When Equality is Unfair – Social Norms Between and Within Levels of a Hierarchy[†]

Johannes Abeler[*], Steffen Altmann[*], Sebastian Kube[**], Matthias Wibral[*]

April 19, 2006

Abstract:

We analyze experimentally the role of vertical and horizontal fairness in labor relations with implicit contracts. We focus on two prominent wage schemes in firms with more than one agent: wage equality as an extreme form of wage compression and individual wage setting. We find that efforts and efficiency are significantly higher under individual wages. This is not caused by differences in monetary incentives since it is profit-maximizing for agents to exert high efforts under both wage schemes. The difference in effort provision is rather due to a violation of social norms within the firm: under equal wages, the agent exerting the higher effort receives a lower payoff as effort is costly. Reciprocal agents suffering from this norm violation get discouraged and decrease their effort. Our results suggest that individual reward and punishment opportunities are crucial for making reciprocity a powerful contract enforcement device.

JEL classification: C92, J33, J41, M12, M52.

Keywords: laboratory experiment, wage setting, gift exchange, reciprocity, social norms, implicit contracts, multiple agents.

[†] Financial support from the Bonn Graduate School of Economics is gratefully acknowledged. We thank Rachel Croson, Lucie Dörner, Ernst Fehr, Uri Gneezy, David Huffman, Michael Kosfeld, Peter Kuhn, Sandra Ludwig, Felix Marklein, Martin Sefton, Ferdinand von Siemens and especially Armin Falk, Wolfgang Luhan, and Philipp Wichardt for helpful discussions. Valuable comments were also received from seminar participants in Alessandria, Barcelona, Bonn, Budapest, Jena, Karlsruhe, and Zürich.

[*]IZA Bonn and University of Bonn, Schaumburg-Lippe-Str. 7–9, 53113 Bonn, Germany. email:[last name at]iza.org

[**]University of Karlsruhe, Department of Economics, Kollegium IV am Schloss, 76128 Karlsruhe, Germany. email:kube[at]wior.uni-karlsruhe.de

"To treat people fairly you have to treat people differently."

Roy Roberts, at that time VP of General Motors[1]

1 Introduction

Traditional economic theory posits that institutions influence decisions by imposing constraints under which purely selfish individuals maximize their utility. There is ample evidence, however, that in many situations individuals are also guided by social norms or preferences such as reciprocity or equality. If this is the case the actual incentives shaped by an institution arise from the interaction of this institution with the social preferences. Therefore social preferences and institutions have to be analyzed jointly.

Norms and social preferences are particularly important when explicit contract enforcement mechanisms are weak or non-existent, which is often the case in the labor market. For bilateral work relations, experimental research has impressively shown that reciprocity is an effective means to enforce implicit contracts (e.g., Fehr et al. 1997, Fehr and Falk 2002, Hannan et al. 2002).

The more complex (and more realistic) situation of multilateral work relations, in contrast, is not well studied. The interaction of one principal with more than one agent adds another dimension that has to be taken into account: the vertical relationship between principal and agent is now potentially influenced by the horizontal relationship between the agents. While there seems to be general agreement—at least among practitioners (e.g., Bewley 1999)—that concerns about co-workers can greatly influence the agents' performance, the implications for compensation policy are not well understood.

In this paper, we analyze wage compression as a prototypical example for the interplay of institutions and social norms. It has been argued that compressed wages reduce envy and resentment within the workforce and improve work morale and thus productivity (Frank 1984, Akerlof and Yellen 1990). We therefore compare experimentally the performance of two wage schemes that are polar opposites with respect to the enforced degree of horizontal equality—wage equality as an extreme case of wage compression and full discretion over

[1] Quoted in Baker et al. (1988).

wages. We are interested in how the different dimensions of social preferences influence the efficiency of the two wage schemes (i.e., the performance of agents that they elicit).

In our set-up, a firm consists of one principal and two agents. At first, the agents exert costly effort. The principal pays them a wage after observing their efforts. In one treatment he can wage discriminate between the two agents, in the other he is obliged to pay the same wage to both agents. Neither effort nor wages are contractible, allowing us to concentrate on the influence of fairness concerns with respect to both the horizontal and vertical dimension.

The experiment yields the following findings. First, the two wage institutions at hand exhibit dramatic differences with respect to the performance they elicit: efforts under individual wages are significantly higher than efforts under equal wages. In addition, efforts decline over time when equal wages are paid. Second, this strong treatment effect is not caused by differences in monetary incentives. From a purely monetary viewpoint it is profitable in both treatments to exert high efforts given the actual wage distribution in the experiment. Third, we show that, instead, the frequent violation of a norm of "multilateral reciprocity" in the equal wage treatment can explain the difference between the treatments. Multilateral reciprocity requires that the principal rewards a higher effort with a higher payoff and also that, within a firm, the agent who works more gets a higher payoff compared to his co-worker. Reciprocal agents who are initially willing to exert high effort get discouraged by these norm violations and choose lower effort levels under equal wages. In contrast, the individual wage scheme leaves multilateral reciprocity mostly intact, resulting in a high level of efficiency.

Our results show that it is important to carefully analyze the interactions between social norms and monetary incentives when evaluating the effectiveness of incentive schemes or institutions in general. Regarding compensation practice in firms, our findings highlight the importance of taking the horizontal dimension of social preferences into account. However, doing so by paying equal wages to a group of agents may actually do more harm than good because equal wages entail the wrong reciprocity signals and induce reciprocal agents to provide lower effort. In light of our results it seems crucial to maintain the possibility to individually sanction bad performance and reward good performance for reciprocity to work as a contract enforcement device.

The paper contributes to several strands of the literature. In the traditional literature on incentive provision in groups, the inefficiency of equal wages stems from the fact that marginal products and wages are not aligned. This can lead to free-riding between selfish agents (Holmström 1982, Erev et al. 1993). Our experiment, in contrast, suggests a psychological rationale for using individual wages. Since the subjects perceive equal wages as unfair, it is the presence of reciprocal agents and not their absence that makes equal wages inefficient.

Our findings also inform the growing literature on behavioral contract theory (e.g., Kandel and Lazear 1992, Demougin and Fluet 2003, Fehr et al. 2004, Bartling and von Siemens 2004, Rey-Biel 2004, Dur and Glazer 2004). These models rely on assumptions regarding the players' reference groups, the importance of different variables for social comparisons, and the channels through which these comparisons influence decisions. Our results provide input for the evaluation and improvement of the existing theories along these lines. Moreover, they demonstrate that experiments can be effectively used to study such questions in complex, multilateral environments.

We also contribute to the literature analyzing the influence of relative income on satisfaction and performance. It has been shown that relative income affects people's well-being (e.g., Clark and Oswald 1996, Solnick and Hemenway 1998, Easterlin 2001, Frey and Stutzer 2002). However, it is less clear how this influences performance, i.e., whether low relative income leads to frustration and reduced performance (Clark et al. 2006, Torgler et al. 2006) or to an increase in performance due to a "positional arms race" (Neumark and Postlewaite 1998, Layard 2002, Bowles and Park 2005). The controlled laboratory environment of our experiment allows us to reconcile these differing views. Our results suggest that the comparison process goes beyond a one-dimensional comparison of income. Specifically, we argue that performance is not influenced by relative income per se but depends on whether social norms are violated or intact.

There are only a few experimental studies in which one principal interacts with several agents (Rossi and Warglien 2001, Güth et al. 2001, Meidinger et al. 2003, Maximiano et al. 2004, Thöni and Gächter 2005). Most closely related to our paper is the work of Charness and Kuhn (2005). In contrast to our results, they find that co-workers' wages do not matter much for agents' decisions. However, their design differs from ours in several

important points. While Charness and Kuhn focus on productivity differences, we look at the effect of actual output differences between agents. Furthermore, we allow for richer comparisons between the agents, as in their design agents are not aware of the magnitude and direction of the productivity differences. The difference in results points out the importance of reference groups: in our view, Charness and Kuhn's results are important for groups of workers that are rather loosely related and know little about each other, while our focus is on direct colleagues who have a good understanding about their peers' abilities and efforts.

Also related is the study of Potters et al. (2004). For a joint production task they find that it is efficiency enhancing to give one group member the right to allocate rewards in comparison to a revenue-sharing institution. Their results suggest that the general mechanisms at work in our experiment do not depend on our specific set-up.

The rest of this paper is structured as follows. In the next section we describe the experimental design. Section 3 discusses theoretical predictions. In section 4 we present and discuss our results and section 5 concludes.

2 Experimental Design

In the stage game of our experiment, one principal and two agents form a firm. First, the agents decide simultaneously and independently how much effort they want to exert. Exerting effort is costly for the agents. Effort choices can range from 1 to 10 and are associated with a convex cost function displayed in Table 1. The principal reaps the benefits of production: every unit of effort increases his payoff by 10.

