No. 48

UNIVERSITY OF COLOGNE

WORKING PAPER SERIES IN ECONOMICS

WAGE TRANSPARENCY AND PERFORMANCE: A REAL-EFFORT EXPERIMENT

BEN GREINER
AXEL OCKENFELS
PETER WERNER

Department of Economics
University of Cologne
Albertus-Magnus-Platz
D-50923 Köln
Germany

http://www.wiso.uni-koeln.de

Wage Transparency and Performance: A Real-Effort Experiment

Ben Greiner[*], Axel Ockenfels[†], and Peter Werner[†,‡]

January 20, 2011

Abstract

Without transparency about peer wages in a real effort experiment, a change of wages does not affect performance. With transparency, however, higher paid workers tend to work more accurately, and lower paid workers shirk more under piece rates.

Keywords: labor market experiments, real effort, social comparison, wage schemes

JEL Classification: C91, J33, M52

1. Introduction

Wage distributions within firms are often more compressed than implied by standard economic theory (Frank, 1984a). One explanation that has been put forward is that workers care about their relative position (Frank, 1984b; Akerlof and Yellen, 1990; Bewley, 1999). Indeed, laboratory studies generally provide evidence for a detrimental effect of disadvantageous relative positions (see Burchett and Willoughby, 2004; Clark et al., 2010; Gächter and Thöni, 2010) and high wage dispersion (Rivas, 2009) on real or hypothetical effort exertion. Our paper extends this literature by investigating the role of transparency of wage dispersion after wage changes. In our experiment, identical wages are paid for the first part of an experimental real-effort task. In an otherwise identical second part, we then introduce wage differentials, which are either transparent (that is, known to workers) or not. Moreover, we investigate the role of transparency under both a fixed and a piece rate wage scheme.

We find that introducing wage differentials yields performance changes only when being transparent. In this case, high wage subjects increase their efforts, while low earners under piece rates increase work quantity at the expense of quality, and in this sense become more prone to shirking. Contrary, under purely private wage

[*]University of New South Wales, School of Economics, Sydney, NSW 2052, Australia.

[†]University of Cologne, Department of Economics, D-50923 Köln, Germany.

[‡]Corresponding author. Tel: +49/221/470-4354, Fax: +49/221/470-5068.

E-Mail addresses: bgreiner AT unsw.edu.au (B. Greiner), ockenfels AT uni-koeln.de (A. Ockenfels), peter.werner AT uni-koeln.de (P. Werner).

information we do not observe performance differences between high wage and low wage workers.

In Section 2 we present the details of our experimental design and introduce our hypotheses. Our results are reported in Section 3 and we briefly conclude in Section 4.

2. Experimental Design and Hypotheses

The experiment consisted of two parts of a laborious working task. In each part participants were handed out a pile of 50 hardcopy questionnaire forms.[1] Each form contained an ID number and 15 three-digit decimal numbers, filled in by participants in a previous experiment. The task was to copy those numbers into an input mask at the computer screen. Each of the two parts lasted 20 minutes.[2] At the beginning of each session, subjects were randomly assigned to either group A or group B. For the first part, all subjects were paid identical wages. In the second part, each participant in group A received a 60% wage increase, while the wages of participants in group B were cut by the same share. Wage changes were not related to any performance measure, and subjects could not assess their own performance relative to others.

We implemented a 2x2 design. In all conditions, A (B) participants were informed that they were assigned to a B (A) participant working on the exact same set of questionnaires, which would allow us to easily conduct consistency checks and search for mistakes after the experiment. Under *private information*, participants were only told their own wage. In the *public information* information condition, participants were also told the other participant's wage. In the second design dimension, participants either received a *flat wage* for both parts, or a *piece rate wage* only depending on quantity (the number of completed forms) but not on quality (the share of correctly transcribed forms).[3]

Workers in all conditions were informed that the correctness of entries would not be checked prior to payment. So, assuming that delivering correct entries is more costly than delivering random entries, our *Hypothesis 1 [effort costs]* is that quality is close to zero in all conditions.[4] Also, because there is a higher incentive to transcript forms under piece rates than under flat rates, we hypothesize that quantities are larger under piece rates (*Hypothesis 2 [piece rates]*), as observed by, e.g., Lazear (2000) in a field context.

From a standard economic point of view, our changes of wages in the second part of the experiment should not affect our Hypotheses 1 and 2. However, gift-exchange and empirical evidence (Bellemare and Shearer, 2009) suggest that paying lower

[1] 100 questionnaire forms were randomly distributed over both parts. All participants received the same set of forms.

[2] Instructions translated from German can be found in Online Appendix A available at http://ben.orsee.org/supplements/.

[3] Since from an employer perspective it might also be reasonable to make the piece rate contingent upon single entries rather than upon complete forms, we note that the results described in Section 3 are similar if we had used single entries for our analysis.

