IZA DP No. 3959

Treating Equals Unequally: Incentives in Teams, Workers' Motivation and Production Technology

Sebastian Goerg
Sebastian Kube
Ro'i Zultan

January 2009

Treating Equals Unequally: Incentives in Teams, Workers' Motivation and Production Technology

Sebastian Goerg
University of Bonn

Sebastian Kube
Max Planck Institute for Research on Collective Goods and IZA

Ro'i Zultan
Max Planck Institute for Economics

Discussion Paper No. 3959
January 2009

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 3959
January 2009

ABSTRACT

Treating Equals Unequally: Incentives in Teams, Workers' Motivation and Production Technology

The importance of fair and equal treatment of workers is at the heart of the debate in organizational management. In this regard, we study how reward mechanisms and production technologies affect effort provision in teams. Our experimental results demonstrate that unequal rewards can potentially increase productivity by facilitating coordination, and that the effect strongly interacts with the exact shape of the production function. Taken together, our data highlight the relevance of the production function for organization construction and suggest that equal treatment of equals is neither a necessary nor a sufficient prerequisite for eliciting high performance in teams.

JEL Classification: C92, D23, D63, J31, J33, J41, M12, M52

Keywords: team incentives, equity, production function, social preferences, laboratory experiment, discriminating mechanism, mechanism design

Corresponding author:

Sebastian Kube
Max Planck Institute for Research on Collective Goods
Kurt-Schumacher-Str. 10
53113 Bonn
Germany
E-mail: kube@coll.mpg.de

1 Introduction

"Some contend that differentiation is nuts – bad for morale. They say that differential treatment erodes the very idea of teamwork. Not in my world. You build strong teams by treating individuals differently. ... Everybody's got to feel they have a stake in the game. But that doesn't mean everyone on the team has to be treated the same way."

Jack Welch, former chairman and CEO of General Electric[1]

A general feature of incentive schemes in organizations is a non-uniform distribution of benefits among its agents, which usually accounts for the heterogeneity in agents' ability and performance. As long as the discrimination is based on individual differences, i.e., as long as *unequal* agents are rewarded unequally, there should be little scope for fairness considerations to induce dissonance among the agents.[2] However, a recent theoretical model developed by Eyal Winter (2004) shows that it might even be optimal to treat *equal* agents unequally – depending on externalities given by the production function. This surprising result, derived under the standard assumptions of fully rational, self-centered and money-maximizing behavior, seems to stand in sharp contrast to the implications from research on fairness and equity preferences, whose bottom line is that *"even a small intrinsic concern for justice, .. may have significant effects on .. wage structure"* (Konow (2000), p. 1089; see also Bolton and Ockenfels (2000), Fehr and Schmidt (1999), Mowday (1991), Young (1994) or Selten (1978)). In the present paper, we experimentally explore the interaction in teams and test within the framework of Winter's model whether the psychological cost of the inequality induced by a discriminating mechanism deters from the efficiency of the theoretical optimal mechanism. Thus, to the best of our knowledge, we report the first empirical evidence on the interplay between equity, coordination and production function within teams.[3]

[1]Quote from Welch and Byrne (2001).

[2]A necessary assumption for this statement is that agents are aware of the individual differences and do not misperceive the direction of the differences; which might for example not hold true if agents are overconfident about their own performance (see Ross and Sicoly (1979) for early evidence on overconfidence about contribution to a joint project).

[3]The existing literature on team production and teamwork, e.g., Alchian and Demsetz

The general model as described in Winter (2004) features n risk-neutral agents who work on a project. Each agent i decides simultaneously whether to work ($e_i = 1$) or shirk ($e_i = 0$). Exerting effort is connected with costs c, with c being constant across all agents. Individual effort is assumed to be non-observable and non-contractible. Instead, agents' rewards are contingent on the success of the project, i.e., agents receive individual rewards $b = (b_1, ..., b_n)$ if the project succeeds and 0 otherwise. The probability $p(k)$ of the project's success is specified as a function of the number k of agents exerting effort, mapping the effort profiles to $[0, 1]$. In this sense, $p(k)$ can be interpreted as the project's technology or production function. We assume $p(k)$ to be strictly increasing in k. Depending on the exact specification of $p(k)$, the production function can be modeled to have increasing or decreasing returns to scale. By increasing returns to scale we mean that the production function is one of complementarity, i.e., that $p(k+1)-p(k)$ increases in k; whereas a production function of substitutability has decreasing returns to scale, i.e., $p(k + 1) - p(k)$ is decreasing in k ($k \in [0, ..., n - 1]$).[4]

In the following, a reward vector b is said to be strongly incentive-inducing if it induces all agents to exert effort as a unique Nash equilibrium, and it is optimal if it does so at minimal cost of rewards. The mechanism is symmetric if rewards are constant across all agents. It can be shown that such a *symmetric*, optimal, strongly incentive-inducing mechanism exists if and only if the production function is one of *substitutability*. Contrarily, a production function of *complementarity* implies the optimal, strongly incentive-inducing mechanism to be fully *discriminating* – even if all agents are perfectly symmetric!

(1972), Nalbantian and Schotter (1997) or Irlenbusch and Ruchala (forthcoming), usually focusses on the problem of free-riders and provides means to organize and discipline selfish workers. Complementing this line of research, our paper points to the difficulties that can arise if incentive schemes originally designed for selfish agents are applied to other-regarding agents; thus, interestingly, in our setup it is the absence of selfish agents, and not their presence, that constitutes a potential source of inefficiency for work teams.

[4]For the sake of simplicity we only consider the two extreme cases of increasing or decreasing returns to scale here. In general, the production function could take any form, as long as it satisfies the assumption of $p(k)$ being strictly increasing in k.

Consider that a technology of increasing returns to scale is a sufficient, but not a necessary, condition for full discrimination. In fact, it is only necessary that an agent's incentive to exert effort increases with the number of other agents who do so, which for example might also be caused by some psychological effect like peer pressure (cp. Kandel and Lazear (1992), Barron and Gjerde (1997), Falk and Ichino (2006) or Mas and Moretti (2007) and the references therein).

The purpose of the present study is to experimentally test the key findings of Winter's model, namely whether subjects' behavior is indeed sensitive to the externalities given by the production technology, and whether a major incentive advantage really exists when discriminating among perfectly identical agents; or if the psychological cost of the unequal treatment of equals drives a wedge between the initially predicted and the actually observed efficiency.

Ideally, these questions would be examined with 'cloned' workers acting in 'cloned' work environments which differ only with respect to the production function and the reward schemes. To come close to this ideal world, we introduce a simple and parsimonious laboratory experiment that allows us to analyze the interaction between production function, equity considerations, and reward scheme, while at the same time ensuring that agents are perfectly identical. In the experiment, three players work on a joint project and exert costly efforts. Their payoffs, given as reward minus cost, depend on the number of some goods produced by the joint project. Across the different treatments, the characteristics of the production function (either a function of complementarity or of substitutability) as well as of the reward scheme (either a symmetric or a discriminating mechanism) are manipulated.

We find that, as predicted by Winter's model, the subjects in our experiment respond to the shape of the production function. The discriminating reward scheme under the production function of complementarity achieves almost maximum efficiency, whereas it leads to significantly lower efficiency rates under the production function of substitutability. Moreover, our data suggest that subjects' effort choices are highly sensitive to their own reward, but largely unresponsive to the rewards of the other two subjects in their group: The disadvantaged player (receiving the low reward) regularly exerts effort under the production function of complementarity, notwithstanding the unequal treatment of equals. Contrarily, the symmetric reward scheme significantly hampers efficiency, demonstrating that equal treatment of equals is not necessarily a prerequisite for eliciting high performance in teams, and that unequal treatment can facilitate coordination within the workforce.

