Essays on Social Preferences, Incentives, and Institutions

Inaugural-Dissertation
zur Erlangung des Grades eines Doktors
der Wirtschafts- und Gesellschaftswissenschaften
durch die
Rechts- und Staatswissenschaftliche Fakultät
der Rheinischen Friedrich-Wilhelms-Universität
Bonn

vorgelegt von
Matthias Wibral
aus Olpe

Bonn 2009

Dekan:	Prof. Dr. Christian Hillgruber
Erstreferent:	Prof. Dr. Armin Falk
Zweitreferent:	Prof. Dr. Matthias Kraekel

Tag der mündlichen Prüfung: 08.09.2009

Diese Dissertation ist auf dem Hochschulschriftenserver der ULB Bonn
(http://hss.ulb.uni-bonn.de/diss_online) elektronisch publiziert.

para mi guerillera -

siempre sigue luchando

Acknowledgments

Armin Falk has been a perfect supervisor. I hope that his enthusiasm for research, his creativity, and his amazing ability to transmit both have found their way into the following pages.

Paul Heidhues, Matthias Kräkel and other faculty members of the Bonn Graduate School of Economics provided very helpful comments and support at all stages of this dissertation. Urs Schweizer and Georg Noeldecke have deeply impressed me with their relentless effort to keep the Bonn Graduate School of Economics running and to improve it.

Work is easy and fun when you work with friends. I am very grateful to Sebastian Kube, Felix Marklein, and especially Johannes Abeler and Steffen Altmann for their advice, patient help, and countless bad jokes and good laughs. I was also very lucky to have the classmates that I had. The BGSE class of 2003 has been a perfect social storm. Many other fellow graduate students also deserve credit for making the BGSE a fun and inspiring place to do research. All members of the chair of Armin Falk provided great help and support in the ”final” stages of this dissertation.

Several sources have prevented that the old saying ”A dissertation can last longer than you can stay liquid” became true for me. Financial support from the DFG through GRK 629, the BGSE, the Institute for Empirical Research in Economics at the University of Zurich, and the DAAD is gratefully acknowledged. During parts of this dissertation I enjoyed the hospitality of the Universitat Pompeu Fabra, the Institute for Empirical Research in Economics at the University of Zurich, and the Institue for the Study of Labor and I would like to thank all the people who made these places pleasant.

I am indebted to my friends who bore with me through the ups and downs of the past years, in particular Peter Flore, Daniel Hilgers and Harald Kaluza. I am also deeply grateful to my family for their support and love. My mom deserves credit for inspiring my interest in social preferences by being the best living example of them. My dad has been an inexhaustible source of knowledge, in particular on the intricacies of philosophy and the English language. Annette and Angel have always provided a different perspective and a home away from home. I am grateful to Ute for descending on me with her exploding backpack and taking my mind off things, and to Michael for taking my mind on things.

Finally, one person has accompanied this dissertation from its very beginning and made all the difference. Without Suni's love, tireless encouragement, unconditional support and gentle stubbornness this dissertation would still await its completion.

Contents

List of Figures

List of Tables

Introduction

Behavioral economics incorporates non-pecuniary motives and bounded rationality into economic models in order to better understand human behavior. Over the past three decades this effort has spawned a substantial literature and contributed to a better understanding of a wide variety of economic phenomena. This diversity is also reflected in this dissertation. The topics of the different chapters—the intrapersonal relationship of trust and reciprocity, the impact of different wage schemes on effort, the efficiency of online reputation systems, and the behavioral effects of the structure of promotion tournaments—do not seem to have much in common at first sight. Nevertheless, all of them share several essential features. First, all chapters aim at shedding light on different aspects of the same underlying question. How do non-pecuniary motives affect human behavior and the efficiency of institutions? Second, they all study situations in which interactions are not governed by explicit, fully contingent contracts. Finally, all chapters use laboratory experiments to address their research question. In the following paragraphs I will briefly summarize the main findings of each chapter in the context of the respective literature and, in passing by, show the links between them.

A major strand of the behavioral economics literature studies the nature of social preferences such as altruism (Andreoni and Miller 2002), inequity aversion (Fehr and Schmidt 1999, Bolton and Ockenfels 2000), or reciprocity (Rabin 1993, Falk and Fischbacher 2006). An important first step is to establish empirical regularities which formal models of social preferences can be built on. In a second step the explanatory power and the generality of these theories are tested with field or laboratory data. Laboratory experiments have been particularly successful in this process since clean tests of theories often require a high degree of control over the decision environment.[1]

Chapter 1 contributes to this literature by analyzing the intrapersonal relationship between trust and reciprocity.[2] Both phenomena have been studied extensively since they are generally considered as conducive to economic performance and efficiency (e.g. Knack and Keefer 1997). The fact that trust varies so much across countries and individuals seems particularly puzzling. Various factors have been discussed as potential determinants of trust at an individual level including gender (Buchan et al. 2008), betrayal aversion (Bohnet and Zeckhauser 2004) or risk attitudes (Dohmen et al. 2006). In this chapter we study to which degree a person's inclination to reciprocate explains her own trusting behavior. While this question is interesting in its own right, it is also interesting from a theory testing perspective because prominent theories of social preferences implicitly assume a negative correlation between trust and reciprocity. We therefore ask whether the reciprocal really trust less than the selfish.