[4] The probability that a single 3-digit random number is correct is 1/1000, the probability that a complete form of 15 numbers is correct under randomness is 10^{-45}.

(higher) wages may crowd out (increase) subjects' willingness to exert effort and subsequently work quality (*Hypothesis 3 [wage changes]*).[5] If, additionally, agents care about relative positions, one might expect that performance depends on what is known about the wage differences in Part 2. More specifically, we hypothesize that the responses to wage changes as described by Hypothesis 3 are stronger under public information (*Hypothesis 4 [transparency]*).

The experiment took place in early 2006 at the Cologne Laboratory for Economic Research. Participants were invited using ORSEE (Greiner, 2004). We conducted four sessions with altogether 126 subjects (mostly students in Business or related fields), each session lasting approximately one hour. The computerized input mask was implemented utilizing the z-Tree software (Fischbacher, 2007). Participants in the piece rate treatments received 0.20 Euros per completed form in the first part, and 0.08 or 0.32 Euros in the second part. To ensure comparability of overall financial incentives, wages in the fixed rate treatments were set equal to the average payoff in the piece-rate conditions. Specifically, fixed wages were 6.80 Euros in the first part, and 2.70 or 10.90 Euros in the second part. The overall average payoff was 16.10 Euros (accumulated over both parts), including a show-up fee of 2.50 Euros.

3. Experimental Results

Despite incentives to cheat, work quality is generally high: the average share of correct forms under flat wages accounts for 75.7% in Part 1 and 82.6% in Part 2; the corresponding values under piece rates are 71.8% and 74.6%, respectively. Thus, Hypothesis 1 can be rejected.

Regarding Hypothesis 2, we find only weak evidence. Under flat wages, participants complete on average 29.0 forms in Part 1 and 32.6 forms in Part 2, whereas under piece rates, subjects fill in 31.2 forms in Part 1 and 36.8 forms in Part 2.[6] Comparing individual working quantities between the payment schemes with two-sided Mann-Whitney-U (MWU) tests yields $p = .128$ for Part 1 and $p = .010$ for Part 2.

Coming to Hypothesis 3, we first observe that average working quality and quantity increase significantly in almost all experimental groups from Part 1 to Part 2, indicating experience effects.[7] In order to isolate the incentive effects of the wage differential from experience effects, we compare relative performance changes of subjects (measured in percent of Part 1 performance). Using two-sided MWU tests we find no significant differences between low and high wage earners under *private information* – neither under piece rates nor under fixed wages (all p-values $> .1$). Thus, our experiment provides no statistically significant support for Hypothesis 3

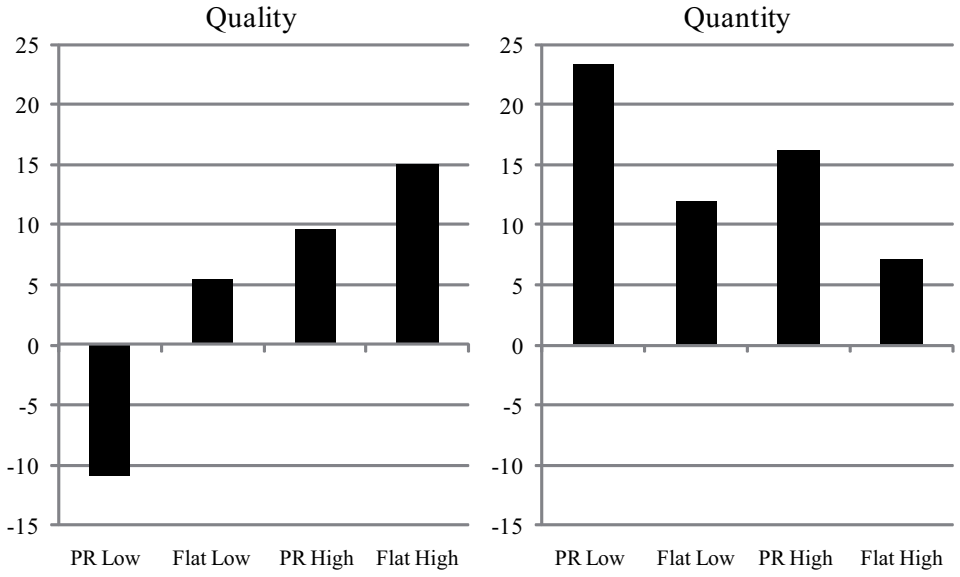
[5] Some experimental studies find that wage change effects are not persistent (Gneezy and List, 2006) and that positive effects of wage increases are weaker than negative responses to wage cuts (Kube et al., 2010).

[6] For a complete set of descriptive statistics see Online Appendix B available at http://ben.orsee.org/supplements/.