The insights gained from our experiment are of significant importance for research on optimal mechanism design in general, but especially in the context of work contracts and organizations. As Winter puts it: *"A large number of models in personnel economics establishes that unequal treatment of unequal agents may have major incentive advantages. The particular importance of demonstrating the optimality of treating* ***equals*** *unequally is that it potentially implies an additional gain for inequality in each of these models"* (Winter (2004), p. 766). We complement this assertion by ascertaining it in an empirical way.

In this regard, we contribute to the question of "equality versus inequality", which is at the heart of the debate in organizational management. Internal inequity is thought to have a tendency to lead to morale problems and to interfere with teamwork (cp. Akerlof and Yellen 1990, Milgrom and Roberts 1992, or Bewley 1999, chapter 6), whereas equal wages are usually associated with positive effects (e.g., increased peer monitoring or lower transaction costs, see Knez and Simester 2001 or Prendergast 1999). However, as Lazear (1989, p. 561) puts it, *".. it is far from obvious that pay equality has these effects."* For example, equal wages do not account for heterogeneity in agents' ability and performance, and payment is not linked to the individual's marginal product, which in turn can lead to free-riding among selfish agents (cp. Holmstrom 1982). Moreover, as we demonstrate in our setup, equal rewards make it hard to form exact beliefs about the others' effort. In contrast, the asymmetry that is created by unequal rewards has the potential to facilitate coordination within the workforce, because it reduces strategic uncertainty about each others' actions.

In real-life organizations, this discrimination is often implemented through non-monetary rewards, e.g., prestige, or by using artificial classifications or (job) titles for seemingly similar tasks, e.g., 'Project Head' or 'Team Captain'.[5] It is often hidden to avoid negative reactions of inequality-averse workers, or fixed by an internal (pay) structure. For example, lawyers, consultants and accountants are paid according to seniority. This special form of *hidden* discrimination creates common knowledge about the stakes that everyone has in the project's success, and thus fosters cooperation and coordination; while at the same time it does not invoke equity concerns because everyone knows that his turn will come to be senior partner. The experimental results in the present paper show that under a production function of complementarity even *transparent* discrimination contributes to efficiency, yet hidden discrimination is effective.

Our study differs from existing experimental studies that analyze the interaction between social preferences and reward schemes in several points. First, the evidence up to now mainly stems from bilateral gift-exchange games between a principal and a single agent (e.g., Fehr et al. (1993, 1997)). What is usually observed in this setup is a positive wage-effort relationship; if the principal shares a large part of the total output with the worker, the worker feels treated fairly and reciprocates by exerting a high effort. While this suggests that most workers care about fairness along a vertical dimension, our question about possible horizontal comparisons within the workforce is usually not addressed.[6] Second, the existing studies are mainly conducted in an incomplete-contract framework where effort and/or wage is non-contractible, while we allow for complete contracts.[7] Third, the usual experimental setup features a principal who can set wages anew in each period, but this introduces uncontrolled elements of intentionality and reputation. Agents can

[5]The 'Team Captain', as the one carrying the responsibility and possible blame for unsuccessful results, is highly motivated to exert effort. Therefore, he functions to incentivize the other team members in the same way as the high-reward agent in our model induces cooperation and high productivity. Cp. also Winter (2004), p. 769.

[6]Two exceptions are notable which feature a multi-agents setup. In Charness and Kuhn (2007), two workers differ in productivity. The authors find that co-workers' wages do not matter much for agents' decisions. Contrarily, Abeler et al. (2006) demonstrate that paying equal wages to workers exerting different efforts leads to a strong decline in efficiency over time.

[7]In Keser and Willinger (2000) agents' actions are hidden, but wage payments can be made contingent on the observed output. However, again the focus is on the vertical comparison between a principal and a single agent. Fehr et al. (2007) provide a direct comparison on the efficiency of incomplete and complete contracts in a bilateral setup.

withhold effort to punish and enforce principals to pay higher wages in the future, which to us not only seems difficult to reconcile with real-world work-relationships, but additionally is outside the scope of Winter's model. Fourth, to the best of our knowledge we are the first to pay attention to the important role of the production function in a labor market setting.[8] Our finding that agents' behavior is sensitive to the shape of the production function should be taken into account in future empirical research on the interaction between social preferences and reward schemes.

The remainder of this paper is organized as follows: In the next section, we describe the experimental design and derive theoretical predictions. Subsequently, the experimental results are presented and discussed in Section 3, and Section 4 concludes.

[8]Normann et al. (2007) examine the relation between production function and the existence of large-buyers' discounts.

2 Experimental Setup

Our experiment is designed to explore the interplay between production function, monetary rewards and social preferences within a team of agents. More specifically, we check if workers' behavior is sensitive to the type of the production function that they face in their joint project. Additionally, we test for the tension between equity and efficiency in a team environment, as induced by the interaction between social (other-regarding) preferences and reward schemes.

2.1 Experimental Design

In the game we have three agents working on a joint project. Each agent i individually decides whether to work ($e_i = 1$) or shirk ($e_i = 0$), and the individual cost of exerting effort is 90 Taler[9]. In case that an agent exerts effort, the costs of 90 are deducted from his individual reward.[10] The output of the project, i.e. the number of produced units, depends on the number of agents $\sum_i e_i$ choosing to work, and on our treatment variable **production function**:

	number of units produced if...			
production function	$\sum_i e_i = 0$	$\sum_i e_i = 1$	$\sum_i e_i = 2$	$\sum_i e_i = 3$
complementarity (COM)	20	40	65	100
substitutability (SUB)	20	55	80	100

The first case (COM) describes a production function of *complementarity.* The technology has increasing returns to scales, since the number of produced units (the output of the project) is $p(0) = 20$ if all agents shirk, $p(1) = 40$ if two agents shirk, $p(2) = 65$ if only one agent shirks and $p(3) = 100$ if all agents work, thus $p(3) - p(2) > p(2) - p(1) > p(1) - p(0)$. In the second

[9]Taler is our experimental currency. Talers earned in the experiment were converted at a rate of 80 Taler = 1 Euro.

[10]Exact control over the underlying cost and production functions is crucial for testing the theoretical model - which would not have been the case under, e.g., a real effort task.

case (SUB), we have a production function of *substitutability*. The technology has decreasing returns to scale, since $p(3)-p(2) < p(2)-p(1) < p(1)-p(0)$.

Agents' rewards are made contingent on the output of the project and the **reward mechanism** or remuneration scheme, which we vary across treatments. The reward mechanism in treatment 444COM is symmetric. Each agent in the group receives a reward of 4 Taler per unit produced. Contrarily, the mechanism implemented in treatments 345COM and 345SUB is a discriminating one: agents' reward per unit produced is either 3, 4, or 5 Taler (with each possibility occurring exactly once). At the same time, the sum of the individual rewards does not differ across the reward mechanisms. For example, the total reward costs in case that all agents shirk equals $3(4 \cdot 20) = 240$ under the *symmetric* reward mechanism, and $3 \cdot 20 + 4 \cdot 20 + 5 \cdot 20 = 240$ under the *discriminating* reward mechanism.