A common workhorse in research on trust and reciprocity is the trust game (Berg et al. 1995) or variants thereof. The trust game is played by two players, a trustor and a trustee. In the first stage the trustor can send a share or all of his endowment to the trustee. On the way to the trustee this amount is multiplied with a factor greater than 1 by the experimenter, i.e., sending money leads to efficiency gains. In a second stage the trustee can then decide how much (if anything at all) of the amount he received he wants to send back to the trustor. The amount sent by the trustor is usually taken as a measure of trust, since he cannot be sure that the trustee will send anything back, i.e. there is potential for moral hazard. The (relative) amount returned by the trustee is used as a measure of the latter's reciprocity or trustworthiness, since any amount which he sends back directly reduces his own payoff. The basic structure of this game, i.e., a first mover who can generate efficiency gains but may be reluctant to do so because of second mover moral hazard, is used in chapters 1-3 of this dissertation.

[1]A good example of this is the experimental work on the role of intentions in theories of social preferences (e.g. Falk et al. 2003 and McCabe et al. 2003).

[2]Chapter 1 is based on joint work with Steffen Altmann and Thomas Dohmen (Altmann et al. 2008).

To address the question whether the reciprocal trust less we employ a variant of the trust game that allows us to measure both variables for each individual. We find a strong and positive relationship between a person's reciprocity and his trusting behavior. Reciprocal players exhibit much higher levels of trust than more selfish ones, even when personal characteristics and preferences such as gender or risk attitudes are controlled for. Our results show that the relationship between trust and reciprocity may be more complex than captured by most models. In chapter 1 we discuss how existing models could be modified to accommodate our findings.

The basic moral hazard problem inherent in the trust game described above can be observed in many economic interactions. Examples include credit markets (Brown and Zehnder 2008) or online trade (chapter 3). Arguably the most important application, however, is the labor market. Labor contracts are often incomplete, for example, because it is difficult to exactly specify the tasks of a particular job or because work effort may be observable but not verifiable in court. A large literature demonstrates the potential of "gift exchange" to mitigate moral-hazard problems in the labor market: since many agents repay a gift in the form of higher wages by providing higher efforts, effort can be elicited under incomplete contracts even in one-shot situations where no future gains can be expected (e.g., Akerlof 1982, Fehr et al. 1997, Maximiano et al. 2007). A prerequisite for gift exchange is that wages are perceived as fair. In a bilateral setup, fairness perceptions mainly depend on the absolute amount paid. However, in a firm with more than one employee, horizontal fairness concerns between the workers could play an important role. The potential of gift exchange as a contract enforcement device therefore is likely to depend on the mode of payment. A key question in this context is how to treat agents relative to each other. In chapter 2, we study this question by focusing on two important fairness principles: horizontal equality and equity.[3]

On the one hand, it has been argued that *horizontal equality* is crucial for a wage scheme to be considered as fair. Differential pay of co-workers could cause

[3]Chapter 2 was developed jointly with Johannes Abeler, Steffen Altmann, and Sebastian Kube. An earlier version of this chapter was circulated under the title "Reciprocity and Payment Schemes: When Equality is Unfair".

resentment and envy within the workforce, and ultimately lower performance (e.g., Pfeffer and Langton 1993, Bewley 1999). If workers care foremost about equality, a wage scheme that guarantees equal wages for co-workers should lead to an efficiency-enhancing gift-exchange relation. On the other hand, the importance of the *equity principle* has long been discussed in social psychology, personnel management, and economics (e.g., Homans 1961, Fehr and Schmidt 1999, Konow 2003). In a work environment, the equity principle demands that a person who exerts higher effort should receive a higher wage compared to his co-worker. Whenever workers differ in their performance, horizontal wage equality violates the equity principle since a higher effort is not rewarded with a higher wage.

In chapter 2 we use a laboratory experiment to analyze the interaction between the institution of wage equality and gift exchange. In the experiment, one principal is matched with two agents. In a first stage the agents exert costly effort. After observing their efforts, the principal pays them a wage. In the main treatment he can choose the level of the wage but he is obliged to pay the same wage to both agents (*equal wage treatment* or EWT). In the control treatment, the principal can wage discriminate between the two agents (*individual wage treatment* or IWT). In both treatments, neither efforts nor wages are contractible.

