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

Sebastian J. Goerg *Max Planck Institute for Research on Collective Goods*

Sebastian Kube *University of Bonn*

Ro'i Zultan *Max Planck Institute for Economics*

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 schemes 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.

1 Introduction

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

We thank seminar participants at the University of Bonn, Mannheim and Cologne, and in particular Gary Charness, Christoph Engel, Oliver Guertler, Bernd Irlenbusch, Matthias Kraekel, Stefanie Aniela Lehmann, Sandra Maximiano, Reinhard Selten, and Eyal Winter for helpful discussions. We also thank Darius Wieczorek for programming parts of the experiment, and Brian Cooper for proofreading. Financial support from the Fritz Thyssen Foundation (Az. 10.06.1.118) and from the Minerva Foundation is gratefully acknowledged. Contact the corresponding author, Sebastian Kube, at kube(at)coll.mpg.de.

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.[1] 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.[2]

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

[1] 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).

[2] The existing literature on team production and teamwork, e.g., Alchian and Demsetz (1972), Nalbantian and Schotter (1997) or Irlenbusch and Ruchala (2008), 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.

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 ($k \in [0, ..., n-1]$).[3]

In the following, a reward vector b (i.e. a reward mechanism) 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!

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

[3]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.

identical. In the experiment, three players work on a joint project and exert costly efforts. Their total sum of effort determines the number of some goods produced by the joint project for a given production function. The payoff of a player is given by the productivity (i.e., the number of produced goods) multiplied by an individual reward, minus the cost of effort. We create four different treatments by manipulating 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).

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'.[4] 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.[5] Second, the existing studies are mainly conducted in an

[4]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.

[5]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. (forthcoming) demonstrate

incomplete-contract framework where effort and/or wage is non-contractible, while we allow for complete contracts.[6] Third, the usual experimental setup features a principal who can set wages anew in each round, but this introduces uncontrolled elements of intentionality and reputation. Agents can 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.[7] 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.

2 The Experiment

Experimental Design Exact control over players' risk-preferences and over the underlying cost- and production-functions is crucial for testing Winter's theoretical model. We therefore use a deterministic representation of his model for the experiment.[8] In our game we have three agents working on a joint project. Each agent i individually decides whether to work (effort level $e_i = 1$) or shirk (effort level $e_i = 0$), and the individual cost of exerting effort is 90 Taler (our experimental currency). The individual payoff of agent i is

that paying equal wages to workers exerting different efforts leads to a strong decline in efficiency over time.

[6] 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.

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

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

given by the total number of goods produced multiplied by agent's individual reward per unit produced, minus his effort costs. 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**:

Table 1: 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 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 scheme** or remuneration scheme, which we vary across treatments. The reward scheme in treatments 444COM and 444SUB is symmetric. Each agent in the group receives a reward of 4 Taler per produced unit. Contrarily, the mechanism implemented in treatments 345COM and 345SUB is a discriminating one: agents' reward per produced unit 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 scheme, and $3 \cdot 20 + 4 \cdot 20 + 5 \cdot 20 = 240$ under the *discriminating* reward scheme.

Implementation Our experiment was conducted in a labor market framing, avoiding loaded terms (e.g., 'shirk' or 'success'). We used the same procedure in each treatment condition. 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.[9] In case that their belief fully matched the actual behavior, subjects were paid an additional 20 Taler. Only then we announced that five additional rounds of the game would follow, in which everything was kept constant (individual rewards, costs, production function and group composition). This was done to allow for possible learning to take place.[10] After our experiment, subjects had to complete a social-value orientation test[11] and a socio-economic questionnaire.

The computerized experiments were run in 2007 and 2009 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 rounds) or 36 independent decisions (only first round) per treatment.[12] A session lasted approximately 70 minutes.

[9]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 and 444SUB we also asked separately for the behavior of the two other players in the group.

[10]We acknowledge that repeated play may have promoted reciprocal strategies, so that the results might have been different under random rematching. However, as will be shown in the next section, the results do not support this conjecture, as they are qualitatively in line with the baseline model, and remain stable following the announcement of the additional rounds

[11]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.

[12]Subjects were recruited using ORSEE by Greiner (2003). 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 (2007). Sessions for treatment 444SUB were added during the revision process in 2009. 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.

Subjects were paid for their decision and their belief in the first round, and additionally for one randomly selected round (which was constant across all subjects within a session) out of the subsequent five rounds. Talers earned in the experiment were converted at a rate of 80 Taler = 1 Euro. Subjects received a show-up fee of 4 Euros and earned on average approx. 7 Euro in the main experiment.