Our experiment was conducted in a labor market framing, avoiding loaded terms (e.g., 'shirk' or 'success'). We used the same procedure in our three treatment conditions 444COM, 345SUB and 345COM. Upon arrival, participants were randomly divided into groups of three. In the treatments with a discriminating reward scheme, the three possible rewards were randomly assigned within each group. The written instructions were distributed and read out aloud. Afterwards, subjects could pose questions in private, and had to answer a set of computerized control questions to ensure that everybody had understood the game and to make subjects familiar with the operation of the program. Then subjects were told their own reward and the other players' rewards, and simultaneously had to decide between working or shirking. Afterwards, it was announced that we were additionally interested in their beliefs about the other subjects' behavior, and each subject had to state what they expected the first and the second other player in their group to choose.[11] In case that their belief matched the actual behavior, subjects were paid an additional 20 Taler. Only then we announced that five additional periods of the game would follow, in which everything was kept

[11]E.g., a player receiving a reward of 3 Taler per unit had to choose between '4' and '5' shirk, '4' and '5' work, '4' works and '5' shirks, or '4' shirks and '5' works. To keep the procedure constant, in 444COM we also asked separately for the behavior of the two other players in the group.

constant (individual rewards, costs, production function and group composition). This was done to allow for possible learning to take place. After our experiment, subjects had to complete a social value orientation test[12] and a socio-economic questionnaire.

The computerized[13] experiments were run in 2007 at the BonnEconLab at the University of Bonn. Participants were randomly recruited via email invitation out of approximately 3000 persons from the BonnEconLab's subject pool (including mostly undergraduate students from a large variety of fields). For each treatment, we ran two sessions with 18 subjects each; totalling 12 independent matching groups (all periods) or 36 independent decisions (only first period) per treatment. Unfortunately, in one session in treatment 444COM, only 15 subjects showed up, so that we are missing one of the twelve independent observations in this treatment. A session lasted approximately 70 minutes. Subjects were paid for their decision and their belief in the first period, and additionally for one randomly selected period (which was constant across all subjects within a session) out of the subsequent five periods. On average, subjects earned approx. 7 Euro.

2.2 Behavioral Predictions

In this section we derive the possible equilibria under the classical assumptions of agents being fully rational, self-centered, money-maximizing, and risk-neutral[14]. As will be seen, the degree of efficiency, defined as the sum of agents exerting effort in equilibrium, is sensitive to the production function and to the reward mechanism. Subsequently, we demonstrate how the predictions change once we introduce agents who are additionally motivated

[12] The 'ring test' is described for example in Griesinger and Livingston (1973) or Liebrand (1984); see also Beckenkamp (1995) for an early application in Economics.

[13] The experiment was programmed in Pascal using RATimage by Abbink and Sadrieh (1995). The questionnaire and the ring test were conducted using zTree by Fischbacher (1999). Screenshots of the program can be found in the appendix.

[14] Our game as described in Section 2.1 can be rewritten in a probabilistic way, which is the interpretation used by Winter (2004). We instead opted for a deterministic representation to impose risk-neutrality over the final outcome of the project, i.e., to pay the expected value of a lottery rather than to actually implement the lottery. This allows us to abstract from subjects' individual risk preferences.

by equity considerations. In this case, the degree of inequality of the reward mechanism fundamentally affects equilibrium behavior.

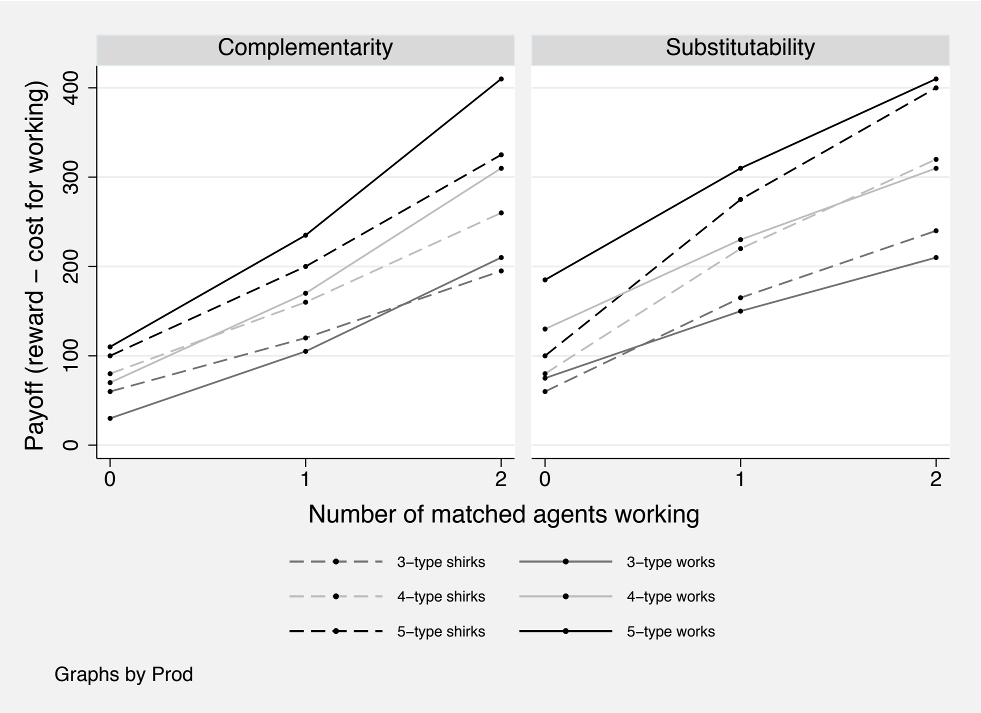
We start by calculating the reward per unit produced that is needed to make an agent just indifferent between working and shirking, which depends on the (belief about the) decisions of the other two players in the group. Let X_0 denote the reward that is needed if an agent believes that both the other two agents in the group will shirk, and let X_1 and X_2 be the corresponding values when expecting one, resp. none of the others to shirk. Under a production function of *complementarity*, X_0 is given as $40X_0 - 90 = 20X_0 \Leftrightarrow X_0 = 4.5$, i.e., the payoff from working must equal the payoff from shirking under the belief that both the others shirk. Analogously, we find that $X_1 = 3.6$ and $X_2 = 90/35 \approx 2.6$.

This implies that the high-reward player in 345COM, receiving a reward of 5 per unit produced, will always work, irrespective of his beliefs (since $5 > X_0 > X_1 > X_2$). Anticipating this, the feasible beliefs for the medium-reward player are such that he also has a dominant strategy to work (since $4 > X_1 > X_2$). The only feasible belief of the low-reward player is thus to expect both the others to work, in which case his reward induces him to work as well (since $3 > X_2$). Hence the *discriminating* scheme enables players to form exact beliefs about the other players' decisions, although they move simultaneously – and repeated elimination of strongly dominated strategies leads to the unique Nash equilibrium of all players exerting effort.

Contrarily, this line of reasoning is not applicable when using the *symmetric* reward mechanism. Each player works only if he has the belief that at least one other player exerts effort as well (since $X_0 > 4 > X_1 > X_2$). This implies that in 444COM we have two equilibria in pure strategies: Either all agents work, or all agents shirk (with all work being the payoff- and risk-dominant equilibrium). Besides that, also an equilibrium in mixed strategies exists in which the probability of shirking is approximately 0.77 (and all players know that each of the other players will shirk with this probability).

If we switch to the production function of *substitutability*, first consider that a naive principal might be tempted to prefer this technology over the

Figure 1: *Player's payoff function by type and decision*



In both treatments, the payoffs for the 5-type players from working dominate the payoffs from shirking, and so the possibility of zero matched agents working can be eliminated for the other players. Now working dominates shirking for the 4-type players under complementarity, while the opposite is true for the 3-type players under substitutability. The remaining player in both treatment now maximizes her payoff by working. Thus the equilibria are derived through repeated elimination of dominated strategies. The two pure equilibria of the egalitarian treatment 444COM are revealed by the crossover of the payoff functions of the 4-type player under complementarity. Note that the gain from working can be seen to increase (diminish) under complementarity (substitutability).

previous one. For any given effort sum, the number of units produced is always equal or higher under substitutability than under complementarity. However, in 345SUB the discriminating reward scheme is not optimal anymore, because the threshold-order is reversed under a production function of *substitutability* (i.e., $X_0 \approx 2.6$, $X_1 = 3.6$ and $X_2 = 4.5$). Thus, the low-reward player shirks in equilibrium, while the other two players work; and all players hold corresponding beliefs. The payoff structure of the game and the equilibria derivation are illustrated in figure 1.