Effort level e_i	1	2	3	4	5	6	7	8	9	10
Cost of effort $c(e_i)$	0	1	2	4	6	8	10	13	16	20

Table 1: *Cost of effort. The marginal cost of effort rises from 1 (for efforts 1–3) to 4 (for efforts 9–10).*

After observing the effort decisions of the agents in his firm the principal decides on wages for the two agents. Wages can range from 0 to 100. Neither efforts nor wages

are contractible. Treatments differ in whether the principal is allowed to pay different wages to the agents or not. In the equal wage treatment the principal can only choose one wage w that is paid to each of the agents. In the individual wage treatment he can wage discriminate between the two agents by choosing wages w_1 and w_2 for agent 1 and 2, respectively. At the end of each period, every member of a firm is informed about efforts, wage(s), and the resulting payoffs for everyone in the firm. The payoff functions for the players are summarized in Table 2. In the following, the individual wage treatment is abbreviated as IWT and the equal wage treatment as EWT.

Treatment	EWT	IWT
Payoff Principal	$\pi_P = 10(e_1 + e_2) - 2w$	$\pi_P = 10(e_1 + e_2) - (w_1 + w_2)$
Payoff Agent i	$\pi_{A_i} = w - c(e_i)$	$\pi_{A_i} = w_i - c(e_i)$

Table 2: *Payoffs of players*

This stage game is played for twelve periods. Our set-up differs from other studies of gift exchange in two important ways. Most importantly, a principal is matched with two agents instead of one. This allows us to analyze the interaction of wage scheme and social preferences in a much richer environment. In particular, we can study both fairness concerns between principal and agent and concerns about co-workers. Additionally, the agents move first while in most experiments the principal moves first. Our move order allows the principal to base his wage decision on the actually exerted effort and has important implications for efficiency (see section 4.4).[2]

The experiment was conducted in a labor market framing.[3] We implemented a stranger design to abstract from confounding reputation effects, i.e., at the beginning of each period new firms were formed anonymously and randomly within a matching group. A matching group consisted of three principals and six agents. The subjects kept their roles during the whole experiment.

All participants started the experiment with an initial endowment of 400 points that also

[2]For an experiment in which principals decide according to productivity differences, see Charness and Kuhn (2005).

[3]An English translation of the instructions is available from the authors upon request.

served as their show-up fee. Points earned were converted at an exchange rate of 0.01 Euro/point. The experiment was conducted at the BonnEconLab at the University of Bonn in April 2005. For each treatment, four sessions with a total of 8 matching groups were carried out (144 participants). The experiment lasted approximately 70 minutes. On average subjects earned 8.30 Euro including the show-up fee of 4 Euro.

3 Theoretical Predictions

In the unique subgame perfect Nash equilibrium of our stage game, a rational and selfish principal will not pay anything to the agents since wage payments only reduce his monetary payoff. Anticipating this, both agents will provide the minimal effort of one in the first stage although it would be efficient to exert the maximal effort. The finite repetition of the stage game in randomly rematched firms does not change this prediction. The equilibrium is the same for both treatments.

However, many experiments have shown that this equilibrium is almost never played. In gift-exchange experiments one typically observes that efforts and wages exceed the smallest possible value. Moreover, wages and efforts are positively correlated.

This is often explained with a notion of social preferences, i.e., players are not completely selfish but also care about the payoffs of the other players. One prominent approach assumes that players dislike unequal payoffs, e.g., Fehr and Schmidt (1999). Their model is consistent with many stylized facts from experiments with bilateral interaction. Using their preferred parameter constellation for the multilateral setting at hand, the model predicts the same outcome as the subgame perfect Nash equilibrium: minimal wages ($w = 0$) and minimal efforts ($e = 1$). Assuming extreme values for the guilt parameter ($\beta > 2/3$ for 40% of subjects) does not change this result much. In this case, the model predicts an average effort of 1.6 in both treatments and slightly higher wages than in the subgame perfect equilibrium. However, with these parameters, the predictions are still the same for both treatments.[4]

The positive correlation between efforts and wages can also be interpreted as an expression

[4] A derivation of the equilibria is provided in the appendix.

of reciprocity, i.e., one party rewards kind actions of the other party. The implications of reciprocity are clear for the bilateral gift-exchange, however, the extension of the concept to a situation with a horizontal and a vertical dimension of interaction is not straightforward. Now reciprocity could imply two things.

First, it could be that reciprocity only applies to the vertical relationships between principals and agents. Since this neglects the existence of the second agent we call this *bilateral reciprocity*. For our set-up it implies that a higher effort should be rewarded with a higher wage. Since a higher wage works towards norm fulfillment *and* provides monetary incentives we should observe a treatment difference in effort provision only if the monetary incentives between the treatments differ. In particular, if the wage-effort relations are so steep in both treatments that high efforts are payoff-maximizing, we should observe no treatment difference in efforts due to bilateral reciprocity concerns.

Alternatively, all players could be concerned about all relationships within a firm, in particular about how the principal treats the agents relative to each other (*multilateral reciprocity*). This concept of reciprocity is stronger than bilateral reciprocity as it does not only prescribe that a higher effort is rewarded with a higher wage. It additionally requires that the agent who works more also gets a higher payoff compared to his co-worker, i.e., that agents' ranking by effort and by payoff are equal.[5] If an agent exerts a higher effort than his co-worker but earns a lower profit, he feels exploited. But his co-worker also experiences some disutility in this firm. He works less but earns more than his colleague, so he feels guilty about this undeserved profit. To avoid this disutility, an exploited agent should decrease his effort in period $t+1$ and his period-t co-worker should increase his effort.[6]

It is reasonable to assume that a disadvantageous norm violation (exploitation) is experienced more strongly than an advantageous norm violation (guilt).[7] One should therefore

[5]Note that this concept of reciprocity does not solely rely on intentions (of the principal) but rather captures procedural fairness more generally (cf. Dufwenberg and Kirchsteiger 2004 or Falk and Fischbacher 2006 for formal models of reciprocity).

[6]In the experiment we employed a stranger matching, so to be formally correct one should think of agents having a belief about the average effort in the agent population. By observing the effort of their co-worker, agents update this belief and choose their effort in the next period accordingly.

[7]For bilateral relationships, support for this assumption can be found in, e.g., Loewenstein et al.

expect the effort increase of agents who face an advantageous norm violation to be less strong than the effort decrease to a disadvantageous violation. This will result in an overall negative trend of effort over periods if the norm is frequently violated.

It is likely that treatments differ with respect to the fulfillment of multilateral reciprocity. In the equal wage treatment, principals are forced to pay the same wage to both agents. Thus the norm of multilateral reciprocity is violated every time agents choose different effort levels. On the contrary, in the individual wage treatment it is always possible for a principal to fulfill the norm. Thus, one should expect to observe more norm violations in the EWT than in the IWT. If players care about multilateral reciprocity, efforts in the IWT should therefore be higher on average than under equal wages.

4 Results

In this section we present the results of the experiment and discuss possible explanations for the observed behavior. We first report the effort choices of the agents and then explore possible reasons for the treatment difference in effort levels. For this, we take a closer look at the monetary incentives induced by principals' wage setting under the different institutional rules. We then discuss non-monetary incentives arising from the interaction between wage-setting institutions and social preferences of the agents. Finally, we examine the consequences for efficiency.

4.1 Effort Choices

Figure 1 shows the development of average efforts over time. Two things are striking about the graph. First, efforts are considerably lower in the equal wage treatment. While agents in the IWT on average exert an effort of 8.21, their counterparts in the EWT average only 4.40 (Mann-Whitney test: $p < 0.01$). Second, efforts decrease over time under equal wages which is not the case when individual wages are paid (Wilcoxon test for periods 1–6 against 7–12: IWT, $p = 0.56$; EWT, $p < 0.01$). This means that the effort difference between the treatments becomes even larger during the experiment. The treatment difference is also present when individual matching groups are considered: the

(1989), Babcock et al. (1996), Fehr and Schmidt (1999) or Bolton and Ockenfels (2000).

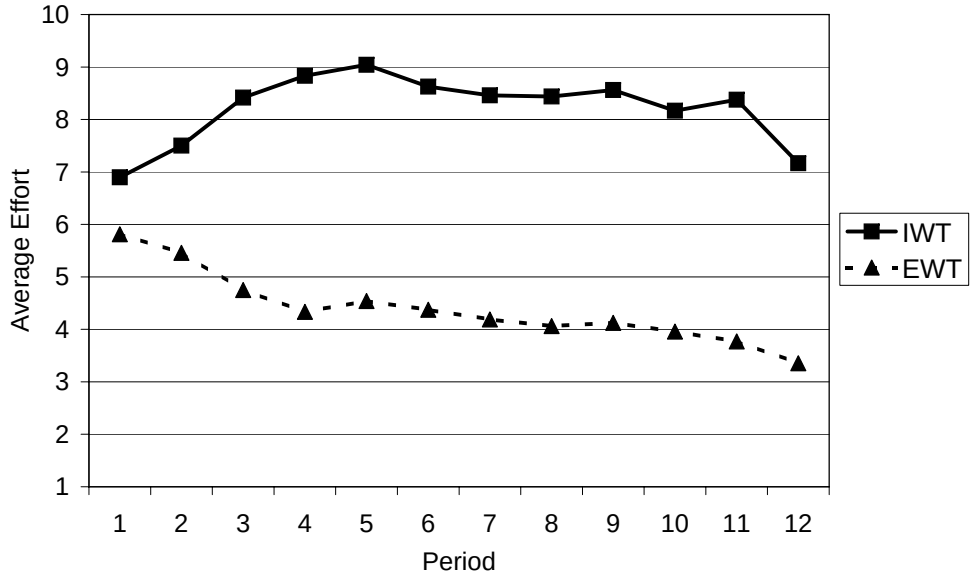

Figure 1: *Average effort per period. The effort is aggregated per period over all matching groups. Agents in the IWT exert significantly more effort.*

highest average effort of an EWT matching group (5.88) is still lower than the lowest average effort of an IWT matching group (7.47).

