe effort to respond more strongly to wage decreases than to wage increases. More generally, Baumeister et al. (2001) conclude in their extensive survey of the psychological literature that the predominance of "bad over good" may be considered as a "general principle or law of psychological phenomena (p. 323)".

[4] Actually, GL even proposed exploring negative reciprocity for future research in their epilogue.

important role in actual labor markets. In line with GL, the evidence for the role of positive reciprocity is less conclusive. In fact, we found even less evidence for gift exchange in the short run.[5]

The remainder of this paper is organized as follows: In the next section, we describe the experimental design. Subsequently, in Sections 3 and 4, the experimental results are described and discussed.

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, building upon GL's experimental design. We recruited undergraduate students from all over the campus via posters. The announcement read that it was a one-time job opportunity for one day (six hours) that will be paid *presumably* 15 Euro per hour.[6] About 200 students applied during the two-months' announcement phase. A research assistant randomly 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 would pay *presumably* 15 Euro per hour. Upon arriving at working day, the subjects were seated in front of a computer terminal with a table of randomly picked books beside them. Their task was to enter into an electronic data base the books' author(s), title, publisher, year and ISBN number. 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 hourly wage and started working.

We conducted three different treatments. The hourly wages paid were 15 Euro in our benchmark treatment (treatment *Neutral*), 20 Euro in the positive reciprocity treatment (treatment *Kind*) and 10 Euro in the negative reciprocity treatment (treatment *Unkind*).[7] No reason was given to the sub-

[5] In a field experiment with a different task, Hennig-Schmidt et al. (2005) also do not find evidence for any form of positive reciprocity.

[6] The announcement said "presumably," or "prospective" (the exact German wording was "voraussichtlich"), in order to raise appropriate expectations without cheating.

[7] 10 Euro still exceed the hourly wages usually paid to a student helper at the University

jects about why they were paid more or less than the announced 15 Euro. We had 10 subjects in the benchmark and negative reciprocity treatment, and 9 subjects in the positive reciprocity treatment, since one subject did not show up to work.

The experiment took place over a 10 days' period, with 3 subjects per day (one in each treatment). The subjects showed up successively at different times and were separated from each other, in different rooms at an online computer terminal, without being monitored.[8] The computer application in which they entered the details of the books, recorded the exact time of each log, allowing us to exactly reconstruct the amount of books each person entered over time and the time span between two books entered. After 6 hours elapsed, everyone had to complete a questionnaire asking (among other things) how they perceived the task and the wage. In order to observe them in a 'natural work environment', subjects were not told that they were taking part in an experiment.

After completing the questionnaire, all subjects were asked if they were willing to work immediately an additional hour for 15 Euros (irrespective of their initial treatment). In case they rejected this offer, they were paid their total wage and left. Otherwise, they worked one additional hour and received their total wage payment afterwards. The data gained from the additional hour enables us to compare the performance of a subject under different wages. Unfortunately, only three subjects in *Neutral* and two subjects in *Kind*, but seven subjects in *Unkind* agreed to work an additional hour.

---

of Karlsruhe, which is 7.53 Euro. We paid higher wages in the benchmark treatment, since we wanted to focus on the effect of the unexpected *change* rather than the absolute height of the wage, which would otherwise have been extremely low in the *Unkind* treatment.

[8] All this was done in order to minimize the possible bias of other influences, like peer effects, monitoring effects, time of start or day of work. Furthermore, all subjects interacted with the same research assistant to eliminate experimenter effects. In order not to induce an experimenter demand effect, the research assistant did know neither the purpose of the study nor the reason for the wage cut/rise.

3 Results

Figure 1 illustrates the average effort (measured by the number of books logged) per 90 respectively 15 minutes time intervals for 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 (two sided) from the corresponding nonparametric Wilcoxon rank-sum tests for the null hypothesis of equal effort between treatments. Over the entire period we observe a substantial difference in effort between subjects in the *Neutral* and *Unkind* treatments. This effect is highly significant from a statistical and economic point of view (see columns three and four in Table 1). On average, workers in the *Unkind* treatment logged 56 (or 27 percent) books less than in the benchmark treatment.[9] Moreover, as hypothesized, and in contrast to the findings of GL for positive reciprocity, there is no indication for any adaptation over the observed periods (i.e. the average treatment effect remains large and significant over time)[10].