In fact, Winter (2004) shows that the total costs of an optimal, strongly incentive-inducing reward scheme are always lower under complementarity than under substitutability. With respect to this, a sophisticated principal who believes that his agents are risk-neutral money-maximizers should always try to 'design' the project in a way that workers' efforts are complements rather than substitutes – for example, he might prefer a function-based technology over a process-based one. Yet Winter's prediction crucially depends on the assumption of subjects being *self-centered* money-maximizers. By contrast, part of the literature (not only) in Behavioral and Experimental Economics suggests that, beside pure money maximization, a non-negligible fraction of subjects is strongly motivated by *other-regarding* considerations – in particular, subjects exhibit a basic desire for equity, including a preference for equal treatment of equals (cp. Selten (1978), Mowday (1991), Roemer (1996)), and a preference for equal payoff distributions (cp. Fehr and Schmidt (1999) or Bolton and Ockenfels (2000)).

In the presence of equity considerations, any discriminating reward mechanism comes at some hidden costs which incentivize agents to shirk, even under an initially incentive-inducing mechanism! Slight equity preferences[15] are already enough to let the superiority of the discriminating rewards vanish in 345COM. If agents' loss of utility of another agent receiving a higher payoff than their own is as low as 1/6 of the loss of utility of reducing their own payoff by the same amount, all-shirk becomes the unique equilibrium in 345COM.[16] Even worse, due to the recursive nature of the equilibrium in Winter's model, the sheer belief that one or both of the other agents might have equity preferences can alone lead to a loss of efficiency – even if all agents themselves are strictly self-centered money-maximizers. By contrast in 444COM, equity preferences provide additional incentives *not* to shirk: If a subject expects the other two players in his group to work, shirking will reduce his payoff and lead to a less equitable payoff distribution ($(260, 170, 170)$

[15]Throughout the paper, equity preferences are defined over payoffs rather than effort levels (cp. Mohnen et al., forthcoming).

[16]The intuition behind this hypotheses can easily be seen if we reconceive above equilibria derivations using an extended utility function which incorporates equality preferences, e.g., the function described in Fehr and Schmidt (1999). Using their model, all-shirk is a possible equilibrium in 345COM if $\alpha \geq 1/3$ and $\beta = 0$ – which is a very conservative estimate in comparison with empirical estimations. Since the exact calculations are rather tedious and lengthy, they are available from the authors upon request.

instead of $(310, 310, 310)$); which is something that (not only) an inequality-averse subject would never prefer.

Taken together, if subjects are motivated by equity considerations instead of being self-centered and money-maximizing, in the symmetric treatment we cannot necessarily predict a lower efficiency than in the discriminating treatment anymore. Already a meager amount of (uncertainty about) equality preferences reverses the initially low amount of strategic uncertainty in 345COM, whereas the symmetric mechanism in 444COM is rather robust in this respect. This is recapitulated in the following table, which lists the possible equilibria in pure strategies for self-centered subjects in the first row, and in the second row the equilibria that might additionally emerge in the presence of equity-considerations:

Treatment	345SUB	345COM	444COM
production function	substitutability	complementarity	
reward mechanism	discriminating		symmetric
self-centered	(0,1,1)	(1,1,1)	(1,1,1),(0,0,0)
inequality-averse	(0,0,1)	(0,0,0),(0,0,1),(0,1,1)	*no additional equ.*

3 Experimental Results

In this section, we present the results of our experiment with regard to our research questions. First, we show that workers' behavior is indeed sensitive to the type of production function they face in their joint project, and that the unequal treatment of equals does not necessarily hamper full effort provision. Accordingly, we then present data on a change in the reward scheme from a discriminating to an egalitarian one, which suggests that equal treatment of equals does not necessarily promote full effort provision within a team of agents.

3.1 Sensitivity to the Production Function

Figure 2 shows the mean effort over all periods, conditional on player's reward type, for our three treatments 345COM, 345SUB and 444COM.

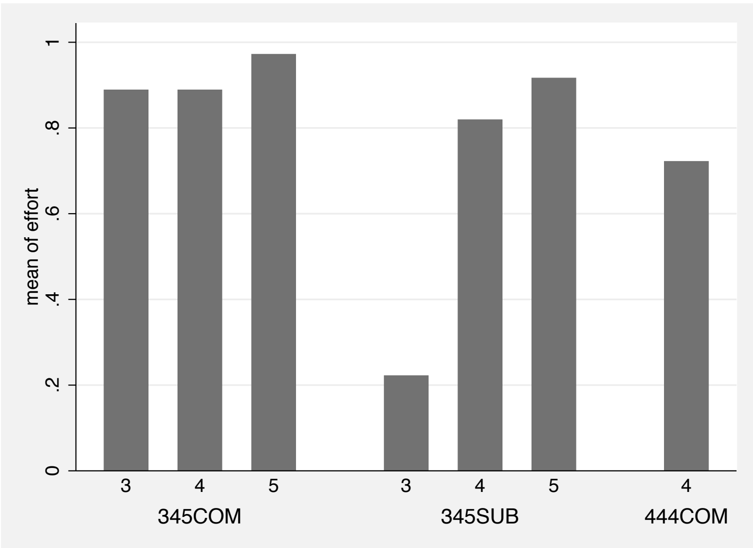


Figure 2: Mean effort per reward type for the three treatments

Focussing on the discriminating reward scheme, overall effort levels are much higher under a production function of complementarity than under a production function of substitutability. 91.7% of all effort decisions in 345COM are to work, compared to only 65.3% in treatment 345SUB (rank-sum test, $Prob > |z| = .0004$). In 345COM, 6 out of 12 (9/12) groups exert full effort in all periods (all but one period), whereas the same is never observed in 345SUB.

The difference in efficiency between 345COM and 345SUB is predicted to stem from a difference in the behavior of the low-reward type in equilibrium. As can be seen in Figure 2, the average effort level of the low-reward type in 345SUB is significantly lower than that of the other two types (22.2% vs. 81.9% and 91.7%, signed-rank test, $Prob > |z| = .0074$ and .0039). It is also significantly lower than the effort level of the same type in the complementarity treatment (22.2% vs. 88.9%, rank-sum test, $Prob > |z| = .0001$). Subjects' individual beliefs are in line with the finding. In 345COM, medium- and high-reward players believe that the low-reward player will work in 85% of all cases, while in 345SUB the low-reward player is expected to work in only 33% of all instances (rank-sum test, medium-reward: $Prob > |z| = .0004$, high-reward: $Prob > |z| = .0007$).

The effort levels of the medium- and high-reward types in 345COM (88.9% and 97.2%) do not differ from the corresponding levels that we observe in 345SUB. Overall, when standard equilibrium predictions dictate effort exertion, the observed effort levels are over 80%. In the one case in which the equilibrium strategy is to shirk, indeed almost 80% of the decisions are to shirk.

Comparing the sums of effort per matching group in the first round and the last round, we find no indication of a significant time trend (signed-rank test: 345COM $Prob > |z| = .65$; 345SUB $Prob > |z| = .31$).[17] The behavior in the first round reveals a similar picture as above. Again, the number of low-reward players choosing to exert effort is much higher in 345COM than in 345SUB (16.7% vs. 75%; Fisher's exact test $p = .012$). For the other two player types the differences across treatments are only marginal and statistically insignificant (medium-reward type: 91.7% vs. 91.7%, $p = 1$; high-reward type: 100% vs. 91.7%, $p = 1$).