[7] Two-sided Wilcoxon Matched Pairs Signed Ranks (WMPSR) tests yield a significant increase in quantity in 7 out of 8 treatment groups ($p < .002$), and a significant quality increase in 5 of those groups ($p < .05$).

postulating that workers respond reciprocally to wage changes (see Hennig-Schmidt et al., 2010, who also found no gift-exchange effect in the context of a real-effort experiment).

FIGURE 1: Relative Performance Change per Treatment under Public Information in %



Note: "PR" and "Flat" stand for piece rate and flat wages, and "High" and "Low" for high and low wages, respectively.

However, in line with Hypothesis 4, high and low wage workers adjust their performance differently under *public information*. As Figure 1 shows, under piece rates low wage subjects lack behind with respect to quality ($p = .019$, MWU) while increasing quantity stronger than high wage subjects ($p = .060$, MWU). In the fixed wage treatment, differences between high and low wage subjects have the same signs as in the piece rate treatment, but are not significant at conventional levels.

A closer look at work quality reveals that participants either completed a form diligently, with few mistakes, or did not put in much effort at all, resulting in many mistakes in the same form. We classify a form as "shirked" if the form had 6 or more wrong fields, a threshold which is about two standard deviations (SD = 2.19) above the average number of mistakes per form of 0.70.[8] Incorrect forms with 1 to 5 mistakes were classified as "inaccurate".

We observe that low earners under piece rates largely reduce their efforts in Part 2. First, the average share of shirked form transcriptions increases weakly significantly from 2.1% in Part 1 to 6.5% in Part 2 ($p = .094$, WMPSR). Moreover, the average proportion of inaccurate forms increases significantly from 22.6% to 24.1% ($p = .005$, WMPSR). With flat wages we find no corresponding effect among

[8] Our results are robust against varying thresholds and alternative classifications of shirked forms.

low earners. There is no significant increase of shirked forms (0.2% versus 1.7%, $p >$ .1, WMPSR), and the share of inaccurate forms becomes (insignificantly) smaller (25.1% and 18.8%, $p = .125$, WMPSR).

Highly paid subjects tend to put more effort into quality. The shares of inaccurate forms drop (weakly) significantly from 23.6% to 15.1% under piece rates and from 22.6% to 10.9% under flat wages ($p = .061$ and $p = .001$, respectively, WMPSR), whereas the shares of shirked form transcriptions remain roughly constant ($p > .1$, WMPSR).[9]

TABLE 1: OLS Regressions of Individual Performance in Part 2

Dependent Variable			Quality in Part 2		Quantity in Part 2	
Treatment dummies						
Piece Rate	Public Info	High	0.241***	[0.067]	7.815***	[1.795]
Piece Rate	Public Info	Low	0.103	[0.067]	11.323***	[1.956]
Piece Rate	Private Info	High	0.166**	[0.064]	8.643***	[2.052]
Piece Rate	Private Info	Low	0.219***	[0.063]	7.594***	[1.726]
Flat Wage	Public Info	High	0.289***	[0.068]	4.947***	[1.754]
Flat Wage	Public Info	Low	0.209***	[0.067]	6.348***	[1.763]
Flat Wage	Private Info	High	0.221***	[0.067]	7.808***	[1.837]
Flat Wage	Private Info	Low	0.207***	[0.069]	7.295***	[1.788]
Quality Part 1			0.785***	[0.078]		
Quantity Part 1					0.895***	[0.048]
N			126		126	
R-squared			0.99		0.98	

Standard errors are given in brackets. ** and *** denote significance on the 5% and 1%-level.

Our conclusions do not change if we take into account measures for individual ability. Table 1 lists simple linear regression models with Part 2 performance (quality and quantity) as the dependent variable. To control for ability we include performance in Part 1, which is, as expected, a strong and highly significant predictor in both specifications. Additionally we include dummies for each experimental treatment condition (the constant is suppressed). Most dummy variables are positive and significant. Comparing the sizes of regression coefficients with Wald tests confirms our previous statements. Estimated coefficients for high and low wage workers do not differ under private information. Under public information and piece rates, low wage subjects lack behind with respect to quality ($p = .003$) and increase quantity stronger than highly paid subjects ($p = .032$). Under flat wages and public information, we observe a weakly significant difference between high and low wage subjects with respect to quality ($p = .087$).

[9]The shares of shirked forms in Part 1 (Part 2) are 2.1% (2.4%) for high earners under piece rates, and 1.8% (0.8%) for high earners under flat wages.

4. Conclusions

Summing up, we find little evidence for the importance of wage changes for the performance of our laboratory workers. However, we provide controlled laboratory evidence that the transparency of wage dispersion strongly affects performance: with public information, increasing wages promotes effort of high earners, and lowering wages leads to more shirking. The performance effects are mitigated under flat wages.