Concerning positive reciprocity, the average treatment effect for the wage increase is marginally small during the first 90 minutes' period. Only between minutes 181 and 270 we find an effect which is significant at a 10 percent level (see column two of Table 1). In line with the results of GL, there is only weak support for the existence of positive reciprocity; in fact, we do not even find much evidence for gift exchange in the short run. Workers provided on average 10 percent more effort (or logged 21 more books) due to the wage increase. This difference is statistically not significant and less than half of the absolute magnitude of the discrepancy between the treatments *Unkind* and *Neutral*.

[9] For the chosen calibration the average costs per book logged do not differ significantly between the treatments *Unkind* (0.43 Euro) and *Neutral* (0.44 Euro). Our attention is focused on the impact of wages on effort, since any statements about efficiency would arbitrarily depend on the chosen parameters.

[10] Alternatively, one could define adaptation as the difference between the number of books entered in the first and the last 90 minutes. In this sense, the neutral treatment has a roughly 14% increase whereas the unkind treatment has a roughly 30% increase, suggesting that there might be convergence as in GL. We are grateful to John List for pointing this out. However, the null hypothesis of equal differences between the treatments *Unkind* and *Neutral* cannot be rejected using a two-sided Wilcoxon rank-sum test (p=0.45).

The regression results in Table 2 are very much in line with the preceding nonparametric analysis. In column (1) we use the same linear Random Effects model specification as GL and add an indicator for the *Unkind* treatment and the corresponding interaction terms with the *Timet* dummies. The coefficient for *Unkind* is highly significant whereas the one for *Kind* does not reach any conventional significance level. Moreover, all of the $Unkind * Timet$ interaction effects are insignificant, suggesting that there was no adaptation over the observed period. On the other hand, as indicated by the estimated $Kind * Timet$ coefficients, the effect of *Kind* is significantly higher after the first 90 minutes elapsed.[11] A further interesting result - which is also clearly visible in Figure 1 - is that the number of books logged increased substantially over time, which we interpret as a learning effect. As reported in column (2) all results discussed so far are robust if higher frequency data - consisting of 15 minutes time intervals - is[12] used.[13, 14]

In addition to the effect on the quantity of effort, we also investigated the impact of our treatments on the *quality* of effort. Following the method of Hennig-Schmidt et al. (2005) we measured the quality of output by the ratio of faultless logs to the total number of books entered.[15] We found that our treatments had no significant impact on the quality of the effort provided (see last column in Table 3).[16] The regression results in column (3) of Table 2

[11] Nevertheless, the null hypothesis that all $Kind * Timet$ interaction effects are equal to zero cannot be rejected (Wald test $\chi^2 = 5.70$ $Prob > \chi^2 = 0.1271$).

[12] "Is" or "are" depending on the background knowledge of Latin.

[13] Note that we have chosen an alternative specification for the higher frequency data in order to avoid the large number of coefficients that would have resulted if we used dummy variables. Instead of the $Timet$ dummies we constructed the variable $Time$ which takes values from 1 to 24 indicating the time period.

[14] As an alternative to the random effects Generalized Least Squares regression we also ran all regressions using OLS with clustered standard errors. The coefficient estimates for $Kind * Time3$ (p=0.126), $Kind * Time4$ (p=0.141) and - using 15 minutes intervals - $Kind * Time$ (p=0.156) become insignificant. All other results remain unchanged in terms of significance and magnitude.

[15] Our measure of quality accounts for spelling mistakes in the titles of the books. Two research assistants searched for errors in the titles by running an automatic spell checking program.

[16] Using a two-sided Wilcoxon rank-sum test, we cannot reject the null hypothesis of equal output quality between *Kind* and *Neutral* (p=0.624) as well as *Unkind* and *Neutral* (p=0.7624).

demonstrate that our main results are robust if we use only error-free entries as our measure of workers' effort.