Taken together, under a discriminating reward scheme the effort provi-

[17] Cp. the corresponding time-series data in Figures 3 and 4 in the appendix.

sion is almost maximal when using a complementarity production function; whereas significantly lower effort rates are observed when using one of substitutability. We thus conclude from our experimental data:

Result 1: *In line with Winter's model, agents' observed behavior is sensitive to the production technology. Treating equals unequally by using a discriminating reward scheme leads to almost full efficiency under a production function of complementarity – whereas the same reward scheme does not perform well under a production function of substitutability.*

3.2 Sensitivity to the Reward Scheme

Before we turn our attention to the influence of the reward scheme, let us again point out that the discriminating reward scheme leads to almost maximum efficiency under a production function of complementarity, and that even the low-reward players exert effort in this situation. Notice that this is not to say that equity considerations are completely absent in 345COM. For example, the average rate of effort provision over all periods is significantly different between the low- and the high-reward type (88.9% vs. 97.2%, signed-rank test, $Prob > |z| = .0261$).[18] Also the beliefs of medium- and high-reward players about the low-reward player's decision in the first period reveal some influence of equity considerations, because 42% (wrongly) expect him to shirk.[19] But although we observe some small signs of equity considerations in our data, the influence on the final outcome is negligible in view of the high degree of efficiency in 345COM. This finding is surprising in the face of past findings from other experiments exploring the importance of equity concerns – and the following observations from our equitable reward mechanism treatment 444COM gives rise to further surprise.

Given a production function of complementarity and keeping the total

[18] Yet, note that i) this is likely to be caused by a ceiling effect; ii) in 5 out of the 8 instances where the low-reward player shirks, the behavior might also be explained by self-centered preferences (best response given the individual belief); iii) behavior in the first period actually does not differ significantly across types.

[19] However, this vanishes with growing experience. The frequency changes to 85% if we take all periods into account.

cost of the reward mechanism constant, the observed mean efficiency in the first period is lower under the symmetric reward mechanism (78.9%) than under the discriminating one (88.9%), albeit non-significant (rank-sum test, $Prob > |z| = .1552$). Over the course of the experiment, the difference grows larger and significant (average efficiency over all periods of 72.2% vs. 91.7%, rank-sum test, $Prob > |z| = .0649$; cp. Figure 2). On average, every reward type in the discriminating treatment provides more effort than the players in the symmetric treatment. Only 3 out of 11 groups (4/11) exert full effort in all periods (all but one period), compared to 6 out of 12 (9/12) groups in 345COM. Moreover, the standard deviation of group efficiencies is significantly higher in 444COM than in 345COM (0.233 vs. 0.158, Conover's squared-ranks test, $Prob > |z| = .0143$). Notice that the difference is not an artifact resulting from the high degree of efficiency in 345COM (which puts a bound on the variance), as the group efficiencies in 345SUB, in which the overall efficiency is similar to that in 444COM, show an even lower standard deviation of 0.068 (cp. Figure 6 in the Appendix).

Our result suggests that equal treatment of equals does not necessarily promote full effort provision within a team of agents. A potential reason for the observed difference in efficiency between the symmetric and the discriminating mechanism might be the introduction of the additional 'all-shirk'-equilibrium in treatment 444COM. Even though it is payoff- and risk-dominated by the 'all-work'-equilibrium, the multiplicity of equilibria introduces strategic uncertainty (cp. van Huyck et al. (1990)).[20] Players formulating beliefs are uncertain whether the other players in their group will work or shirk, which is visible in our data: 83% expect both other players to work in 345COM, whereas only 62% do so in 444COM (rank-sum test, $Prob > |z| = .0979$). This translates into low efficiency rates and a high variance of group efficiencies in 444COM, suggesting that strategic considerations shaped by the reward mechanism are crucial, and outweigh possible equity preferences of the subjects.[21] The asymmetry of the reward mech-

[20]Note that strategic uncertainty should also be present in 345COM (cp. Section 2.2), because 'all-work' *and* 'all-shirk' are potential equilibria once we allow for equity considerations. Yet, we observe almost full efficiency in this treatment.

[21]One might consider that the result may be driven by a difference in the subject pop-

anism facilitates coordination among the agents. Under the discriminating reward scheme, subjects can anticipate that the high-reward player will exert effort, which in turn incentivizes the medium- and low-reward players to do so as well. On the other hand, the identical rewards under the symmetric mechanism make it hard for the subjects to form beliefs about the action of the other players, so that they are all in the dark. We thus conclude from our experimental data:

Result 2: *Treating equals equally is neither a necessary nor a sufficient prerequisite for eliciting high performance in teams. Asymmetry facilitates coordination under a production function of complementarity, i.e., we observe higher efficiency rates under a discriminating reward mechanism than under a cost-equivalent symmetric one – which is again in line with Winter's model.*

4 Conclusion

In this paper, we studied the interaction in teams. More specifically, we experimentally explored whether workers' behavior is sensitive to the externalities given by the production technology, and whether a major incentive advantage exists when discriminating among perfectly identical agents. In our experiment, three workers simultaneously decide on their individual provision of costly effort to a joint project. Treatments differ in the shape of the project's production technology and of the reward mechanism. Under a production technology of complementarity, the use of a symmetric reward mechanism elicits substantially lower efforts and efficiency than a cost-equivalent discriminating reward mechanism. The same discriminating reward mechanism underperforms when it is utilized under a production function of substitutability.

Our findings have important implications for the design of organizations

ulation between treatments. However, pairwise comparisons of the corresponding results of the social value orientation test reveals no significant differences across treatments (345COM vs. 345SUB $p = .66$; 345COM vs. 444COM $p = .52$; 444COM vs. 345SUB $p = .98$); in all treatments, subjects' 'value orientations' indicate a general preference for equitable treatment, i.e., they do differ significantly from being strictly self-centered (345COM: $p \leq .001$; 345SUB: $p \leq .001$; 444SUB: $p \leq .001$).

in practice. First, they clearly point to the relevance of the production function for organization construction – a factor which has so far received little attention in the literature. Designing (production) tasks in a way that makes workers' efforts complements rather than substitutes may lead to a major cost advantage. Insofar as peer pressure constitutes a complementarity in effort exertion, the strengthening of social ties amongst the workforce alone might have a strong impact on productivity.

Second, and closely related, is our finding that unequal treatment of equals does not necessarily hamper efficiency. Whenever the organizational technology is one of complementarity, i.e., whenever the impact of a worker's input increases in the size of the others' input, the usage of a discriminating reward scheme might be potentially efficiency-enhancing. The main reason for this is that asymmetric rewards facilitate coordination, because workers can anticipate that those who have high stakes at hand will certainly exert effort – which in turn incentivizes the other workers to exert effort as well. Consider that discrimination must not necessarily be in monetary terms, but might also take the form of hierarchies. While a vast body of literature in personnel economics already promotes the implementation of hierarchies (e.g., Lazear and Rosen (1981)), our results suggest that hierarchies might enhance performance despite the absence of the existing literature's usual assumptions of monitoring or authority relations.

In this regard, we more generally contribute to the ongoing research on behavioral phenomena in organizations. As James Konow (2000) puts it: *"Many of the successes of economics can probably be attributed to its pushing the assumption of self-interest to the extreme. To proceed further, however, it may be necessary to incorporate richer behavioral assumptions that include fairness and other moral standards."* (Konow (2000), p. 1089). While we agree in principle, it should be added that it is additionally necessary to identify the situations in which behavior is in line with the classical model – which is ultimately an empirical question. Only then can we really understand how to model the richer behavioral assumptions in a way to advance Economics.