The difference in agents' behavior can also be seen in the frequencies of effort choices (Figure 2). In the individual wage treatment agents choose the maximal effort of 10 in 49% of the cases, 84% of the choices are higher than 6. Under equal wages, agents choose an effort higher than 6 in only 26% of cases. The effort decisions are more spread out in the EWT, the minimal effort of 1 being the modal choice with 24% of the choices.[8] The above observations can be summarized as follows.

Result 1: *Agents exert significantly more effort in the individual wage treatment than in the equal wage treatment.*

The above-minimal effort levels and the differences between treatments are not consistent with the predictions of the subgame perfect Nash equilibrium and the equilibrium of the model by Fehr and Schmidt (1999). In the following we will shed more light on the strong

[8]Recent evidence suggests that men and women respond differently to incentive schemes (e.g. Ivanova-Stenzel and Kübler (2005), Gneezy et al. (2003), Vandegrift and Yavas (2005)). However, we do not find significant differences in effort provision between our male and female participants (Mann-Whitney test based on matching group averages of male and female participants: $p = 0.23$ for the EWT, $p = 0.15$ for the IWT).

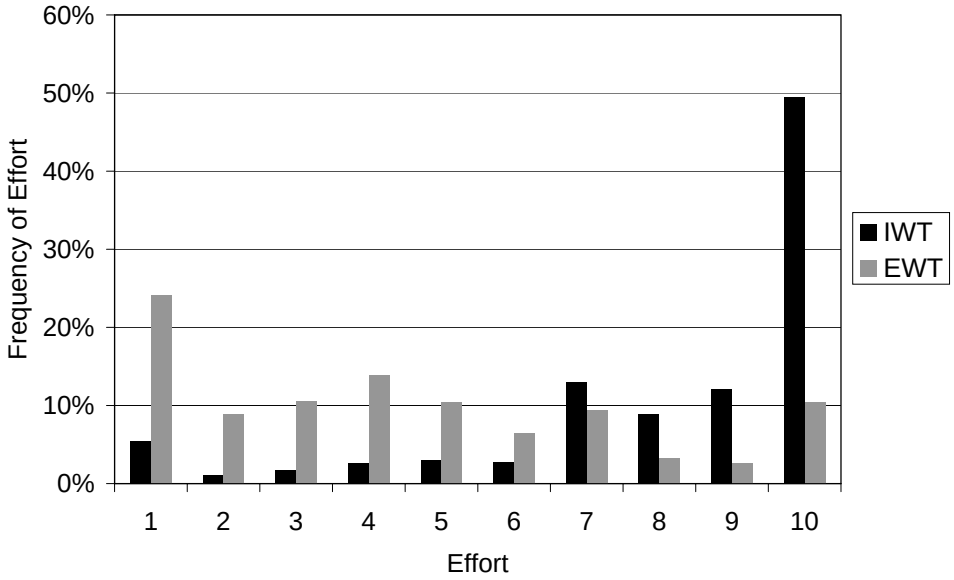

Figure 2: *Frequency of effort choices. The effort distribution in the IWT is more concentrated and shifted to the right in comparison to the EWT. The maximal effort of 10 is chosen in 49% of the cases in the IWT, while this is true for only 10% of choices in the EWT.*

difference in the agents' performance between the treatments. To this end, we first take a closer look at the monetary incentives that the agents face under the different institutional rules.

4.2 Monetary Incentives

For a first impression of the monetary incentives we plot the average wage per effort level in the two treatments in Figure 3.[9] The graph exhibits the upward sloping wage-effort relation of many gift-exchange experiments and shows that on average principals reciprocate higher effort levels with higher wages in both treatments.

The results of a regression analysis presented in Table 3 (column 1) confirm this impression. We regress the agent's wage w_i on his effort e_i, a treatment dummy IWT, an interaction term of the treatment dummy and his effort, and a constant. IWT is equal

[9]For both treatments we take the wage which was paid for each individual effort decision and calculate averages for a given effort level. Thus every wage decision of the principal enters twice in the equal wage treatment.

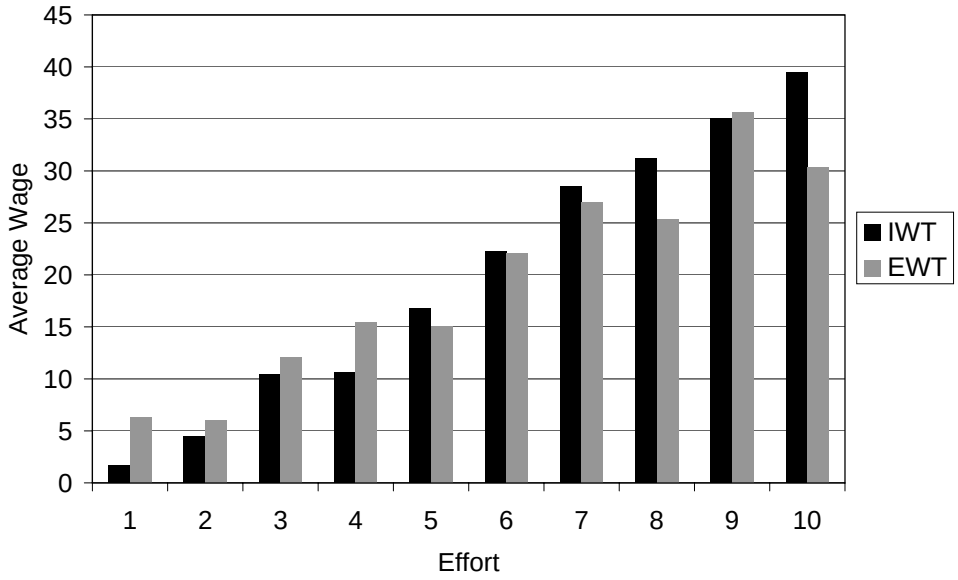

Figure 3: *Average wage for a given effort. Higher efforts lead to higher wages in both treatments.*

to 1 for the individual wage treatment and equal to 0 for the equal wage treatment.[10]

Both treatments show significantly positive wage-effort relations. The slope of the wage-effort relation is steeper in the individual wage treatment. An effort increase of 1 is rewarded with a wage increase of 4.2 in the IWT and with a wage increase of 3.1 in the EWT. These results hold if we control for the effort level of the co-worker (column 2) and an additional time trend (column 3).

Result 2: *We observe gift exchange in both treatments: a higher effort leads on average to a higher wage. The wage increase is stronger in the IWT.*

Note that the wage is only the gross profit for agents as they have to pay the cost of effort exertion (see Table 1). To determine the relation between (net) profits and efforts, we regress the agent's profit π_{A_i} on the same variables as in the wage regression presented above. We again control for clustering on matching groups. The results are reported in columns 4 to 6 of Table 3.

[10]We allow for dependent observations within matching groups and assume that only observations in different matching groups are independent. The reported robust standard errors are adjusted for this clustering. In order to estimate the influence of the co-worker's effort e_j we have to split the sample such that only one observation per firm is included in the analysis. The results reported here do not depend on which worker's effort is selected as "e_i".

Dep. Variable	w_i	w_i	w_i	π_{A_i}	π_{A_i}	π_{A_i}
e_i	3.117***	2.942***	2.934***	1.031*	0.854**	0.835**
	(0.506)	(0.332)	(0.309)	(0.535)	(0.348)	(0.331)
$IWT \times e_i$	1.129*	1.350***	1.417***	0.773	0.995*	1.082**
	(0.583)	(0.447)	(0.432)	(0.615)	(0.469)	(0.446)
cons	2.539	-9.065***	-8.720***	5.927**	-5.815***	-5.023***
	(2.570)	(1.583)	(1.320)	(2.614)	(1.523)	(1.286)
IWT	-5.535*	9.148**	12.668***	-3.744	11.004***	14.658***
	(3.072)	(3.255)	(3.589)	(3.235)	(3.274)	(3.479)
e_j		2.741***	2.731***		2.774***	2.750***
		(0.283)	(0.242)		(0.280)	(0.243)
$IWT \times e_j$		-3.156***	-3.167***		-3.178***	-3.180***
		(0.408)	(0.392)		(0.403)	(0.390)
t			-0.041			-0.093
			(0.194)			(0.176)
$IWT \times t$			-0.600**			-0.643**
			(0.253)			(0.227)
N. Obs.	576	576	576	576	576	576
R^2	0.409	0.500	0.505	0.100	0.238	0.251

Table 3: *Wage and profit regressions. Robust standard errors are given in parentheses. The dummy "IWT" is equal to 1 for the individual wage treatment. Significance at the 10%, 5% and 1% level is denoted by *, **, and ***, respectively.*

The profit-effort relation is significantly positive in both treatments, too. An effort increase of 1 leads to a profit increase of 1.8 in the IWT and to a profit increase of 1.0 in the EWT. Again, the slope is steeper for the IWT but the result is qualitatively the same for both treatments. These results still hold if one controls for the co-worker's effort e_j (column 5) and an additional time trend (column 6). The coefficient of the effort level is even more significant in these specifications.