2.1 Behavioral Predictions

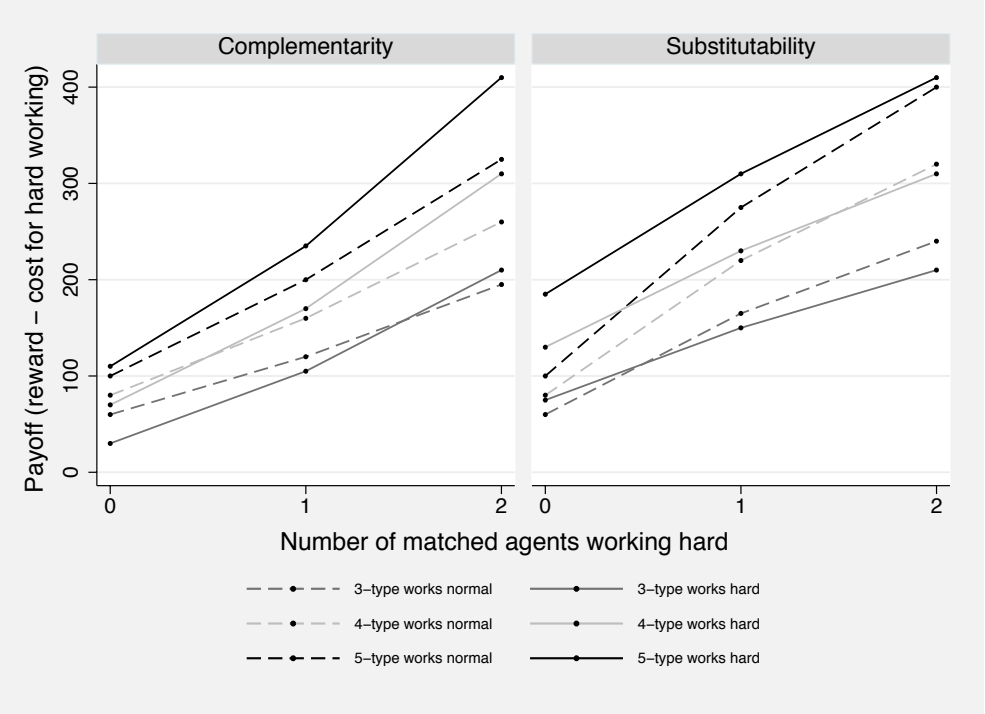
Figure 1 shows players' payoffs as a function of his reward-type and his decision. As can be seen, the reward per unit produced that is needed to make an agent just indifferent between working and shirking depends on the (belief about the) decisions of the other two players in the group. Let X_0 denote 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 an optimal 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 scheme. 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).

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 treatments now maximizes her payoff by working. Thus the equilibria are derived through repeated elimination of dominated strategies. The multiple pure equilibria of the egalitarian treatments are revealed by the crossover of the payoff functions of the 4-type player under the corresponding production function. Note that the gain from working can be seen to increase (diminish) under complementarity (substitutability).

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

In 444SUB, an agent receives a higher payoff from exerting effort if no

more than two of the other agents exert effort (since $X_2 > 4 > X_1 > X_0$). Hence there are three asymmetric equilibria in pure strategies in which one of the agents shirks while the other two agents exert effort. As in 444COM, an additional (symmetric) mixed strategies equilibrium exists, in which the probability of shirking is approximately .22.

The predictions above crucially depend 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 scheme comes at some hidden costs which incentivize agents to shirk, even under an initially incentive-inducing mechanism! Slight equity preferences[13] are already enough to let the superiority of the discriminating rewards vanish in 345COM. If agents' loss of utility from 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.[14] 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 ($(260, 170, 170)$ instead of $(310, 310, 310)$); which is something that (not only) an inequality-averse subject would never prefer. Under substitutability, the effect of discrimination is rather robust to equity preferences. In 345SUB, a smaller

[13]Throughout the paper, equity preferences are defined over payoffs rather than effort levels (cp. Mohnen et al. (2008).

[14]The intuition behind this hypotheses can easily be seen if we reconceive above equilibrium 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.

number of agents is exerting effort in equilibrium as common envy increases. By contrast in 444SUB, the predictions crucially hinge on the exact shape of the assumed equity preferences.

The behavioral predictions are summarized in Table 2. It lists the possible equilibria in pure strategies for self-centered subjects in the first row, and the equilibria that might additionally emerge in the presence of equity-considerations in the second row.

Table 2: Treatment variations and equilibria

	Treatment			
	345COM	345SUB	444COM	444SUB
Production Function:	complementarity	substitutability	complementarity	substitutability
Reward scheme:	discriminating	discriminating	symmetric	symmetric
Equilibria:				
Self-centered	(1,1,1)	(0,1,1)	(1,1,1) (0,0,0)	(0,1,1) (1,0,1) (1,1,0)
Inequality-averse	(0,0,0) (0,0,1) (0,1,1)	(0,0,1)	*no additional equilibria*	(0,0,0) (0,0,1) (0,1,0) (1,0,0) (1,1,1)

3 Experimental Results

In this section, first we show that workers' behavior is indeed sensitive to the type of production function they face in their joint project. The unequal treatment of equals does not necessarily hamper full effort provision. 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. Nevertheless, as we finally show, signs of equity concerns are present in our data.