The results in this paper should not be taken as arguments against the importance of equity considerations in general. They rather suggest that equal treatment of equals is neither a necessary nor a sufficient prerequisite for eliciting high performance in teams.[22] Yet the relative importance of equity considerations is likely to depend on the exact details of the organizational setting and framework. In this paper, we presented experimental evidence for some of these settings, and stressed the interaction between production technologies and reward mechanisms. Other interesting variations of the organizational settings include a change in the timing of effort choices, the introduction of heterogeneity among the workforce or the use of 'symbolic' instead of monetary differentiation. Extending our simple design allows for studying these and other interesting aspects in the future.

[22]Taking a theoretical viewpoint, one might argue that this result is not surprising since it immediately arises from Winter's model. However, theoretical results need not necessarily be compatible with actual behavior. This is nicely captured by the following quote from Falkinger et al. (2000, p. 248) in the context of public goods provision mechanisms: *"[The theoretical] mechanisms are desirably simple and do well in theory. It is, however, important to note that the fact that a mechanism does well in theory, does not tell us much about its effectiveness in the laboratory and in practice. In principle, it could well be the case that although the Nash equilibrium in the presence of the mechanism implies an efficient provision of the public good, subjects' actual behavior will generate significant under- or overprovision. .. Deviations of actual behavior from the equilibrium predicted by theory .. can arise because subjects' motivations differ from the theoretically assumed preferences."* Subject to these considerations, our empirical findings become all the more important.

References

Abbink, Klaus, and Karim Sadrieh. 1995. "RATImage - Research Assistance Toolbox for Computer-Aided Human Behavior Experiments." SFB Discussion Paper B-325, University of Bonn.

Abeler, Johannes, Steffen Altmann, Sebastian Kube, and Matthias Wibral. 2006. "When Equality is Unfair." IZA Discussion Paper 2500.

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

Alchian, Armen A., and Harold Demsetz. 1972, "Production, Information Costs, and Economic Organization." *American Economic Review*, 62(5): 777-795.

Barron, John E., and Kathy Paulson Gjerde. 1997. "Peer Pressure in an Agency Relationship." *Journal of Labor Economics*,15(2): 234-254.

Beckenkamp, Martin. 1995. *Wissenspsychologie. Zur Methodologie kognitionswissenschaftlicher Ansätze.* Heidelberg, Kroening: Asanger Verlag.

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

Bolton, Gary E., and Axel Ockenfels. 2000. "ERC: A Theory of Equity, Reciprocity and Competition." *American Economic Review*, 90: 269-299.

Charness, Gary, and Peter J. Kuhn. 2007. "Does Pay Inequality Affect Worker Effort? Experimental Evidence." *Journal of Labor Economics*, 25: 693-723.

Falk, Armin, and Andrea Ichino. 2006. "Clean Evidence on Peer Effects." *Journal of Labor Economics*, 24(1): 39-57.

Falkinger, Josef, Ernst Fehr, Simon Gächter, and Rudolf Winter-Ebmer. 2000. "A simple mechanism for the efficient provision of public goods – Experimental Evidence." *American Economic Review*, 90(1): 247-264.

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

Fehr, Ernst, Georg Kirchsteiger, and Arno Riedl. 1993. "Does Fairness Prevent Market Clearing? An Experimental Investigation." *Quarterly Journal of Economics*, 108: 437-459.

Fehr, Ernst, Alexander Klein, and Klaus M. Schmidt. 2007. "Fairness and Contract Design." *Econometrica*, 75(1): 121-154.

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

Fischbacher, Urs. 1999. "z-Tree - Zurich Toolbox for Readymade Economic Experiments – Experimenter's Manual." IEW Working Paper.

Griesinger, Donald W., and James W. Livingston. 1973. "Toward a model of interpersonal motivation in experimental games." *Behavioral Science*, 18: 173-188.

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

Irlenbusch, Bernd, and Gabriele Ruchala. forthcoming. "Relative Rewards within Team-Based Compensation." *Labour Economics.*

Kandel, Eugene, and Edward Lazear. 1992. "Peer Pressure and Partnerships." *Journal of Political Economy*, 100(4): 801-817.

Keser, Claudia and Marc Willinger. 2000. "Principal's Principles when Agents' Actions are Hidden." *International Journal of Industrial Organisation*, 18: 163-185.

Knez, Marc, and Duncan Simester. 2001. "Firm-Wide Incentives and Mutual Monitoring at Continental Airlines." *Journal of Labor Economics*, 19:743-772.

Konow, James. 2000. "Fair Shares: Accountability and Cognitive Dissonance in Allocation Decisions." *American Economic Review*, 90(4): 1072-1091.

Lazear, Edward P. 1989. "Pay Equality and Industrial Politics." *Journal of Political Economy*, 97: 561-580.

Lazear, Edward W., and Sherwin Rosen. 1981. "Rank-Order Tournaments as Optimum Labor Contracts." *Journal of Political Economy*, 89(5): 841-864.

Liebrand, Wim. 1984. "The effect of social motives, communication and group size on behavior in an n-person multi-stage mixed-motive game." *European Journal of Social Psychology*, 14: 239-264.

Mas, Alexander, and Enrico Moretti. 2007. "Peers at Work." U.C. Berkley Working Paper.

Milgrom, Paul, and John Roberts. 1992. *Economics, Organization and Management.* Englewood Cliffs, New Jersey: Prentice Hall, 1st edition.

Mowday, Richard T. 1991. "Equity Theory Predictions of Behavior in Organizations." In: *Motivation and Work Behavior*, 5^{th} edition, ed. Richard M. Steers and Lyman W. Porter. New York: McGraw-Hill.

Nalbantian, Haig R., and Andrew Schotter. 1997. "Productivity under group incentives: An experimental study." *American Economic Review*, 87(3): 314-341.

Normann, Hans-Theo, Bradley Ruffle and Chris Snyder. 2007. "Do Buyer-Size Discounts Depend on the Curvature of the Surplus Function? Experimental Tests of Bargaining Models." *RAND Journal of Economics*, 38(3): 747-767.

Mohnen, Alvine, Kathrin Pokorny, and Dirk Sliwka. forthcoming. "Transparency, Inequity Aversion, and the Dynamics of Peer Pressure in Teams: Theory and Evidence." *Journal of Labor Economics.*

Prendergast, Canice. 1999. "The Provision of Incentives in Firms." *Journal of Economic Perspectives*, 37:7-63.

Roemer, John E. 1996. *Theories of Distributive Justice.* Cambridge: Harvard University Press.

Ross, Michael, and Fiore Sicoly. 1979. "Egocentric Biases in Availability and Attribution." *Journal of Personality and Social Psychology*, 37: 322-336.

Selten, Reinhard. 1978. "The equity principle in economic behavior." In: *Decision theory and social ethics, issues in social choice.* ed. Gottinger and Leinfellner, 289-30, Dordrecht: Reifel Publishing,

Van Huyck, John B., Raymond C. Battalio, and Richard O. Beil. 1990. "Tacit Coordination Games, Strategic Uncertainty, and Coordination Failure." *American Economic Review*, 80 (1): 234-248.

Welch, Jack, and John A. Byrne. 2001. *Jack: Straight from the Gut.* New York, NY: Warner Books Inc.

Winter, Eyal. 2004. "Incentives and Discrimination." *American Economic Review*, 94 (3): 764-773.

Young, Peyton. 1994. *Equity in theory and practice.* Princeton, Princeton University Press.