Result 3: *A higher effort leads to a higher profit in both treatments. The increase is stronger for the individual wage treatment.*

At the beginning of this section, we reported the effort choices of the agents (see Figures

1 and 2). Agents in the individual wage treatment behave in line with the monetary incentives, most agents choose high efforts throughout the experiment. However, the actual effort choices in the equal wage treatment seem puzzling from a monetary point of view. Although the monetary incentive structure is quite similar to the IWT, actual effort choices are not. The behavior of most agents in the EWT stands in contrast to the monetary incentives we just presented.[11] They mostly choose low effort levels that lead to low expected profits.[12]

The analysis of wages and profits shows some differences in the incentive structure between treatments, but the discrepancy in agents' behavior can hardly be explained by these rather slight differences. We therefore turn to non-monetary reasons to explain the large and increasing effort gap between the treatments.

4.3 Non-Monetary Incentives

Focussing only on monetary incentives disregards the fact that many people also have social preferences. In the analysis of the principals' behavior, we have already seen that a considerable fraction of them adheres to the norm of reciprocity: almost all principals reciprocate a higher effort with a higher wage. As indicated in section 3, for agents with reciprocal preferences the two compensation schemes at hand might entail different non-monetary incentives, which could explain the difference in effort provision.

In the previous section we have already shown that a norm of bilateral reciprocity is fulfilled in both treatments: a higher effort is on average rewarded with a higher wage and a higher profit. If an agent focusses only on his relationship with the principal, both the monetary and the non-monetary incentives are similar across treatments. Thus the notion of bilateral reciprocity is not consistent with the treatment difference in efforts. Agents apparently are concerned about more than just bilateral reciprocity.

We suggested in section 3 that a multilateral version of reciprocity might be important for

[11] In the EWT, very few agents constantly provide high efforts. These agents receive the highest profits in this treatment.

[12] This result still holds when we control for the fact that the subjects only have access to the information of their former firms.

the agents. This concept implies that the social comparison to his co-worker influences an agent's perceptions of fairness within the firm. In particular, a reciprocal agent who works more than his co-worker also expects to receive a higher payoff.

In order to explore whether this form of reciprocity is important for agents, we compare the reactions of agents who experienced a violation of this norm to the reaction of those who did not. Table 4 shows agents' effort changes from period t to $t+1$ depending on whether multilateral reciprocity is fulfilled in period t. We distinguish three cases: no norm violation, advantageous norm violation (lower effort and higher profit), and disadvantageous norm violation (higher effort and lower profit).[13] The top panel of Table 4 shows data of the equal wage treatment. After experiencing a disadvantageous violation of reciprocity, the majority of agents decrease their effort and only few increase their effort in the following period. The opposite is true after an advantageous norm violation: agents tend to increase their effort and only few reduce it. When the norm is fulfilled, most agents keep their effort constant and slightly more agents increase their effort than decrease it. Apparently, agents dislike being exploited (disadvantageous norm violation) and dislike to feel guilty (advantageous norm violation). After a norm violation they change their effort provision in the direction that makes a violation less likely to occur in the next period. This is consistent with the predictions of multilateral reciprocity.[14]

Behavior in the individual wage treatment (bottom panel) is surprisingly similar. After a disadvantageous norm violation efforts tend to be decreased. When the norm is fulfilled, most agents keep their effort constant. The only difference between the treatments can be observed when agents experience an advantageous norm violation: in the IWT agents tend to decrease their effort. However, the reciprocity norm is violated in the IWT almost exclusively when the principal pays no or only a very small wage. This could be a reason

[13]To be more precise, agent i faces a disadvantageous norm violation if in his firm and a given period $e_i \geq e_j \wedge \pi_i < \pi_j$ or $e_i > e_j \wedge \pi_i = \pi_j$ holds. An advantageous norm violation for i occurs if $e_i \leq e_j \wedge \pi_i > \pi_j$ or $e_i < e_j \wedge \pi_i = \pi_j$. In all other cases the norm is fulfilled, i.e., if $e_i > e_j \wedge \pi_i > \pi_j$, or $e_i = e_j \wedge \pi_i = \pi_j$, or $e_i < e_j \wedge \pi_i < \pi_j$.

[14]Similar effects are observed by Thöni and Gächter (2005) in a related set-up. They allow agents to revise their effort decision after learning their co-workers' effort choice. In the revision stage, the majority decrease the effort difference to their co-worker, i.e., the agents with initially higher effort revise their decision downwards while the agents with lower effort revise it upwards.

	Effort Down	Effort Constant	Effort Up	No. Obs.
EWT				
No Violation	19.1	54.4	26.5	68
Adv. Violation	12.2	43.5	44.4	230
Disadv. Violation	52.6	33.9	13.5	230
Total	30.7	40.7	28.6	528
IWT				
No Violation	19.2	51.8	29.0	448
Adv. Violation	45	27.5	27.5	40
Disadv. Violation	35.0	57.5	7.5	40
Total	22.4	50.4	27.3	528

Table 4: *Frequency of effort reactions. Fractions (in %) of agents who decrease effort / keep effort constant / increase effort after experiencing no / an advantageous / a disadvantageous violation of multilateral reciprocity.*

for the additional downward bias.

If behavior is so similar between treatments, how can a preference for multilateral reciprocity cause the treatment effect? The last column of Table 4 shows how often the three cases occur in the two treatments. In the EWT, the norm is violated for 87% of firms. In the IWT, it is violated only in 15% of the cases. Thus, even if the behavior in a given situation is similar, agents in the EWT are far more often exposed to norm violations than agents in the IWT. This illustrates nicely how the same social preference can have very different implications under different institutions.

Result 4: *Agents mostly react to violations of multilateral reciprocity by reducing their effort after a disadvantageous norm violation and by increasing it after an advantageous norm violation. Reactions are similar in both treatments, however, the norm is far more often violated in the equal wage treatment.*

So far we have seen that agents' reactions are in line with the hypothesized behavior of a reciprocal agent who does not only focus on bilateral, vertical reciprocity but evaluates fairness in the firm also with respect to the horizontal, co-worker dimension. Note that

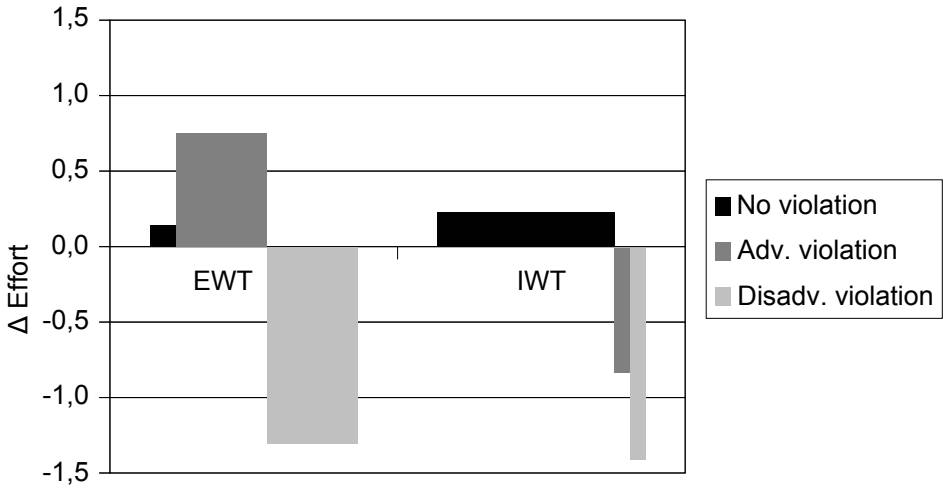

Figure 4: *Magnitude of effort reactions. "Δ Effort" is the average change in effort from period t to period $t+1$ given that the agent experienced no norm violation, an advantageous violation or a disadvantageous norm violation in period t. The width of the bars corresponds to the number of observations.*

for every agent who experiences a disadvantageous norm violation, there is one agent (his co-worker) who faces an advantageous norm violation. Only if both agents' reactions to these violations are not equally strong, a preference for multilateral reciprocity can potentially explain the downward trend in effort provision in the EWT.

We therefore analyze the relative strength of agents' reactions to norm violations. Figure 4 shows the average change in effort provision from period t to $t+1$ after an agent experienced no norm violation, a disadvantageous or an advantageous norm violation in period t. The width of the bars corresponds to the number of observations in the respective category (cf. Table 4). The reaction of disadvantaged subjects in the EWT is almost twice as strong as the reaction of advantaged subjects. The difference is statistically significant (Wilcoxon signed rank test: $p < 0.01$). This indicates that agents suffer more from a disadvantageous norm violation than from an advantageous violation. As already observed above, in the IWT both groups of agents that experience a norm violation decrease their effort.[15]

Note again that in both treatments agents tend to keep their effort constant or even slightly

[15]However, a regression analysis (not reported here) shows that the coefficient of effort change is not significant for agents in the IWT who benefitted from a norm violation.

increase as long as the reciprocity norm is not violated in their firm. This suggests that fulfillment of multilateral reciprocity has a motivating effect. We conclude that the overall negative effort trend and ultimately the low effort levels in the EWT are actually caused by the asymmetry in reactions of reciprocal players and especially the strong negative reaction to a disadvantageous norm violation.