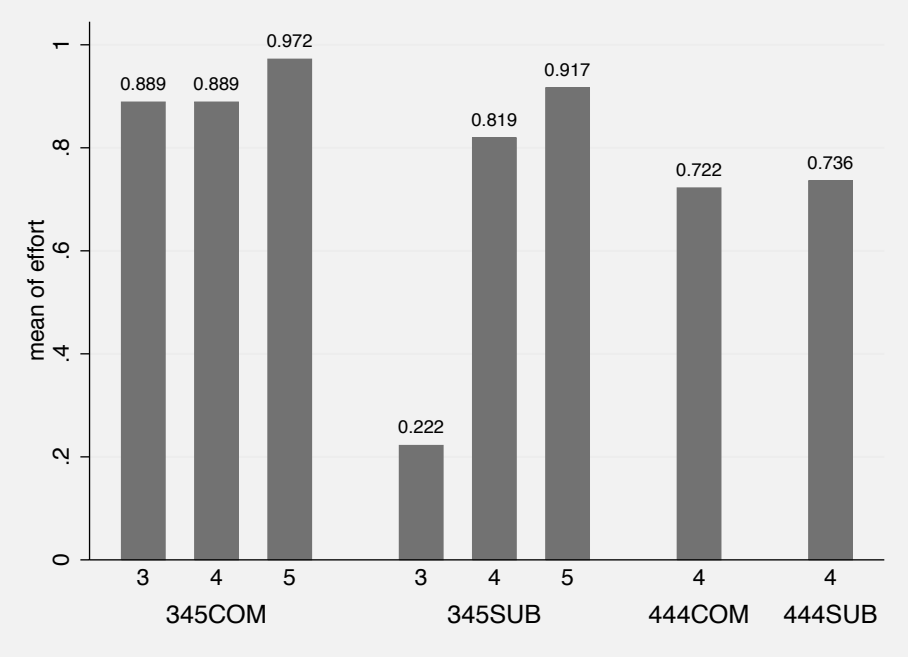


Figure 2: Mean effort per treatment

3.1 Sensitivity to the Production Function

Figure 2 shows mean effort levels over all rounds, conditional on players' reward type and treatment. Table 3 provides summary statistics and test results.

Focussing on the discriminating reward scheme, overall effort levels are significantly 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. In 345COM, 6 out of 12 groups exert full effort in all rounds (9/12 in all but one round), 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. 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%). It is also significantly lower than the effort level of the same type in 345COM (22.2% vs. 88.9%). Also in the first round, the number of low-reward players exerting effort is significantly higher in 345COM than in 345SUB (16.7%

Table 3: Summary statistics and results of statistical comparisons

	Treatment			
	345COM	345SUB	444COM	444SUB
A. Summary statistic:				
Mean effort level round 1	88.9%	66.7%	78.8%	66.7%
Mean effort level rounds 1-5	91.7%	65.3%	72.2%	73.6%
SD round 1	.1641	.1421	.1681	.3482
SD rounds 1-5	.1580	.1421	.1681	.3482
B. Statistical comparison round 1 (p-values Fisher's exact)				
345COM		**.0230**	.2080	**.0230**
345SUB			.1960	.5990
444COM				.1960
C. Statistical comparison rounds 1-5 (p-values, rank-sum test)				
345COM		**.0004**	**.0649**	**.0309**
345SUB			.6850	.4410
444COM				1
D. Mean effort level per reward type:				
low-reward type (3)	88.9%	22.3%		
medium-reward type (4)	88.9%	81.9%	72.2%	73.6%
high-reward type (5)	97.2%	91.7%		
E. Comparison across treatments within reward type (p-values, rank-sum test)				
low-reward type (3)				
345COM		**.001**		
medium-reward type (4)				
345COM		.2058	**.0260**	**.0127**
345SUB			.2526	.2508
444COM				1
high-reward type (5)				
345COM		.3202		
F. Comparison within treatment across reward types (p-values, sign-rank test)				
low vs. medium (3 vs. 4)	.3930	**.0074**		
low vs. high (3 vs. 5)	**.0261**	**.0039**		
medium vs. high (4 vs. 5)	.1577	.1248		

NOTE: All reported p-values are two-sided. A: SD is given over the mean frequencies of work per matching group. B: Effort level of each subject is one independent observation. C: Mean effort level of each matching group is one independent observation. E: Mean effort level of each player is one independent observation. F: Players with different reward types in one matching group are treated as depended observations.

vs. 75%).[15]

The effort levels of the medium- and high-reward types in 345COM (88.9% and 97.2%) do not differ significantly from the corresponding levels in 345SUB. Overall, when standard equilibrium predicts effort exertion, more than 80% do so. In the one case (low-reward type in 345SUB) in which the equilibrium strategy is to shirk, almost 80% of the decisions are to shirk (cp. Table 3-D).