Appendix A: Instructions

This is the English translation of the instructions used in treatments 345COM and 444COM. In treatment 345SUB, the table and examples were adjusted to fit the production function.

Welcome to this decision-making experiment. Please read the following instructions carefully. The experiment will be conducted anonymously, that is to say you will not learn with whom of the other participants you are interacting. Please keep in mind that from now on and throughout the experiment you are **not allowed to talk to the other participants**. If you **have any questions**, please give a signal with your hand and we will come to you. During the experiment you can earn Taler. How much you earn depends on your decisions and the decisions of the other participants in your group. At the end of the experiment these Taler will be converted to Euro at an exchange rate of **80 Taler = 1 EURO**. The Euro amount will be paid out to you. You will be called to collect your earnings. Please turn in all instruction sheets when you collect your earnings.

In this experiment you will be randomly divided into groups of three persons. Together with two other participants you form a group. Each participant decides whether he wants to **work normal or hard**. The more participants choose to work hard, the more units of goods will be produced.

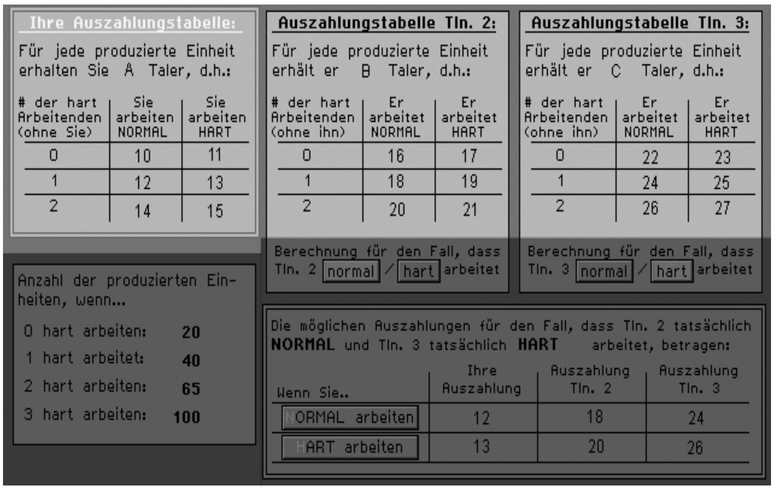
Number (#) of hard working participants	0	1	2	3
Produced units of goods	20	40	65	100

Examples: In case that all participants of the group work normal, 20 units will be produced altogether in your group. If you work hard and another participant in your group works hard as well, 65 units will be produced altogether in your group. etc... good

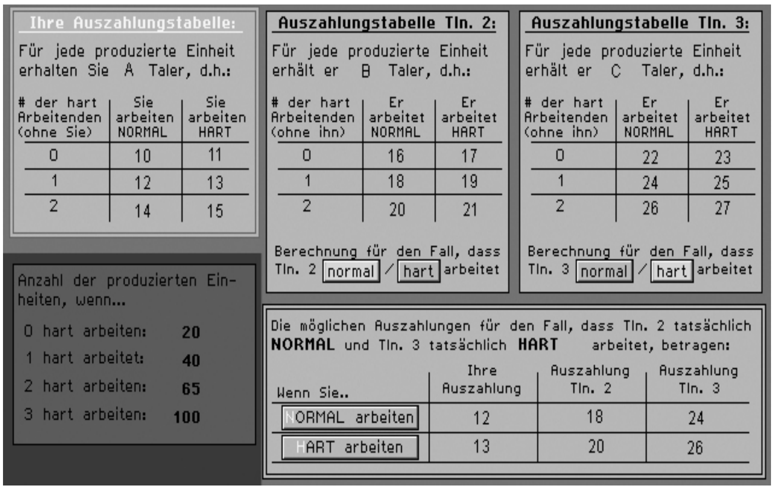
For each unit of goods produced, you receive a certain amount of Taler. At the beginning of the experiment you are informed how many Taler you earn per unit produced. Additionally, you learn how many Taler per unit the other two participants in your group earn. Examples: In the beginning of the experiment you are told that you receive 5 Taler for each unit produced. In case that all participants in your group work hard, 100 units will be produced and you receive 500 Taler. In case that 40 units are produced, you receive $40 \cdot 5 = 200$ Taler. etc...

Costs: If you decide to **work hard**, the amount you receive is reduced by 90 Taler. If you **work normal**, no additional costs arise. Examples: You and another participant in your group work hard, so 65 units are produced. Accordingly, you receive $65 \cdot 5 = 325$ Taler. Since you worked hard, 90 Taler are taken away. Hence, your final payment is $325 - 90 = 235$ Taler. If instead you worked normal, 40 units would be produced. You would receive $40 \cdot 5 = 200$ Taler. Since you worked normal, no Taler are subtracted from this amount. Hence, your final payment would be 200 Taler. etc...

In order to facilitate the decision-making process, each participant is informed in detail about his own possible payoffs and the payoffs of the other two participants in his group. The corresponding information is given in table form. For every participant, a table lists all possible payments dependent on the own decision (to work normal or hard) and the decisions of the other two participants in the group (none, one or both work hard). In these tables, the corresponding costs for working hard have already been subtracted. Below, you see an example with fictional data:



In the lower right part of the screen, you can see another table. At the beginning, the table is empty. In order to display data, you first have to create a hypothetical situation: In the table of participant number 2, click on the corresponding button what you think how he will decide (to work normal or hard). Furthermore, in the table of participant number 3, click on the corresponding button what you think about his decision (to work normal or hard). In the lower table you will then be shown in the first row what the payment for you and the other two participants would be, in case that your chosen situation actually occurs - and that you decide to work normal. The second row lists the possible payments that you and the other two participants would receive, in case that your chosen situation actually occurs - and that you decide to work hard. At any time, you can display data for a different situation. Simply change the situation by clicking on a different button underneath the payment tables of participant number 2 and 3. Below you see another example with fictional data:



Your decision: As soon as you have decided on whether you want to work hard or normal, please click on the according button in the lower right table (on the left hand side). The program will ask you to confirm your decision. Afterwards, your decision will be transferred. Please remain in your cubicle and wait until all participants have reached a decision. Afterwards, you will be informed about the number of units produced in your group and about your payoff. This amount will be paid to you in cash and anonymously at an exchange rate of 80 Taler = 1 EURO.

If you have any questions please give a signal with your hand!

The following instructions were distributed and read out aloud only after the first period.

In the following, the previous procedure will be repeated five times within the same group of persons and with the same numerical values for production function and effort costs. In each of these five periods, you again have to choose between working normal or working hard. In the end, we randomly select one of these five periods. You will receive the payoff for the randomly selected period in addition to your present payoff.

If you have any questions please give a signal with your hand! Otherwise, please click to continue!

Appendix B: Supplementary Data

	Treatment		
Player type	345SUB n=12	345COM n=12	444COM n=33
3	22.2% (9.3)	88.9% (4.3)	-
4	81.9% (6.6)	88.9% (8.3)	72.2% (5.6)
5	91.7% (4.4)	97.2% (1.2)	-
Mean	65.3%	91.7%	72.2%

Observations reflect individual subjects, for each of whom the percentage of 'work hard' decisions out of the six periods was calculated. The standard errors are given in parentheses.