Result 5: *Agents' reactions to a violation of multilateral reciprocity are asymmetric: the negative reaction of the disadvantaged agents is stronger than the positive reaction of the advantaged agents. This asymmetry in agents' reactions results in an overall negative time trend in efforts for the EWT and in the strong difference in efforts between treatments.*

Note that the agents do not react to payoff inequalities per se. The absolute differences between agents' payoffs are not significantly different between treatments (IWT: 6.47, EWT: 7.14, Mann-Whitney test: $p = 0.29$). Albeit not different in absolute terms, the sign differs. In the EWT the hard-working agent ends up with the lower payoff, while in the IWT he gets the higher payoff—which is exactly what multilateral reciprocity is about.

In order to further analyze whether the presence of reciprocal agents drives the results in our experiment, we take our previous findings on agents' reactions and simulate agents' behavior with a simple "reciprocity learning" rule. In this simulation, all agents are assumed to derive utility from money, but to also suffer from violations of multilateral reciprocity. When deciding about their effort in a given period, the simulated agents therefore compare their effort and profit with the effort and profit of their co-worker in the previous period. According to the comparison along these two dimensions, four reactions can be distinguished for the simulated agents. (i) For an agent who had a higher effort and a higher profit, multilateral reciprocity is fulfilled and the pecuniary comparison is also advantageous for him, so he keeps his effort constant. (ii) For an agent who exerted a lower effort and got a lower profit, the norm of reciprocity is satisfied but profit maximization is not, thus he subsequently imitates the other agent by half, i.e. he chooses an effort $(e_{i,t} + e_{j,t})/2$. (iii) An agent with higher effort and lower profit feels exploited as he suffers from a disadvantageous norm violation. Thus he imitates the other agent fully. (iv) Finally, for an agent with lower effort and higher profit the norm violation is advantageous, thus the resulting disutility is smaller than in case (iii). He chooses an

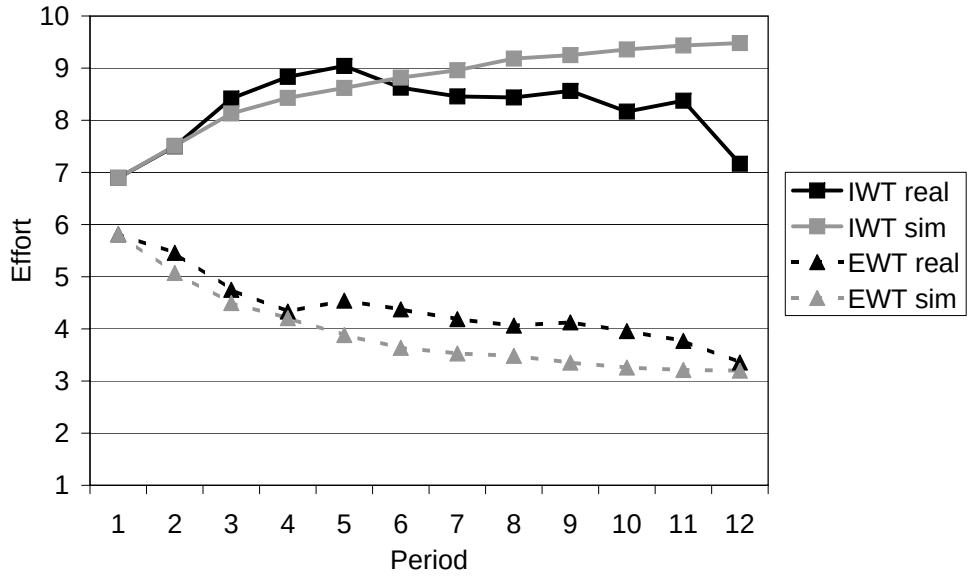

Figure 5: *Simulation of reciprocal agents. The simulation yields effort dynamics similar to those in the real experimental data.*

effort $(e_{i,t}+e_{j,t})/2$. The reactions in cases (i) to (iv) are in line with the aggregate evidence presented in Table 4 and Figure 4.[16]

We use actual effort data from the experiment only for the first period of the simulation. The subsequent effort decisions are based on the simulated profits and simulated efforts of the previous period. The simulated principals pay the average wage for a given effort (IWT) or the average wage sum for a given effort sum (EWT) calculated from the experimental data. Profits are then calculated as wage minus cost of effort exertion. We use the same matching protocol as in the experiment.

Figure 5 shows how effort choices evolve over time in the experimental data (as in Figure 1) and in the simulation. The simulation traces the real data very well and is able to reproduce the large effort difference between the treatments. In the individual wage simulation, efforts increase like the real efforts although the slight downward trend in the second half of the experiment cannot be reproduced. Efforts in the equal wage simulation constantly decrease down to an effort level slightly above 3 in the final period. This pattern is impressively similar to the dynamics in the real data.

Result 6: *A simple simulation based on agents who have preferences for money and*

[16]The qualitative results of the simulation is robust to the exact specification of agents' behavior.

multilateral reciprocity is in line with the efforts observed in the experiment.

Note that the focal agent is different between the simulated treatments: in the equal wage simulation the norm of multilateral reciprocity is not fulfilled when agents choose different effort levels. In these cases, the agent with the higher effort will fully imitate the agent with the lower effort while this agent will increase his effort level only to the average effort of the last period. Therefore in the EWT simulation, the average effort converges to the lowest first period effort as agents are subsequently matched together and imitate the lower effort more strongly than the higher effort. In contrast, in the IWT the norm of reciprocity is mostly fulfilled and thus the agent with the higher effort is imitated more often. Thus, the average effort converges to the highest first period effort.

To see whether this can also be found in the experimental data we classify agents according to their effort decision in the first period. We define the agents with the highest effort in the first period in their matching group as "high-effort providers" and the agents with the lowest effort as "low-effort providers".[17] This type definition is chosen because when agents decide on their effort in the first period, they do not have any information about the behavior of other subjects and all learning and coordination processes occur after this effort choice. Thus if some of the subjects are intrinsically inclined to exert high efforts they should show up in the group of high-effort providers. In contrast, if some of the subjects are intrinsically inclined to exert low efforts they should show up in the group of low-effort providers.

In Figure 6 we follow the high-effort and low-effort providers in both treatments and show their effort decisions over time. In the first period, the groups of high-effort providers and the groups of low-effort providers are close together across treatments.[18] This changes

[17] As there are six agents per matching group, each of these groups corresponds to 16.7% of agents. If more than one agent chooses the highest or lowest effort in the first period, the subsequent effort decisions of these agents are averaged. We carried out the same analysis ignoring matching groups, ordering all effort decisions of the first period and defining the top 20% of all agents as "high-effort providers" and the bottom 20% as "low effort providers". This analysis yields the same results.

[18] In the first period, effort levels are not significantly different between treatments for high-effort providers (Mann-Whitney test: $p = 0.14$) while they are close together but different for the low-effort providers (Mann-Whitney test: $p = 0.03$). Within treatments, the high-effort and low-effort providers choose statistically different effort levels (Wilcoxon signed rank test: $p = 0.01$ (IWT), $p = 0.01$ (EWT)).

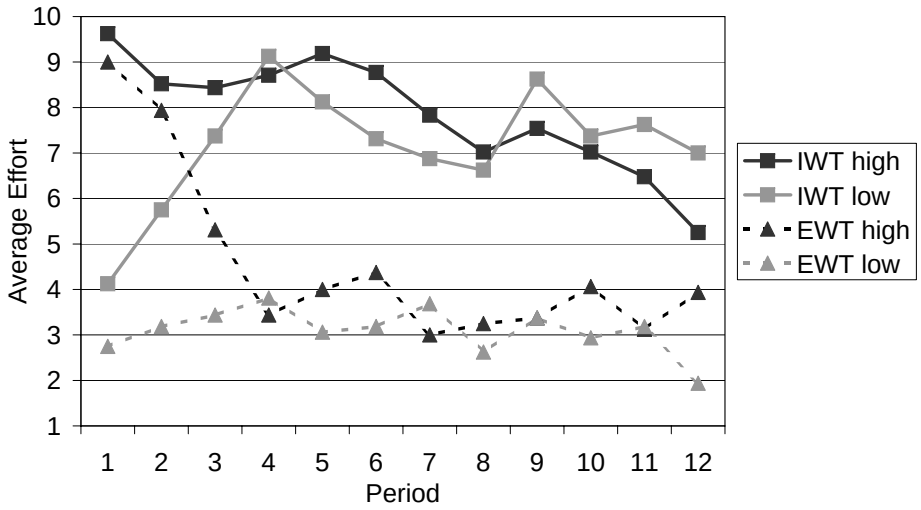

Figure 6: *Effort decisions of high-effort and low-effort providers. In each matching group, the agent with the highest (lowest) effort in the first period is defined as the high (low)-effort provider.*

completely over the course of the 12 periods. In the individual wage treatment, high-effort providers continue to provide high effort levels but the low-effort providers increase their efforts dramatically up to the level of the high-effort providers and even higher in the last periods. In the equal wage treatment, the dynamics are reversed. Here, the low-effort providers keep their effort provision constant and the high-effort providers reduce their efforts to the level of the low-effort providers.[19]

These dynamics underline the importance of the different non-monetary incentives implied by the two wage setting institutions. Wage equality essentially entails the wrong reciprocity signals and therefore discourages agents who are in principle willing to exert high levels of effort. On the contrary, under individual wages where intra-firm reciprocity is intact, good performance spreads. Put simply, one could say that under individual wages, the "good" agents pull the "bad" agents up while under equal wages, the "bad" pull the "good" down.