Subjects' individual beliefs are in line with these findings. 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.[16] 93% of the decisions in 345COM and 77% of the decisions in 345SUB are best responses to stated beliefs.

Result 1: *In line with Winter's model, 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

Given a production function of complementarity and keeping the total cost of the reward scheme constant, the mean efficiency in round one is lower under the symmetric reward scheme (78.9%) than under the discriminating one (88.9%). Over the course of the experiment, the difference grows larger and becomes significant (72.2% vs. 91.7%).[17] On average, every reward type exerts more effort in the discriminating than in the symmetric treatment. Only 3 out of 11 groups exert full effort in all rounds, compared to 6/12 groups in 345COM (4/11 vs. 9/12 in all but one round). Moreover, the standard deviation of group efficiencies is significantly higher in 444COM

[15]Fisher's exact test $p = .012$. 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 (two-sided sign-rank test in 345COM $p = .75$; and in 345SUB $p = .37$). For further details compare the corresponding time-series data of Figures 3 and 4 in the appendix.

[16]Two-sided rank-sum test, medium-reward: $p = .0004$; high-reward: $p = .0007$

[17]Note that although the difference becomes significant over time, we find no significant time trend in 444COM. Comparing the mean effort level of work per matching group between the first and the last round reveals no significant difference (two-sided sign-rank test, $p = .53$; cp. also the time-series data in Figure 5).

than in 345COM.[18]

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 scheme under complementarity 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). 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.[19] In 444COM, this translates into low efficiency rates and a high variance of group efficiencies, suggesting that strategic considerations shaped by the reward scheme are crucial, and outweigh possible equity preferences of the subjects.[20]

The asymmetry of the reward scheme facilitates coordination among the agents under a production of complementarity. In case of 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 scheme make it hard for the subjects to form beliefs about the action of the other players, so that they are all in the dark.

Also under a production function of substitutability, the symmetric reward scheme yields a higher degree of strategic uncertainty than the discriminating reward scheme. This is reflected by our data. In 40% of the rounds, at least one team member changes its effort level in 444SUB (32% in 345SUB). The observed standard deviation of group efficiencies is significantly higher in 444SUB than in 345SUB.[21] However, all of the pure equilibria in 444SUB

[18]SD in 444COM (0.233) vs. SD in 345COM (0.158), Conover's squared-ranks test: $p = .0145$. 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 7).

[19]two-sided rank-sum test, $p = .0979$

[20]Note that strategic uncertainty might also be present in 345COM, because 'all-work' *and* 'all-shirk' are potential equilibria once we allow for equity considerations. Yet, we observe almost full efficiency in this treatment. One might consider that the result may be driven by a difference in the subject population between treatments. However, a comparisons of the corresponding results of the social-value orientation test reveals no significant differences between treatments (Kruskal-Wallis test, $p = .19$).

[21]SD in 444SUB (0.215) vs. SD in 345SUB (0.068), Conover's squared-ranks test: $p =$

require exactly two out of three agents to exert effort, which is the same as in the unique equilibrium in 345SUB. Therefore, the lack of coordination should not decrease efficiency. In fact, in both treatments we observe exactly two-thirds of the subjects exerting effort in the first round. The proportion remains fairly stable in 345SUB throughout the experiment (65.3%) and increases slightly in 444SUB (73.6%).[22]

Result 2: *Treating equals equally is neither a necessary nor a sufficient prerequisite for eliciting high performance in teams. Asymmetry facilitates coordination and increases efficiency under a production function of complementarity. The possible benefit of a discriminating reward scheme on efficiency levels strongly interacts with the production function — which is again in line with Winter's model.*

3.3 Equity Concerns

Although we saw above that unequal rewards can potentially increase productivity, let us point out that this is not to say that equity considerations are absent in our experiment. For example, the average rate of effort provision over all rounds in 345COM is significantly different between the low- and the high-reward type (88.9% vs. 97.2%). This might reflect a slight reluctance of the low-reward players to work because the others then earn more than himself. Yet, in 5/8 instances where the low-reward player shirks, the behavior might also be explained by self-centered preferences, because subjects play a best response given their individual belief (in total, 93% of the decisions in 345COM are best responses to the stated beliefs). Speaking of beliefs, in 345COM also the beliefs of medium- and high-reward players about the low-reward player's decision in the first round reveal some influence of equity concerns, because 42% (wrongly) expect him to shirk.