After the course of the regular working time subjects were offered the opportunity to work one additional hour for 15 Euros.[17] This allows us to test whether the detrimental impact of the wage-cut on effort can be offset by a subsequent wage increase. During the last 60 minutes of the regular working time the workers in the *Unkind* treatment logged on average 29.6 books. Even though the hourly wage was increased by 50 percent, back to 15 Euros, the average effort level remained unchanged at 28.3 books logged (Wilcoxon signed-rank test: z=0.170 and $p > |z| = 0.865$). The ineffectiveness of the wage increase is unlikely the result of exhaustion, since subjects in the treatments *Kind* and *Neutral* logged in on average 43.6 books, which is according to a Wilcoxon rank-sum test ($z = 2.847$ $p > |z| = 0.004$) significantly more.

4 Discussion and Concluding Remarks

Summarizing, we find that wage-cuts have severe implications on workers' effort. Moreover, during the observed period there is no significant indication that workers adapted to the lower wage over time. On the contrary, the detrimental effect is so strong that it cannot be compensated for by setting the wages back to their original level. While these results are broadly supportive for the notion of negative reciprocity in labor markets, the evidence for positive reciprocity is less conclusive. In line with GL we do not find a significant positive relationship between higher wages and effort, not even in the short run.[18] The response of effort to the higher wage is lagged and

[17] Only few subjects were willing to work for an additional hour in the treatments *Kind* (N=3) and *Neutral* (N=2). Therefore, in the following we only consider the 7 subjects that agreed to work for an additional hour in the *Unkind* treatment.

[18] If one defines effort in GL's fundraising task as number of households approached rather than amount of dollar raised, which seems in some sense more appropriate, there is also no indication of positive reciprocity at all, not even in the short run (cf. Footnote 9 in GL).

about half the magnitude of the response to a wage reduction of the same size.[19] The quality of workers' output does not significantly differ between all of our treatments.

At least three potential reasons might have caused the observed differences between GL's and our results concerning the effect in the short run. First, the wage increase implemented by GL (66 percent) was - in relative terms - twice as large as ours (33 percent) and consequently might have been evaluated more positively. Second, as mentioned above the hourly wage offered in *Neutral* (15 Euros) was well above the usually paid 7.53 Euros for student helpers. Thus, subjects could have already perceived 15 Euros as a generous offer and provided an above average effort. In fact, the number of books logged during the first 90 minutes in our *Neutral* treatment (49.7 books) is nearly identical to the effort provided in GL's $Gift$ treatment (51.7 books). Finally, and related to this, the results for the $Kind$ treatment might have been affected by a ceiling effect. If an average of 50 books per 90 minutes corresponds to the physical limit for unexperienced subjects, even an increased effort in $Kind$ could have not resulted in a higher measurable output than in $Neutral$. Only due to the learning effect the workers in $Kind$ were able to gradually increase their output. Note that all potential explanations are related to the actual parametrization of the experiment. This highlights that future research, both in the field and in the lab, should take issues related to parametrization very seriously.[20]

Our most striking finding is the asymmetry between positive and negative reciprocity in our data. The result from this experiment, that negative reciprocity is stronger than positive reciprocity, is in line with previous evidence from surveys and lab experiments. Campbell and Kamlani (1997), for instance, explore the reasons for wage rigidity and survey wage setting

[19] In related work, Al-Ubaydli et al. (2006) also find that the positive relationship between higher wages and effort (in a different task) occurs with a time lag. On the other hand, Maréchal and Thöni (2006) find evidence for *immediate* positive reciprocal reactions in a natural field experiment with sales representatives.

[20] In the taxonomy proposed by Engelmann and Ortmann (2001) the sensitivity of the results towards parametrization is called "first-degree robustness". See also Fehr et al. (2002), Rigdon (2002) or Pereira et al. (2006) for first-degree robustness checks of gift exchange in the lab.

individuals from 184 firms. They find that the "respondents expected that the effect of wages on effort would be strongly asymmetric (p. 778)" in the sense that effort is more sensitive to wage-cuts than to wage increases. This point was also put forth by Bewley (1999) and Blinder and Choi (1990). In a lab experiment, Offerman (2002) comes to the same conclusion showing that negative reciprocity as a reaction to intentionally hurtful behavior is substantively more frequent than positive reciprocal reactions to helpful behavior.[21] Offerman suggests that a self serving bias might be a potential reason for the observed asymmetry. A kind deed fits "...very well in people's positive view of themselves and is no reason for surprise." By contrast, intentionally unkind actions are "...in sharp conflict with the positive self-image of the individual (p. 1435)." Campbell and Kamlani (1997), on the other hand, consider loss aversion (i.e. the tendency to value losses more than gains[22]) as a potential explanation for the phenomenon.