Result 7: *The focal agent is different between treatments: in the IWT the initially low-effort providers align with the high-effort providers over time. In the EWT the initially*

[19]In the last six periods, effort levels are not different within treatments (Wilcoxon signed rank test: $p = 0.67$ (IWT), $p = 0.78$ (EWT)) while they differ between treatments (Mann-Whitney test: $p < 0.01$ (high-effort providers), $p < 0.01$ (low-effort providers)).

high-effort providers align with the low-effort providers over time.

We have seen so far that norm violations caused by the equal wage scheme can explain the low efforts in the EWT. Interestingly, not only the agents but also the principals seem to suffer from the norm violations in the EWT. When the two agents in their firm provide different efforts reciprocal principals face a moral dilemma: they would like to reward high efforts and punish low efforts. However, they can only pay one wage. Several findings suggest that principals suffer from this moral dilemma.

First, in the post-experimental questionnaire, principals of the EWT are asked how important it is for them to (i) reward high efforts and (ii) punish low efforts on a scale from 1 to 6. The results show that rewarding high efforts and punishing low efforts are equally important for the principals (Mann-Whitney test: $p = 0.20$). However, the two cannot be reconciled in the equal wage treatment.

Second, as another item in the questionnaire, the principals are presented three hypothetical game situations that include effort choices, wage choices, and the resulting payoffs for all players. Subsequently they are asked whether they consider the resulting allocation as just. In the first situation, all three player have the same payoff, in the second situation the agents' payoffs are again the same but the principal's payoff is twice as high. Finally, each principal is presented a situation that reflects his own average behavior in the experiment.[20] The principals do not know that they are facing their own past decisions when answering this question. The answers to the first two situations are not significantly different between treatments (Mann-Whitney test: $p = 0.32$ and $p = 0.32$), suggesting that the principals across treatments have similar fairness preferences. However, 60% of the principals in the IWT consider their own decisions fair while only 35% of the principals in the EWT are of this opinion (Mann-Whitney test on matching group shares: $p = 0.03$).

Finally, an estimation of the determinants of wage setting on firm level confirms the questionnaire data. We regressed the wage sum on effort sum, absolute effort difference

[20]The third situation is constructed as follows: We calculate the average effort of the higher-effort and of the lower-effort providers the principals actually faced during the experiment. We then look at the wages the principals paid to the two groups and again calculate averages. Finally, we calculate a hypothetical payoff for all three "average" players by considering the costs of the average efforts.

Dep. Variable	$(w_1 + w_2)$
cons	-15.380***
	(2.392)
IWT	14.158*
	(7.297)
$(e_1 + e_2)$	6.137***
	(0.619)
$IWT \times (e_1 + e_2)$	-2.206***
	(0.741)
$\|e_1 - e_2\|$	-2.152**
	(0.945)
$IWT \times \|e_1 - e_2\|$	2.450**
	(1.068)
N. Obs.	576
R^2	0.478

Table 5: *Wage sum regression. Robust standard errors are given in parentheses. The dummy "IWT" is equal to 1 for the individual wage treatment. Significance at the 10%, 5% and 1% level is denoted by *, **, and ***, respectively.*

and interaction terms (Table 5). Principals reward a higher effort sum with a higher wage sum in both treatments. However, the principals also reward homogenous working groups in the EWT. This is not the case in the IWT where the principal can individually target his wage payments. This suggests that the principals in the EWT reward their agents for not having to solve the moral dilemma.

Result 8: *Principals experience the equal-wage institution as less fair because it confronts them with a moral dilemma if agents provide different efforts.*

4.4 Efficiency

In this section we analyze the consequences of the differences in incentives and effort provision for efficiency. Since higher efforts increase production and since the marginal product of effort always exceeds its marginal cost, the differences in effort provision di-

rectly translate into differences in efficiency. As the average effort is much higher in the individual wage treatment, efficiency is higher in this treatment as well.

Both, the agents and the principals benefit from the increase in efficiency, so the individual wage treatment even Pareto dominates the equal wage treatment. The average period profit of a principal is 56 in the EWT and 100 in IWT (Mann-Whitney test: $p < 0.01$), while agents earn an average period profit of 10 under equal wages vs. 17 under individual wages (Mann-Whitney test: $p < 0.01$). As the players suffer more frequently from norm violations in the EWT, if anything, the treatment difference in utility should be even larger.

The high effort level in the individual wage treatment is worth noting in itself: in all but 3 periods the average effort level is higher than 8 (with a maximum possible effort of 10). This is an extremely high degree of efficiency compared to gift-exchange experiments with similar parameters in which the wage is paid before effort is provided. As an example, Fehr et al. (2004) use the same game parametrization but with an upfront wage payment and observe an average effort of approximately 2. The efforts in our IWT even reach the levels of repeated game experiments on gift exchange (cf. Falk et al. 1999). The degree of efficiency apparently depends strongly on which party is moving first. This corroborates the results of Fehr et al. (2004) who find that adding a bonus payment stage to a usual gift-exchange experiment is efficiency increasing. They claim that in general "the person who loses less from trusting the other person should trust first". Gneezy (2006) makes a similar point, however, he conducts a real effort experiment and thus can control only imperfectly for the cost of effort exertion. These observations highlight the importance of the timing of actions in reciprocal relationships. They also underline the efficiency-enhancing effect of determining a leader as a residual claimant in the production process and giving him the right to allocate rewards (cf. Potters et al. 2004).

Result 9: *The individual wage treatment Pareto dominates the equal wage treatment. Compared to other gift-exchange experiments, the IWT exhibits a remarkably high degree of efficiency.*

5 Conclusions

We analyzed how effective equal wages are in an environment where explicit contract enforcement is absent. In our experiment, subjects interact in firms consisting of two employees and one employer who sets individual wages in one treatment and pays equal wages in the other. The use of individual wages elicits substantially higher efforts and efficiency although the monetary incentives for the workers are similar under both wage schemes. In particular, exerting high effort pays off in both settings. The strong treatment difference is driven by subjects' reciprocal preferences and the fact that a multilateral reciprocity norm is violated far more frequently in the equal wage treatment. Therefore, workers who are initially willing to provide high effort get discouraged under wage equality and lower their effort. On the other hand, the individual wage treatment shows that good performance of agents spreads as long as the social norm within the firm is fulfilled.

Our results have a number of important implications, both for the advancement of existing "behavioral" theories and for the design of wage schemes in practice. First of all, it is doubtful whether wage equality can be reconciled with the use of reciprocity to enforce implicit contracts. Our results rather suggest that the possibility to individually sanction bad performance and reward good performance is a crucial prerequisite to make reciprocity the powerful enforcement device it has proven to be in many bilateral interactions.

Second, our findings offer general insights into the nature of reciprocal preferences in multilateral settings. They indicate that reciprocal agents do not only care about bilateral reciprocity, i.e., their individual relationship to the principal. In addition, they also want reciprocity to be fulfilled in a horizontal dimension, i.e., an agent who performs better than his co-worker expects also to receive higher rewards. This implies that agents actually care about their co-workers. Moreover, the results demonstrate that the evaluation of horizontal fairness goes beyond a one-dimensional wage comparison, but entails a more complex comparison of performance and rewards.

Third, while it is well-known that equal wages can distort monetary incentives, we find that they are also perceived as less fair and thereby decrease efficiency, even if the monetary incentives are not affected. It is thus oversimplifying to argue that equal wages

lead to less envy and therefore higher work morale, as is frequently done in the political discussion. Our results show that social preferences must not be neglected when incentive schemes are designed and that it is additionally necessary to have a sound understanding of the nature of these preferences.

In practice, the discretion to individually reciprocate good performance of subordinates does not have to be in monetary terms. Non-monetary benefits like extra vacation or awards also represent potentially useful devices to motivate workers in this context. These instruments become especially important when it is not possible to wage discriminate on a given hierarchical level, e.g., because the firm's internal pay structure, agreements with a union or the political system dictate wage equality.

Our results should not be interpreted as arguments against wage compression in general but against extreme forms of compression. It is also clear that our design abstracts from some potential benefits of equal wages. Examples are reduced implementation or monitoring costs and high transparency due to their simplicity as well as increased incentives for peer-monitoring. Moreover, the information structure and organizational design of the firm will also influence the impact of the workforce's social preferences. An advantage of our design is that it provides a simple and parsimonious framework that can be successively enriched to study these questions in future research.

Finally, our findings illustrate that the impact of social preferences can strongly differ under different institutional rules. Some of the most interesting aspects of the interaction between social preferences and institutions can only be studied in a multilateral set-up with a hierarchical structure where social preferences can have a horizontal and a vertical dimension. Equality can be unfair. We believe that further research along these lines is an important and fruitful enterprise.