Also in 444COM, we observe signs of equity concerns. Players are very likely to exert effort if they expect both other players to work as well (in this situation 89% of the decisions were to work). This might just be playing the

.0011

[22]The slight increase may be due to subjects trying to coordinate on one of the pure equilibria in the first round, but later approaching the symmetric mixed equilibrium through learning. Nevertheless, no significant time trend is observed in 444SUB. Comparing the mean effort level per matching group between the first and the last round reveals no significant difference (two-sided sign-rank test, $p = .87$; cp. also the time-series data in Figure 6). As noted earlier, the same holds true for 345SUB.

best reply on their stated belief (in total, 79% of the decisions in 444COM are best responses to the stated beliefs). But it might also be because they do not want to increase the payoff inequality in their group by shirking.

Exploring the same situation in 444SUB, we can actually distinguish between the two reasons because now they do not coincide. Here, if a player expects both the others to work, the best reply would be to shirk. Nonetheless, equity concerns might make him want to work so that players' payoffs are equitable. What we observe is that in this case only 18% of the decisions are a best reply to the stated beliefs (in the other cases in 444SUB, the rate is actually 55%) — which again is indicative of equity concerns. However, it is important to note that the discrepancy between beliefs and behavior could also result from rational selfish money-maximizing preferences. In the mixed equilibrium, agents exert effort with high probability. Thus, when asked for a point belief, expecting others to exert effort becomes optimal, which leads to behavior similar to what we observe.

Finally, the data from the social-value orientation test indicate that subjects have a general preference for equitable outcomes, because in all treatments the value-orientations do differ significantly from being strictly self-centered (two-sided rank-sum test, 345COM: $p \leq .001$; 345SUB: $p \leq .001$; 444COM: $p \leq .001$; 444SUB: $p \leq .001$).

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 scheme. Under a production technology of complementarity, the use of a symmetric reward scheme elicits substantially lower efforts and efficiency than a cost-equivalent discriminating reward scheme. The same discriminating reward scheme underperforms when it is utilized under a production function of substitutability.

Our findings have important implications for the design of organizations in practice. First, they clearly point to the relevance of the production func-

tion for organization construction – a factor which has so far received little attention in the literature.[23] 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 implications of our results can be extended beyond the labor context to additional environments in which unequal treatment of equal is shown to be efficient. Relevant applications include differential tax rates (Atkinson &

[23]The role of externalities between coworkers was studies by Gould and Winter (2009), who show that professional baseball players react to the technology in a way consistent with a related model to the one we study.

Stiglitz, 1976) and various trade contexts in which a 'divide and conquer' strategy maximizes gains (see Segal, 2003, and references therein). Thus, equity preferences may hinder the dynamics assumed in different domains. Our paper presents a step forward in understanding the boundaries of equity considerations and their potential implications.

The results in this paper should not be taken as arguments against the importance of fairness considerations in general. For instance, they might be partially explainable by models incorporating social efficiency (e.g. Charness & Rabin, 2002). Still, our findings suggest that equal treatment of equals is neither a necessary nor a sufficient prerequisite for eliciting high performance in teams. 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 schemes. Future research could try to exacerbate the differences in payoffs in order to estimate some kind of metric for the strength of inequity preferences in our setting (we thank an anonymous referee for making this suggestion). 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.

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. forthcoming. When Equality is Unfair. *Journal of the European Economic Association.*

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, no. 5:777-795.

Atkinson, Anthony B., and Joseph E. Stiglitz. 1976. The Design of Tax Structure: Direct versus Indirect Taxation. *Journal of Public Economics* 6:55-75.

Barron, John E., and Kathy Paulson Gjerde. 1997. Peer Pressure in an Agency Relationship. *Journal of Labor Economics*15, no. 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, no. 1:269-299.

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

Charness, Gary, and Matthew Rabin. 2002. Understanding Social Preferences with Simple Tests. *The Quarterly Journal of Economics* 117, no. 3:817-869.

Falk, Armin, and Andrea Ichino. 2006. Clean Evidence on Peer Effects. *Journal of Labor Economics* 24, no. 1:39-57.

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, no. 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. 2007. Zurich toolbox for ready-made economic experiments. *Experimental Economics* 10, no. 2:171-178.

Gould, Eric D., and Eyal Winter. 2009. Interactions between Workers and the Technology of Production: Evidence from Professional Baseball *Review of Economics and Statistics* 91, no. 1:188-200.

Greiner, Ben. 2003. An Online Recruitment System for Economic Experiments. In: Kurt Kremer, Volker Macho: *Forschung und wissenschaftliches Rechnen 2003. GWDG Bericht 63*, 79-93, Goettingen : Ges. fuer Wiss. Datenverarbeitung.

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. 2008. Relative Rewards within Team-Based Compensation. *Labour Economics* 15:141-167.