Our results provide some suggestive evidence for negative reciprocal motivations as a determinant of downward wage rigidity. It seems worthwhile to asses the robustness of this effect of negative reciprocity in future research. Three aspects appear to be especially important. First, our time horizon is limited to one day of work. Extending the time horizon and testing for the robustness of our results is consequently an important next step. Secondly, we avoided to communicate any reasons for the changes in hourly wages. However, the evidence from firm surveys is broadly conclusive that wage-cuts are expected to have a higher acceptance among the workers if they are necessary to prevent bankruptcy and large scale dismissals.[23] Hence, the framing of wage-cuts will probably play an important role. And finally, if workers suffer from nominal illusion, real wage-cuts - by increasing nominal wages at a lower rate than inflation - are more likely to be possible without incurring the risk of a detrimental effect on work morale.[24]

[21] For further consonant evidence in laboratory experiments, see Brandts and Charness (1999), Engelmann and Ortmann (2001) or Thöni and Gächter (2006).

[22] See Kahneman and Tversky (1979).

[23] See Kaufman (1984), Kahneman et al. (1986), Blinder and Choi (1990) Campbell and Kamlani (1997) or Bewley (1999).

[24] See Kahneman et al. (1986), Blinder and Choi (1990) or Agell and Bennmarker (2004).

References

Agell, J. and H. Bennmarker (2004): "Endogenous Wage Rigidity", working paper.

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 (1988): "Fairness and Unemployment", *American Economic Review, Papers and Proceedings*, Vol. 83, pp. 44–49.

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): "Gift Exchange and Incentives", paper presented at the ESA European Regional Meeting in Nottingham.

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.

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.

Brandts, J. and G. . Charness (1999): "Retribution in a Cheap Talk Experiment", Working Paper, Universitat Autonoma de Barcelona.

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.

Cox, J. C., D. Friedman, and S. Gjerstad (2005): "A Tractable Model of Reciprocity and Fairness", mimeo.

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 (2001): "The Robustness of Laboratory Gift Exchange: A Reconsideration", Working Paper, Charles University.

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

Fehr, E., U. Fischbacher, and E. Tougareva (2002): "Do High Stakes and Competition Undermine Fairness? Evidence from Russia", Working Paper, University of Zürich.

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 (Forthcoming): "Fairness and Contract Design", *Econometrica*.

Harrison, G. W. and J. A. List (2004): "Field Experiments", *Journal of Economic Literature*, Vol. 42, pp. 1009–1055.

Hennig-Schmidt, H., B. Rockenbach, and A. Sadrieh (2005): "In Search of Workers' Real Effort Reciprocity - A Field and a Laboratory Experiment", mimeo.

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. and A. Tversky (1979): "Prospect Theory: An Analysis of Decision under Risk", *Econometrica*, Vol. 47, pp. 263–291.

Kaufman, R. T. (1984): "On Wage Stickiness in Britain's Competitive Sector", *British Journal of Industrial Relations*, Vol. 22, pp. 101–112.

Levitt, S. D. and J. A. List (2006): "What Do Laboratory Experiments Tell Us About the Real World?", *Working paper*.

List, J. A. (2006): "The Behavioralist Meets the Market: Measuring Social Preferences and Reputation Effects in Actual Transactions", *Journal of Political Economy*, Vol. 114, pp. 1–37.

Maréchal, M. A. and C. Thöni (2006): "Do Managers Reciprocate? Field Experimental Evidence From a Competitive Market", mimeo.

Offerman, T. (2002): 'Hurting Hurts More Than Helping Helps", *European Economic Review*, Vol. 46, pp. 1423–1437.

Pereira, P. T., N. Silva, and Andrade e Silva J. (2006): "Positive and Negative Reciprocity in the Labour Market", *Journal of Economic Behavior and Organization*, Vol. 59, pp. 406–422.