That said, we caution that our findings are based on a laboratory experiment, employing subjects who know that they are participating in an experiment. This limits what we can say about the external validity of our laboratory study (see Levitt and List, 2007). Also, even though participants in our experiment exert real efforts (rather than just choosing a hypothetical effort level linked to their payoffs), wages and wage cuts in our experiment do not share many characteristics that seem typical for wages and how they are negotiated in usual labor market relationships.[10] On the other hand, we emphasize that our results do not only seem to have intuitive appeal, but are also complemented by recent findings from naturally occurring field data: Ockenfels et al. (2010) found that while transparent bonus differences hamper satisfaction and performance of those workers who fall behind others, intransparent bonus differences do not.[11] We believe that, taken together, this complementary research makes a strong case that wage transparency can systematically matter, both in the laboratory and the field.

Acknowledgements

Financial support from the German Science Foundation (DFG) is gratefully acknowledged. We thank an anonymous referee for helpful comments and suggestions.

References

Akerlof, G.A., Yellen, J.L., 1990. The fair wage-effort hypothesis and unemployment. Quarterly Journal of Economics 105, 255–283.

Bellemare, C., Shearer, B., 2009. Gift giving and worker productivity: Evidence from a firm-level experiment. Games and Economic Behavior 67, 233–244.

Bewley, T.F., 1999. Why wages don't fall during a recession. Harvard University Press, Cambridge.

Burchett, R., Willoughby, J., 2004. Work productivity when knowledge of different reward systems varies: Report from an economic experiment. Journal of Economic Psychology 25, 591–600.

Charness, G., Kuhn, P., 2010. Lab Labor: What can Labor Economists Learn from the Lab? IZA Discussion Paper no. 4941.

[10] Other labor market experiments share similar qualifications and discussed related issues; see Fehr et al. (2009) and Charness and Kuhn (2010) for recent surveys of laboratory labor markets; see List and Rasul (2010) for discussing some of the advantages of using field experiments in the context of labor markets.

[11] There is also anecdotal evidence, such as Frank and Sunstein (2001) who observe that some companies create "a norm against public discussion of salaries, on the theory that people are likely to be happy enough with their own, but would be less happy if they found themselves making comparisons to others".

Clark, A.E., Masclet, D., Villeval, M.C., 2010. Effort and comparison income: Experimental and survey evidence. Industrial and Labor Relations Review 63, 407–426.

Fehr, E., Götte, L., Zehnder, C., 2009. A Behavioral Account of the Labor Market: The Role of Fairness Concerns. Annual Review of Economics 1, 355–384.

Fischbacher, U., 2007. z-Tree: Zurich toolbox for ready-made economic experiments. Experimental Economics 10(2), 171–178.

Frank, R.H., 1984a. Are workers paid their marginal products? American Economic Review 74, 549–571.

Frank, R.H., 1984b. Interdependent preferences and the competitive wage structure. The Rand Journal of Economics 15, 510–520.

Frank, R.H., Sunstein, C., 2001. Cost benefit analysis and relative position. University of Chicago Law Review 68, 323–374.

Gächter, S., Thöni, C., 2010. Social comparison and performance: Experimental evidence on the fair wage-effort hypothesis. Journal of Economic Behavior and Organization 76(3), 531–543.

Gneezy, U., List, J.A., 2006. Putting behavioral economics to work: Testing for gift exchange in labor markets using field experiments. Econometrica 74, 1365–1384.

Greiner, B., 2004. An online recruitment system for economic experiments, in: Kremer, K., Macho, V. (Eds.), Forschung und wissenschaftliches Rechnen 2003, GWDG Bericht 63. Ges. für Wiss. Datenverarbeitung, Göttingen, pp. 79–93.

Hennig-Schmidt, H., Rockenbach, B., Sadrieh, A., 2010. In search of workers real effort reciprocity - a field and a laboratory experiment. Journal of the European Economic Association 8, 817–837.

Kube, S., Marèchal, M.A., Puppe, C., 2010. Do wage cuts damage work morale? Evidence from a natural field experiment. IEW Working Papers no. 471.

Lazear, E., 2000. Performance pay and productivity. American Economic Review 90(5), 1346–1361.

Levitt, S.D., List, J.A., 2007. What Do Laboratory Experiments Measuring Social Preferences Reveal About the Real World? Journal of Economic Perspectives 21(2), 153–174.

List, J.A., Rasul, I., 2010. Field Experiments in Labor Economics. NBER Working Paper 16062.

Ockenfels, A., Sliwka, D., Werner, P., 2010. Bonus payments and reference point violations. IZA Discussion Paper no. 4795.

Rivas, F., 2009. Wage dispersion and workers' effort. Economics Bulletin 29(2), 788–794.