Kandel, Eugene, and Edward Lazear. 1992. Peer Pressure and Partnerships. *Journal of Political Economy* 100, no. 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, no. 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, no. 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. 2009. Peers at Work. *American Economic Review* 99, no. 1:112-145.

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

Mohnen, Alvine, Kathrin Pokorny, and Dirk Sliwka. 2008. Transparency, Inequity Aversion, and the Dynamics of Peer Pressure in Teams: Theory and Evidence *Journal of Labor Economics* 26, no. 4:693-720.

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, no. 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, no. 3:747-767.

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.

Segal, Ilya R. 2003. Coordination and Discrimination in Contracting with Externalities: Divide and Conquer? *Journal of Economic Theory* 113:147-181.

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, no. 1:234-248.

Winter, Eyal. 2004. Incentives and Discrimination. *American Economic Review* 94, no. 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:

Ihre Auszahlungstabelle:

Für jede produzierte Einheit erhalten Sie A Taler, d.h.:

# der hart Arbeitenden (ohne Sie)	Sie arbeiten NORMAL	Sie arbeiten HART
0	10	11
1	12	13
2	14	15

Auszahlungstabelle Tln. 2:

Für jede produzierte Einheit erhält er B Taler, d.h.:

# der hart Arbeitenden (ohne ihn)	Er arbeitet NORMAL	Er arbeitet HART
0	16	17
1	18	19
2	20	21

Berechnung für den Fall, dass Tln. 2 [normal] / [hart] arbeitet

Auszahlungstabelle Tln. 3:

Für jede produzierte Einheit erhält er C Taler, d.h.:

# der hart Arbeitenden (ohne ihn)	Er arbeitet NORMAL	Er arbeitet HART
0	22	23
1	24	25
2	26	27

Berechnung für den Fall, dass Tln. 3 [normal] / [hart] arbeitet

Anzahl der produzierten Einheiten, wenn...

0 hart arbeiten:	**20**
1 hart arbeitet:	**40**
2 hart arbeiten:	**65**
3 hart arbeiten:	**100**

Die möglichen Auszahlungen für den Fall, dass Tln. 2 tatsächlich **NORMAL** und Tln. 3 tatsächlich **HART** arbeitet, betragen:

Wenn Sie..	Ihre Auszahlung	Auszahlung Tln. 2	Auszahlung Tln. 3
[NORMAL arbeiten]	12	18	24
[HART arbeiten]	13	20	26

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:

Ihre Auszahlungstabelle:

Für jede produzierte Einheit erhalten Sie A Taler, d.h.:

# der hart Arbeitenden (ohne Sie)	Sie arbeiten NORMAL	Sie arbeiten HART
0	10	11
1	12	13
2	14	15

Auszahlungstabelle Tln. 2:

Für jede produzierte Einheit erhält er B Taler, d.h.:

# der hart Arbeitenden (ohne ihn)	Er arbeitet NORMAL	Er arbeitet HART
0	16	17
1	18	19
2	20	21

Berechnung für den Fall, dass Tln. 2 [normal] / [hart] arbeitet

Auszahlungstabelle Tln. 3:

Für jede produzierte Einheit erhält er C Taler, d.h.:

# der hart Arbeitenden (ohne ihn)	Er arbeitet NORMAL	Er arbeitet HART
0	22	23
1	24	25
2	26	27

Berechnung für den Fall, dass Tln. 3 [normal] / [hart] arbeitet

Anzahl der produzierten Einheiten, wenn...

0 hart arbeiten:	**20**
1 hart arbeitet:	**40**
2 hart arbeiten:	**65**
3 hart arbeiten:	**100**

Die möglichen Auszahlungen für den Fall, dass Tln. 2 tatsächlich **NORMAL** und Tln. 3 tatsächlich **HART** arbeitet, betragen:

Wenn Sie..	Ihre Auszahlung	Auszahlung Tln. 2	Auszahlung Tln. 3
[NORMAL arbeiten]	12	18	24
[HART arbeiten]	13	20	26

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 round.

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 rounds, you again have to choose between working normal or working hard. In the end, we randomly select one of these five rounds. You will receive the payoff for the randomly selected round 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