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

Rigdon, M. (2002): "Efficiency Wages in an Experimental Labor Market", *Proceedings of the National Academy of Sciences*, Vol. 99, pp. 13 348–13 351.

Thöni, C. and S. Gächter (2006): "Social Interaction in the Workplace", mimeo.

5 Appendix

Figure 1: Number of Books Logged per Time Interval by Treatment

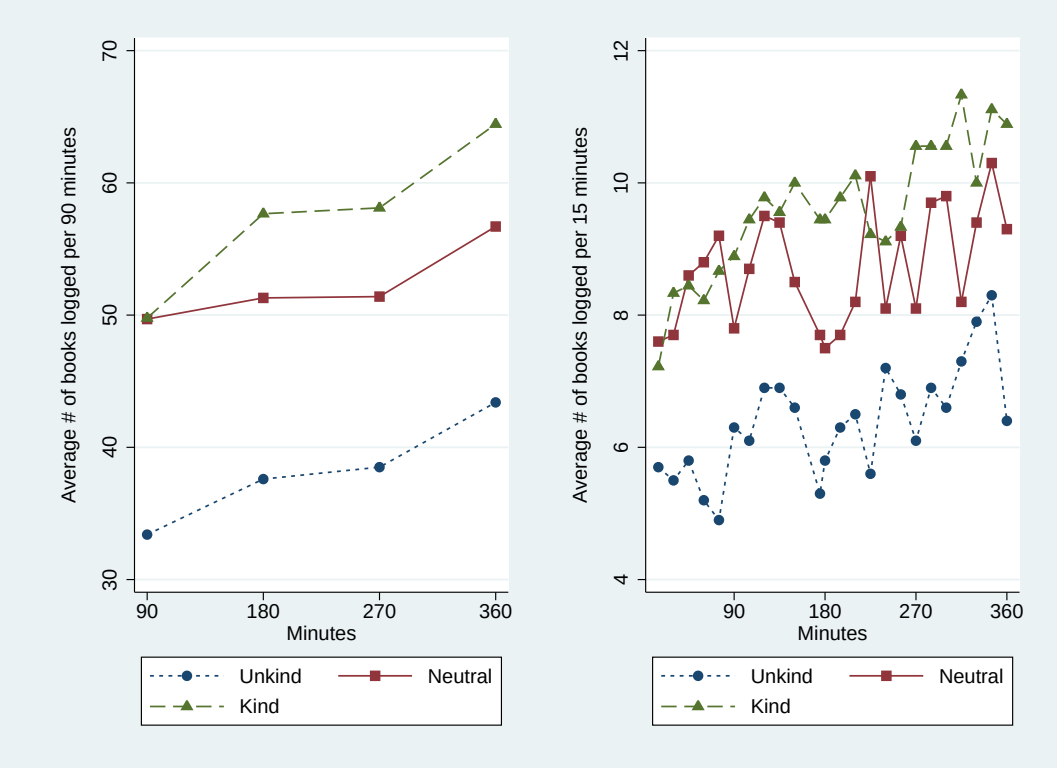

Table 1: Average Treatment Effects by Timeunits: # Books Logged

Time interval	(1) ***Kind - Neutral***	(2) $p > \|z\|$	(3) ***Unkind - Neutral***	(4) $p > \|z\|$
Minutes 1-90	0.1	0.623	-16.3	0.002
Minutes 91-180	6.4	0.164	-13.7	0.010
Minutes 181-270	6.7	0.094	-12.9	0.006
Minutes 270-360	7.7	0.347	-13.3	0.017
Total regular time	20.9	0.165	-56.2	0.003

Note: Columns 1 and 3 report Average Treatment Effect estimates for the treatments *Kind* (20 Euros hourly wage) and *Unkind* (10 Euros hourly wage) in comparison with the benchmark *Neutral* (15 Euros hourly wage) for the different time intervals. The outcome variable is the number of books logged as a measure for the effort provided. Columns 2 and 4 report the corresponding p-values from a nonparametric (two-sided) Wilcoxon rank-sum test for the null hypothesis of equal effort between treatments.