References

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

Babcock, Linda, Xianghong Wang and George Loewenstein (1996): "Choosing the Wrong Pond: Social Comparisons in Negotiations That Reflect a Self-Serving Bias." *Quarterly Journal of Economics*, 111:1–21.

Baker, George P., Michael C. Jensen and Kevin J. Murphy (1988): "Compensation and Incentives: Practice vs. Theory." *Journal of Finance*, 43:593–616.

Bartling, Björn and Ferdinand von Siemens (2004): "Efficiency in Team Production with Inequity Averse Agents." *Working Paper, University of Munich.*

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

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

Bowles, Samuel and Yongjin Park (2005): "Emulation, Inequality, and Work Hours: Was Thorsten Veblen Right?" *The Economic Journal*, 115:F397–F412.

Charness, Gary and Peter Kuhn (2005): "Pay Inequality, Pay Secrecy, and Effort: Theory and Evidence." *NBER Working Paper*, 11786.

Clark, Andrew, David Masclet and Marie-Claire Villeval (2006): "Effort and Comparison Income: Survey and Experimental Evidence." GATE Working Paper 06-01.

Clark, Andrew E. and Andrew J. Oswald (1996): "Satisfaction and Comparison Income." *Journal of Public Economics*, 61:359–381.

Demougin, Dominique and Claude Fluet (2003): "Group vs. individual performance pay when workers are envious." *CIRPÉE Working Paper*.

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

Dur, Robert and Amihai Glazer (2004): "Optimal Incentive Contracts For a Worker Who Envies His Boss." *CESifo Working Paper*, 1282.

Easterlin, Richard A. (2001): "Income and Happiness: Towards a Unified Theory." *The Economic Journal*, 111:465–484.

Erev, Ido, Gary Bornstein and Rachel Galili (1993): "Constructive Intergroup Competition as a Solution to the Free Rider Problem: A Field Experiment." *Journal of Experimental Social Psychology*, 29:463–478.

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

Falk, Armin, Simon Gächter and Judit Kovács (1999): "Intrinsic motivation and extrinsic incentives in a repeated game with incomplete contracts." *Journal of Economic Psychology*, 20:251–284.

Fehr, Ernst and Armin Falk (2002): "Psychological Foundations of Incentives." *European Economic Review*, 46:687–724.

Fehr, Ernst, Simon Gächter and Georg Kirchsteiger (1997): "Reciprocity as a Contract Enforcement Device: Experimental Evidence." *Econometrica*, 65:833–860.

Fehr, Ernst, Alexander Klein and Klaus M. Schmidt (2004): "Contracts, Fairness, and Incentives." *Working Paper, University of Munich*.

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

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

Frey, Bruno S. and Alois Stutzer (2002): "What can Economists Learn from Happiness Research?" *Journal of Economic Literature*, 40:402–435.

Gneezy, Uri (2006): "Bonuses versus predetermined wages: A study of incentives and reciprocity." *Working Paper, University of Chicago*.

Gneezy, Uri, Muriel Niederle and Aldo Rustichini (2003): "Performance in Competitive Environments: Gender Differences." *Quarterly Journal of Economics*, 118:1049–1074.

Güth, Werner, Manfred Königstein, Judit Kovács and Enikö Zala-Mezo (2001): "Fairness Within Firms: The Case of One Principle and Multiple Agents." *Schmalenbach Business Review*, 53:82–101.

Hannan, R. Lynn, John H. Kagel and Donal 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*, 20:923–951.

Holmström, Bengt (1982): "Moral Hazard In Teams." *Bell Journal Of Economics*, 13:324–340.

Ivanova-Stenzel, Radosveta and Dorothea Kübler (2005): "Courtesy and Idleness: Gender Differences in Team Work and Team Competition." *mimeo*.

Kandel, Eugene and Edward P. Lazear (1992): "Peer Pressure and Partnerships." *Journal of Political Economy*, 100:801–817.

Layard, Richard (2002): "Rethinking Public Economics: The Implications of Rivalry and Habit." *mimeo*.

Loewenstein, George, Leigh Thompson and Max Bazerman (1989): "Social Utility and Decision Making in Interpersonal Contexts." *Journal of Personality and Social Psychology*, 57:426–441.

Maximiano, Sandra, Randolph Sloof and Joep Sonnemans (2004): "Gift Exchange in a Multi-Worker Firm." *Tinbergen Institute Discussion Paper, Amsterdam*.

Meidinger, Claude, Jean-Louis Rullière and Marie-Claire Villeval (2003): "Does Team-Based Compensation Give Rise to Problems when Agents Vary in their Ability?" *Experimental Economics*, 6:253–272.

Neumark, David and Andrew Postlewaite (1998): "Relative Income Concerns and the Rise in Married Women's Employment." *Journal of Public Economics*, 70:157–183.

Potters, Jan, Martin Sefton and Eline van der Heijden (2004): "Hierarchy and Opportunism in Teams." *mimeo*.

Rey-Biel, Pedro (2004): "Inequity Aversion and Team Incentives." *Working Paper, University College London*.

Rossi, Alessandro and Massimo Warglien (2001): "An Experimental Investigation of Fairness and Reciprocal Behavior in a Triangular Principal-Multiagent Relationship." *Working Paper, University of Trento*.

Solnick, Sara J. and David Hemenway (1998): "Is More Always Better? A Survey on Positional Concerns." *Journal of Economic Behavior and Organization*, 37:373–383.

Thöni, Christian and Simon Gächter (2005): "Social Interaction in the Workplace." *mimeo*.

Torgler, Benno, Sascha L. Schmidt and Bruno S. Frey (2006): "Relative Income Position and Performance: an Empirical Panel Analysis." *IEW Working Paper*, 268.

Vandegrift, Donald and Abdullah Yavas (2005): "An Experimental Test of Behavior under Team Production." *mimeo*.

6 Appendix

In this appendix we analyze the behavior of the players if they maximize an inequality aversion utility function of the following form (see Fehr and Schmidt 1999):

$$U_i(\pi) = \pi_i - \alpha_i \frac{1}{n-1} \sum_{j \neq i} \max[\pi_j - \pi_i, 0] - \beta_i \frac{1}{n-1} \sum_{j \neq i} \max[\pi_i - \pi_j, 0]$$

Here π is the vector of monetary payoffs of the players, n is the number of the players in the reference group, α_i represents the "envy" of player i if his monetary payoff is smaller than that of player j, and $\beta_i \in [0, 1)$, $\beta_i \leq \alpha_i$ is i's feeling of "guilt" if he has a larger monetary payoff than j.

We assume that the inequality averse players compare their monetary payoff to the (expected) monetary payoff of both other players in their firm. We first assume that the parametrization given by Fehr and Schmidt (1999) holds for our experimental subjects (section 6.1) and subsequently calculate the equilibria with a slightly different distribution of the fairness parameters α and β (section 6.2).

6.1 Parametrization by Fehr and Schmidt

As we will see below, the crucial part of the parameter distribution in this section is the distribution of the guilt aversion parameter β. Using data from various experiments, Fehr and Schmidt derive the following distribution for β: 30% of the population show no guilt aversion at all ($\beta = 0$), 30% are slightly guilt averse ($\beta = 0.25$), and 40% show stronger guilt aversion ($\beta = 0.6$). The distribution of the envy aversion parameter α is: 30% with $\alpha = 0$, 30% $\alpha = 0.5$, 30% $\alpha = 1$, 10% $\alpha = 4$. As Fehr and Schmidt (1999) we assume that α and β are perfectly correlated. This last assumption plays only a role in section 6.2.2.

With this distribution we are able to solve the stage game of our experiment by backward induction. It is clear that for our game, a selfish principal ($\beta = 0$) will never pay a wage on the last stage as this would only reduce his utility. For the wage decision of the inequality averse principals note first that only the feeling of guilt may cause principals to pay a positive wage. After the agents have chosen their effort and before wages are paid the monetary payoff of the principal is always greater or equal than the monetary payoff of

each agent with our choice of parameters.[21] The decision problem of an inequality averse principal is therefore as follows:

$$\begin{aligned}\max_{w_1,w_2} u_P(w_1,w_2) &= 10(e_1+e_2)-w_1-w_2 \\ &\quad -\frac{\beta}{2}\max\{10(e_1+e_2)-w_1-w_2-(w_1-c(e_1)),0\} \\ &\quad -\frac{\beta}{2}\max\{10(e_1+e_2)-w_1-w_2-(w_2-c(e_2)),0\}\end{aligned}$$

s.t.

$$0 \leq w_i \leq 100$$

where e_1, e_2 are the efforts provided by agent 1 and 2 respectively and w_1, w_2 are the wages the principal pays to agent 1 and 2. Recall that the only difference between treatments is the following additional constraint that a principal in the group wage treatment faces:

$$w = w_1 = w_2$$

Rearranging terms, the objective function of the principal looks as follows:

$$\max_{w_1,w_2} 10(e_1+e_2)[1-\beta] - \frac{1}{2}\beta[c(e_1)+c(e_2)] + w_1(\frac{3}{2}\beta-1) + w_2(\frac{3}{2}\beta-1)$$

Under the assumptions on β given above this immediately implies that even the most inequality averse principals (the 40% with $\beta = 0.6$) will never pay a positive wage. An inequality reducing increase of the wage for an arbitrary agent by 1 unit costs 1 but gives the principal only an extra utility of $\frac{3}{2}\beta < 1$. Thus, in our setup stronger guilt aversion would be necessary to make positive wage choices optimal ($\beta \geq \frac{2}{3}$). This is true for both treatments. Thus, under this distribution of β all principals pay no wage.