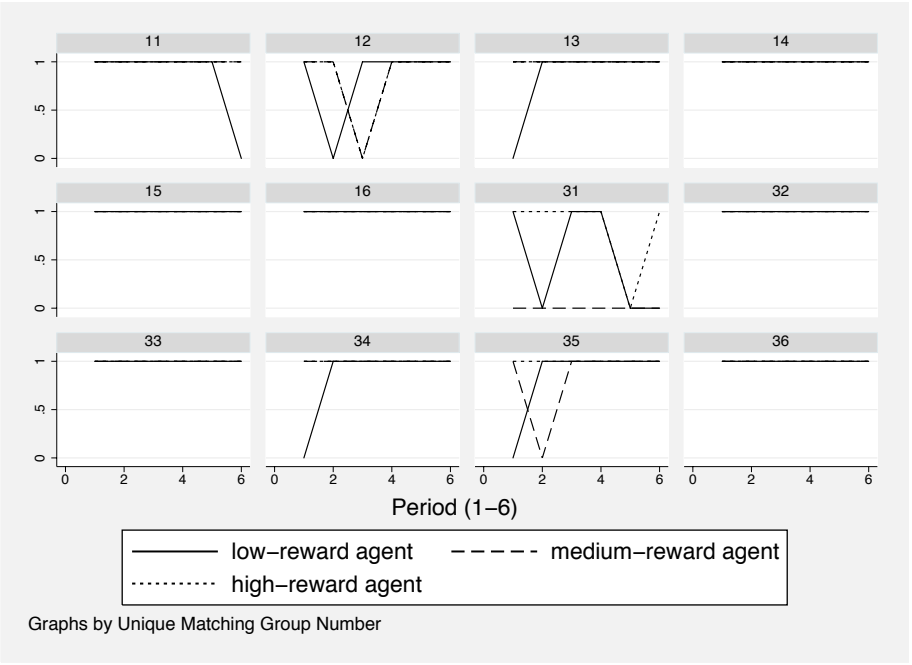


Figure 3: Effort per reward type over time in 345COM

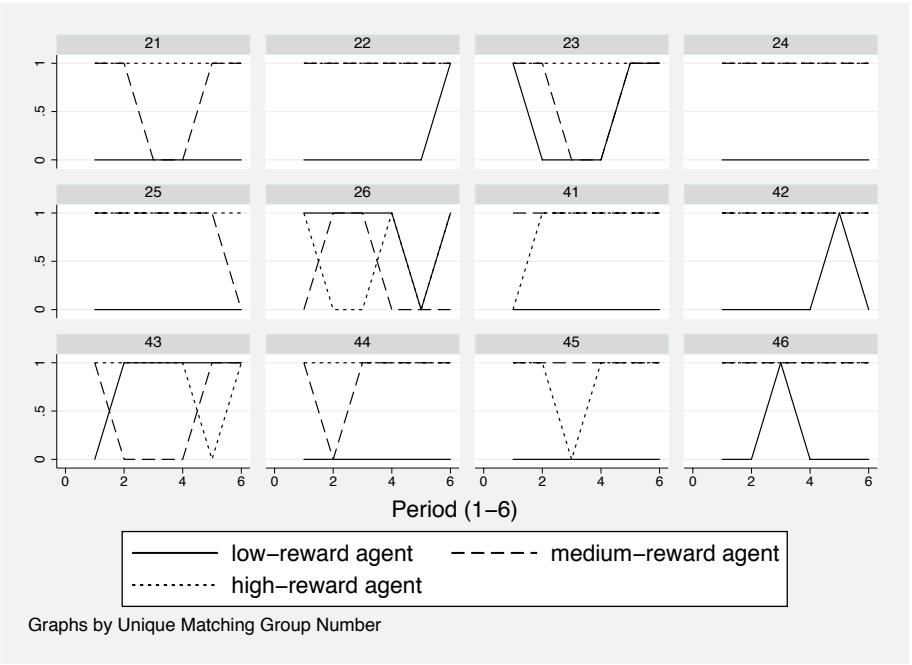


Figure 4: Effort per reward type over time in 345SUB

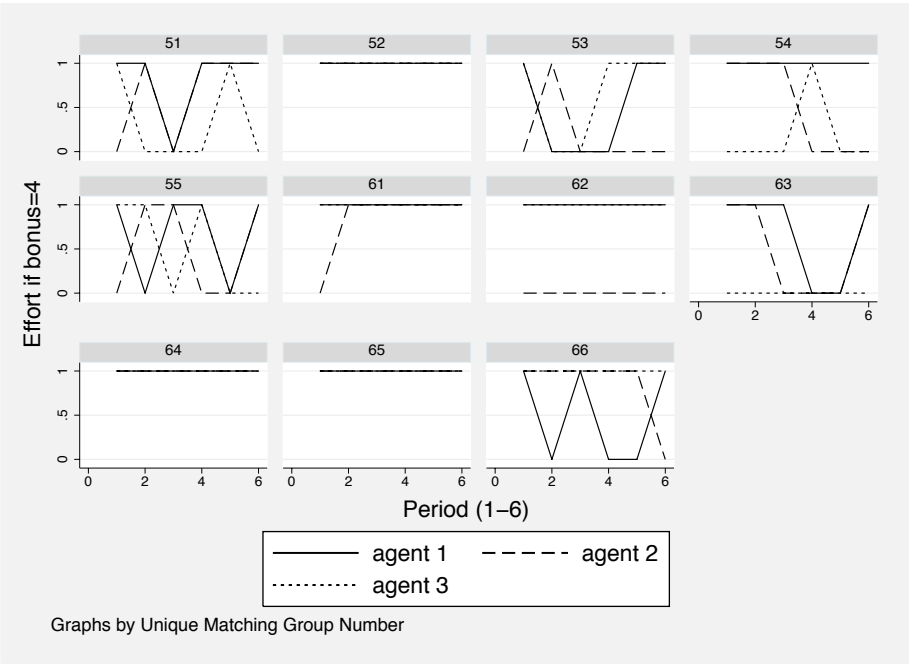