Table 2: Panel Regression Results

	(1)	(2)	(3)
Number of books logged	**90 minutes intervals** Coeff./Rob.se	**15 minutes intervals** Coeff./Rob.se	**Faultless logs** Coeff./Rob.se
Kind	0.078	0.029	1.544
	(4.670)	(0.728)	(4.611)
Unkind	-16.300***	-2.662***	-14.500***
	(5.014)	(0.764)	(5.082)
Time2	1.600		0.500
	(2.574)		(2.872)
Time3	1.700		2.200
	(2.664)		(2.693)
Time4	7.000**		8.500**
	(2.968)		(3.349)
Kind $*$ *Time2*	6.289**		4.833
	(3.069)		(3.312)
Kind $*$ *Time3*	6.633**		4.356
	(3.356)		(3.330)
Kind $*$ *Time4*	7.667*		1.722
	(4.491)		(4.636)
Unkind $*$ *Time2*	2.600		2.900
	(3.766)		(4.106)
Unkind $*$ *Time3*	3.400		2.600
	(3.836)		(3.836)
Unkind $*$ *Time4*	3.000		0.400
	(3.877)		(4.135)
Time		0.054**	
		(0.021)	
Kind $*$ *Time*		0.067**	
		(0.031)	
Unkind $*$ *Time*		0.026	
		(0.027)	
Constant	49.700***	8.040***	45.900***
	(3.681)	(0.591)	(3.829)
Obs.	116	696	116
R^2	0.429	0.250	0.387
Wald χ^2	1011.9	72.7	952.5
Prob$>\chi^2$	0.000	0.000	0.000

Note: The table reports GLS coefficient estimates (robust standard errors in parentheses) from a linear Random Effects model. In columns (1) and (2) the dependent variable is the number of books logged per time interval. In column (3) the dependent measure is the number of faultlessly logged books per 90 minutes time interval. The treatment dummies *Unkind* (10 Euros hourly wage) and *Kind* (20 Euros hourly wage) are interacted with the time interval dummies $Timet$ (or the variable $Time$ which takes values from 1 to 24 in the high frequency data). In column one *Neutral* (15 Euros hourly wage), *Time1*, $Kind * Time1$ and $Unkind * Time1$ are considered as reference categories. Significance levels are denoted as follows: * p<0.1, ** p<0.05, *** p<0.01.

Table 3: Summary Data: Number of Books Logged During 90 Minutes Intervals and 60 Minutes of Overtime

Treatment	**Participant #**	**Minutes 1-90**	**Minutes 91-180**	**Minutes 181-270**	**Minutes 270-360**	**Total regular time**	**Overtime 60 minutes**	**Quality ratio regular time**
Neutral	26	44	39	39	42	164		0.939
(15 Euros)	9	45	45	46	41	177		0.921
	4	38	42	49	60	189		0.915
	8	47	47	48	53	195		0.990
	16	45	52	50	61	208		0.947
	19	57	55	46	52	210	39	0.976
	28	44	56	58	55	213		0.934
	12	55	47	64	57	223	42	0.803
	2	49	65	59	67	240	41	0.958
	21	73	65	55	79	272		0.938
	Average	*49.7*	*51.3*	*51.4*	*56.7*	*209.1*	*40.7*	*0.932*
Kind	3	42	46	50	42	180		0.900
(20 Euros)	5	40	45	43	52	180		0.978
	24	48	54	53	56	211		0.867
	20	45	56	61	59	221	39	0.964
	17	58	62	51	55	226		0.978
	13	53	64	60	73	250		0.884
	11	50	65	61	76	252		0.853
	27	55	59	69	76	259		0.931
	22	57	68	75	91	291	57	0.945
	Average	*49.8*	*57.7*	*58.1*	*64.4*	*230*	*48.0*	*0.922*
Unkind	25	21	22	9	27	79	17	0.937
(10 Euros)	10	23	21	29	27	100		0.980
	6	38	30	35	31	134	23	0.925
	1	32	38	38	39	147	24	0.986
	18	27	34	44	44	149	33	0.980
	15	32	38	42	50	162		0.932
	29	31	35	45	52	163		0.804
	14	30	50	37	50	167	31	0.940
	7	38	40	48	50	176	36	0.869
	23	62	68	58	64	252	34	0.984
	Average	*33.4*	*37.6*	*38.5*	*43.4*	*152.9*	*28.3*	*0.934*