Table 1: *Mean efficiencies (first row) and standard deviation (second row) for all rounds per player type over the treatments*

Player type	345COM vs. 345SUB	345COM vs. 444COM
3	.0001	n.a.
4	n.s.	n.a.
5	n.s.	n.a.
Means	.0004	.0649

Table 2: *Comparison of mean efficiencies by player types between different treatments with two-sided rank-sum test.*

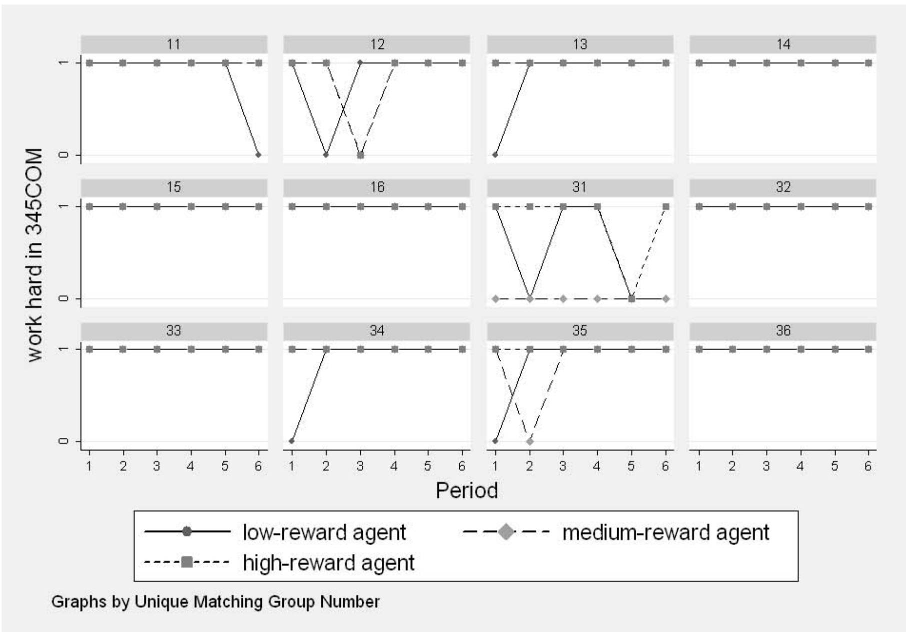


Figure 3: Effort per reward type over time in 345COM

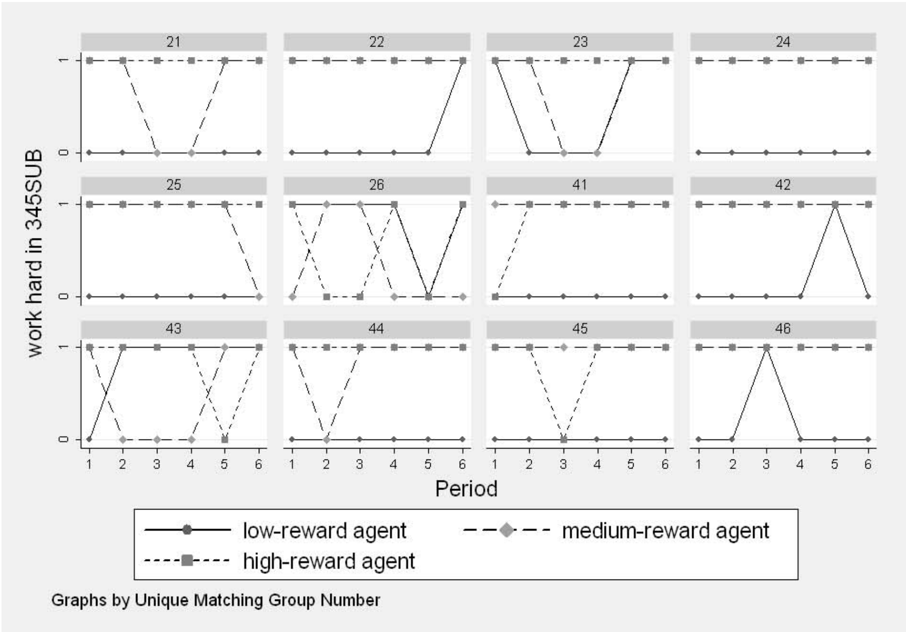


Figure 4: Effort per reward type over time in 345SUB

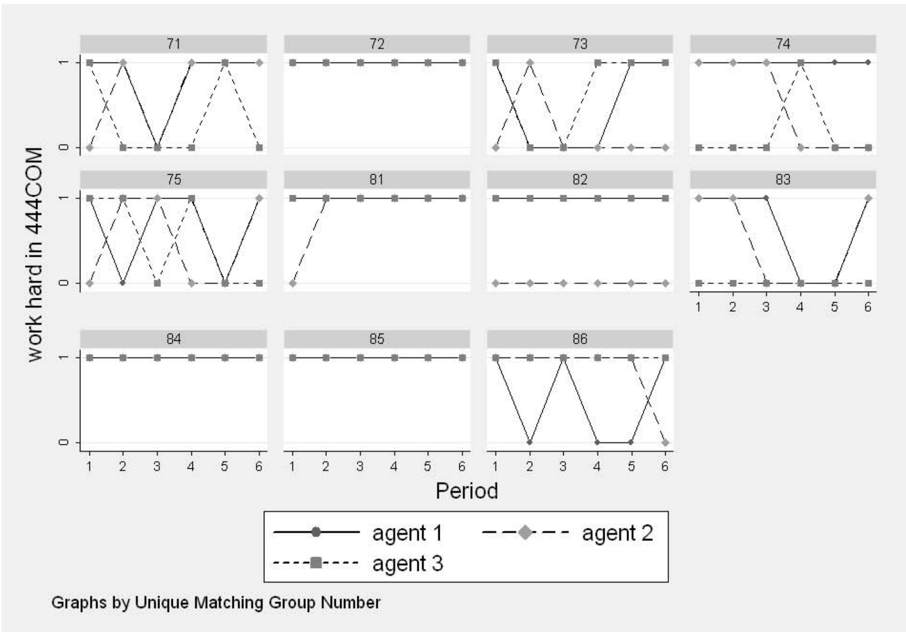


Figure 5: Effort per agent over time in 444COM

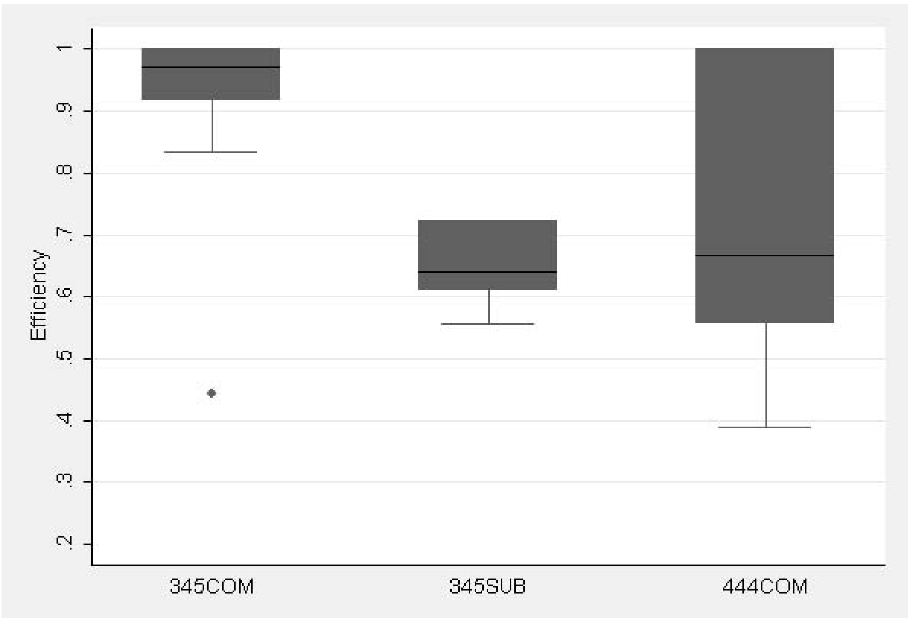


Figure 6: Boxplots of average group efficiency rates