Figure 5: Effort per agent over time in 444COM

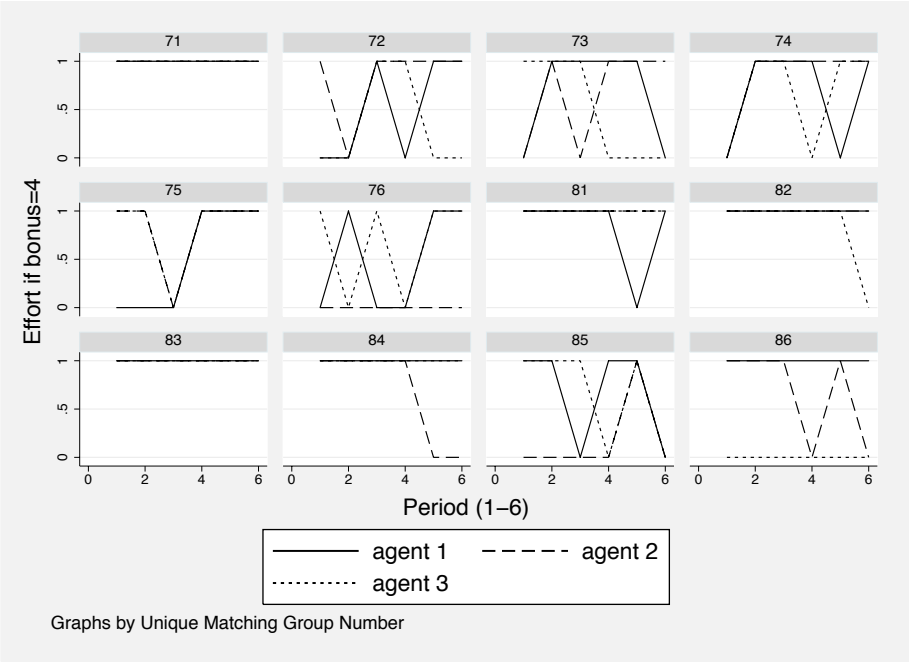


Figure 6: Effort per agent over time in 444SUB

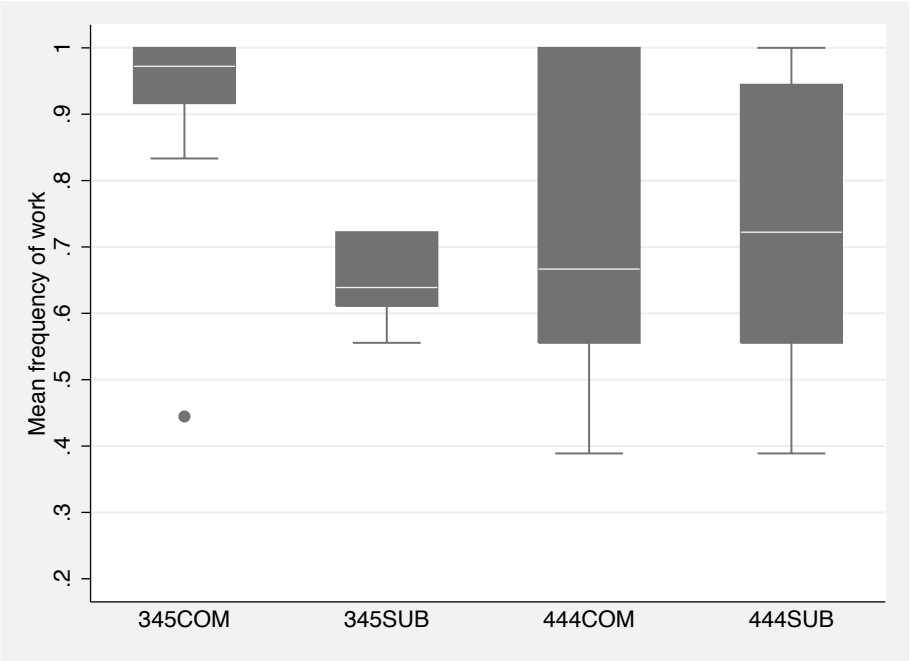


Figure 7: Boxplots of average group efficiency rates