Given this, the analysis of workers' behavior is trivial. As they expect to receive a wage of zero with certainty, both selfish and inequality averse workers choose the minimal effort $e = 1$ on the first stage in both treatments. The Fehr-Schmidt model thus yields the same predictions as the subgame perfect Nash equilibrium assuming rational and selfish players.

[21] The principal has strictly positive payoffs as he gets all goods produced, i.e. $10(e_1+e_2)$, whereas the agents have weakly negative payoffs as they have to bear the cost of production.

6.2 Alternative Parametrization

We now assume that 40% of the players are slightly more guilt averse than in the distribution of preferences given above, i.e. $\beta > \frac{2}{3}$ instead of $\beta = 0.6$ for 40% of the subjects. As we have seen before, this means that positive wage payments can now be optimal for some of the principals. We leave the remaining distribution of parameters unchanged.

6.2.1 Individual Wage Treatment

We first consider the individual wage treatment. From the objective function above it is clear that a principal with $\beta > \frac{2}{3}$ wants to equalize the monetary payoffs of all members of his firm. If the "ex interim" difference (before paying wages) between his payoff and the agents' payoff is large enough, it is optimal for him to pay the following wages to the agents.

$$w_i^* = \frac{10}{3}e_i + \frac{10}{3}e_j + \frac{2}{3}c(e_i) - \frac{1}{3}c(e_j)$$

However, in our stage game, the monetary ex interim payoff of the principal is always so high that he can equalize all monetary payoffs by paying appropriate wages to the agents. For the following analysis we can thus assume that the principals with $\beta > \frac{2}{3}$ pay the wages given above in the IWT.

On the first stage, a completely selfish agent (i.e., an agent with $\alpha = \beta = 0$) maximizes his expected monetary payoff. Because he can only hope to receive a wage if he is matched with one of the principals with high enough β (which occurs with probability 0.4), he faces the following objective function:

$$\begin{aligned} \max_{e_i} \quad & -c(e_i) + \frac{4}{10}\left[\frac{10}{3}e_i + \frac{10}{3}e_j + \frac{2}{3}c(e_i) - \frac{1}{3}c(e_j)\right] \\ = \quad & \frac{4}{3}e_i - \frac{11}{15}c(e_i) + \frac{4}{3}e_j - \frac{2}{15}c(e_j) \end{aligned}$$

Taking into account the cost function described in Table 1 it is optimal for a selfish agent to choose effort $e_i^* = 3$.

An agent with positive fairness parameters takes the distribution of principals and co-workers he may be matched with as given and maximizes his expected Fehr-Schmidt utility. If he takes the effort difference to his co-worker as given, he fears to meet a

principal with $\beta < \frac{2}{3}$ because in this situation no wages are paid and payoffs may therefore differ. We will show in the following that it is thus optimal for a fair agent to choose the minimal effort. We proceed as follows: We first consider the two cases where the inequity averse agent provides (weakly) higher / lower effort than his co-worker separately. We then show that his optimal strategy does not differ for these cases.

Under the assumption that the other agent, j, chooses a higher or equal effort than himself, the objective function of a fair agent i looks as follows:

$$\begin{aligned}\max_{e_i} \quad & -c(e_i) + \frac{4}{10}\left[\frac{2}{3}c(e_i) - \frac{1}{3}c(e_j) + \frac{10}{3}e_i + \frac{10}{3}e_j\right] \\ & -\frac{3}{10}\alpha[10e_i + 10e_j + c(e_i)] - \frac{3}{10}\beta\left[-c(e_i) + c(e_j)\right] \\ & = e_i\left[\frac{4}{3} - 3\alpha\right] + c(e_i)\left[-\frac{11}{15} - \frac{3}{10}\alpha + \frac{3}{10}\beta\right] + e_j\left[\frac{4}{3} - 3\alpha\right] + c(e_j)\left[-\frac{2}{15} - \frac{3}{10}\beta\right]\end{aligned}$$

As $\alpha \geq 0.5$ and $\alpha \geq \beta$ for the agents who are not completely selfish, the optimal choice is the exertion of minimal effort, $e_i^* = 1$.

If we assume that the co-worker j chooses a lower effort than i, i experiences more envy and less regret, but the optimal choice remains the same:

$$\begin{aligned}\max_{e_i} \quad & -c(e_i) + \frac{4}{10}\left[\frac{10}{3}e_i + \frac{10}{3}e_j + \frac{2}{3}c(e_i) - \frac{1}{3}c(e_j)\right] \\ & -\frac{3}{10}\alpha[10e_i + 10e_j + c(e_i)] - \frac{3}{10}\alpha\left[-c(e_j) + c(e_i)\right] \\ & = e_i\left[\frac{4}{3} - 3\alpha\right] + c(e_i)\left[-\frac{11}{15} - \frac{3}{5}\alpha\right] + e_j\left[\frac{4}{3} - 3\alpha\right] + c(e_j)\left[-\frac{2}{15} + \frac{3}{10}\beta\right]\end{aligned}$$

Minimal effort, $e_i^* = 1$, is again optimal for the fair player i.

We conclude that there is a unique equilibrium for the IWT, in which completely selfish agents choose $e = 3$, inequality averse agents choose $e = 1$, and only principals with $\beta > \frac{2}{3}$ pay positive wages such that ex post the monetary payoffs of all players are equal, i.e. $w_i = \frac{2}{3}c(e_i) - \frac{1}{3}c(e_j) + \frac{10}{3}e_i + \frac{10}{3}e_j$. The resulting average effort in the IWT is then 1.6 whereas the average wage is 8.69.

6.2.2 Equal Wage Treatment

In the equal wage treatment, the behavior of a highly guilt averse principal may differ compared to the individual wage treatment. If the two agents in his firm choose different

effort levels, he can no longer equalize all payoffs. As he suffers more from disadvantageous inequality than from advantageous inequality ($\alpha \geq \beta$), he chooses the wage in order to match his monetary payoff with the "lazier" agent (the one with the lower effort level and thus higher payoff), thereby ensuring himself not to feel envious against that agent. The optimal wage payment w^* is therefore

$$w^* = \frac{10}{3}e_1 + \frac{10}{3}e_2 + \frac{1}{3}c(\min[e_1, e_2])$$

For the behavior of an agent it is again convenient to distinguish the case where he provides weakly lower effort than his co-worker from the case where his effort is higher. It will again be shown that all types of agents have a dominant strategy, so that finally this case distinction does not matter. In the case where the agent expects to exert lower or equal effort than his co-worker, his objective function looks as follows:

$$\begin{aligned}\max_{e_i} \quad & -c(e_i) + \frac{4}{10}\left[\frac{10}{3}e_i + \frac{10}{3}e_j + \frac{1}{3}c(e_i)\right] - \frac{3}{10}\alpha\left[10e_i + 10e_j + c(e_i)\right] \\ & -\frac{3}{10}\beta\left[c(e_j) - c(e_i)\right] \\ & = e_i\left[\frac{4}{3} - 3\alpha\right] + c(e_i)\left[-\frac{13}{15} - \frac{3}{10}\alpha + \frac{3}{10}\beta\right] + e_j\left[\frac{4}{3} - 3\alpha\right] - \frac{3}{10}\beta c(e_j)\end{aligned}$$

An agent who expects to exert higher effort than his co-worker, maximizes

$$\begin{aligned}\max_{e_i} \quad & -c(e_i) + \frac{4}{10}\left[\frac{10}{3}e_i + \frac{10}{3}e_j + \frac{1}{3}c(e_j)\right] - \frac{3}{10}\alpha\left[10e_i + 10e_j + c(e_i)\right] \\ & -\frac{3}{10}\alpha\left[c(e_i) - c(e_j)\right] \\ & = e_i\left[\frac{4}{3} - 3\alpha\right] + c(e_i)\left[-1 - \frac{3}{5}\alpha\right] + e_j\left[\frac{4}{3} - 3\alpha\right] + c(e_j)\left[\frac{2}{15} + \frac{3}{10}\alpha\right]\end{aligned}$$

It can be easily seen that agents with positive α and β (where $\alpha \geq \beta$) have a dominant strategy to exert minimal effort, $e_i^* = 1$ whereas completely selfish agents ($\alpha = \beta = 0$) choose $e_i^* = 3$ irrespective of what their co-worker does.

To summarize, the following strategies describe the unique equilibrium for the equal wage treatment: Selfish agents choose $e = 3$, fair agents choose $e = 1$, only principals with $\beta > \frac{2}{3}$ pay positive wages and match their payoff with the agent who exerted lower effort by paying $w = \frac{10}{3}e_1 + \frac{10}{3}e_2 + \frac{1}{3}c(\min[e_1, e_2])$. The resulting average effort for the EWT is thus again 1.6 whereas the average predicted wage is 8.58, marginally lower than the 8.69 in the IWT.