Our main findings underline the importance of the equity principle. Performance differs substantially between the EWT and the IWT: agents who are paid equal wages exert significantly lower efforts than agents who are paid individually. The strong treatment effect cannot be explained by differences in monetary incentives. The actual wage choices of principals imply that providing high effort levels is profitable for agents in both treatments. Instead, we show that the frequent violation of the equity principle in the equal wage treatment can explain the effort differences between the treatments. In both treatments, agents who exert a higher effort and earn a lower payoff than their co-worker strongly decrease their effort in the next period. However, the norm of equity is violated much more frequently under equal wages. Principals in the IWT understand the mechanisms of equity quite well. When efforts differ they do pay different wages, rewarding the harder-working agent with a higher payoff in most cases.

Our results suggest a psychological rationale for using individual wages. Subjects perceive equal wages for unequal performance as unfair and reduce their effort subsequently. The traditional literature on incentive provision in groups comes to a similar conclusion though for a different reason. It is usually argued that the inefficiency of equal wages stems from the fact that marginal products and wages are not aligned. This can lead to free-riding among selfish agents (e.g., Holmström 1982, Erev et al. 1993). We enlarge the scope of this critical view on wage equality: interestingly, in our setup it is precisely the presence of fair-minded agents and not their absence that calls for the use of individual rewards.

Regarding compensation practice in firms, our findings highlight the importance of taking the concerns for co-workers' wages into account. However, doing so by paying equal wages to a group of agents may actually do more harm than good. As soon as agents differ in their performance, equal wages which seem to be a fair institution at first sight might be considered very unfair. While the discouraging effect of equal wages on hard-working agents has long been informally discussed (e.g., Milgrom and Roberts 1992, p.418f) this chapter provides controlled evidence in favor of this intuition. Moreover, it suggests that it is the violation of the norm of equity that causes the discouragement and low performance.

Chapter 2 shows that laboratory experiments can also be used to study the performance of different institutions. In recent years economists have successfully employed experiments to "engineer" markets and institutions, i.e., to improve their design and functioning (Roth 2002). Importantly, the laboratory allows causal inferences about how institutions affect individual behavior and markets outcomes since institutions are imposed exogenously. Chapter 3 uses this feature to study the efficiency of online reputation systems when sellers can change their identity. In contrast to most face-to-face transactions, online trade is prone to severe moral hazard problems since it usually takes place between strangers, payment occurs before the good is shipped and legal enforcement of an agreement may be prohibitively costly. Sellers therefore have an incentive to ship a good of lower quality or not to ship at all. Reputation systems seek to address this problem by disseminating information about past conduct of a user to the whole community. In principle, a reputation profile thus allows buyers to distinguish honest sellers from dishonest ones and to interact only with the former. Online reputation, however, is only connected to the virtual identity of a person, i.e., the user name, and not the person itself. After a bad rating a dishonest seller can comparatively easily abandon his old virtual identity and create a new account under a new user name with no reputation attached to it.

At first sight identity changes seem to severely weaken the disciplinary power of a reputation system. Dishonest behavior can still be punished with a negative rating but these ratings may lose their edge when they can easily be shed by creating a new identity. However, it is difficult to determine whether identity changes are responsible for problems with fraud and dishonest behavior because identity changes are not directly observable in field data. In chapter 3, we therefore use a laboratory experiment to examine empirically how the option to change one's virtual identity affects the efficiency of markets and the performance of reputation systems in inducing trust and trustworthiness. We study two experimental markets in which buyers play a trust game (Berg et al. 1995) with varying sellers. Buyers can rate sellers after each transaction; and before deciding whether to send money to a seller, a buyer can see the rating profile of this seller. New players enter the market over time in both treatments. The only difference is that sellers in one market can change their identity (*change treatment*), i.e., erase their rating profile and start over as new players, while in the other market this is not possible (*no-change treatment*).

The main findings can be summarized as follows. The no-change treatment convincingly demonstrates the great potential of reputation systems. We observe high levels of trust and trustworthiness. In contrast to one-shot trust games without a reputation system buyers make a substantial profit on their investment. Opportunistic sellers, who return little or nothing to the buyer, do not make a higher profit than those returning significant amounts to the buyer. This is because the latter receive higher investments. With identity changes trust and trustworthiness decrease. Lower buyer trust is due to a greater incidence of buyers dropping out of the market and to lower trust in new sellers. Sellers' trustworthiness mainly decreases because a substantial group of sellers behave opportunistically and change their identity multiple times. Interestingly though, trustworthiness is still high enough to make investing profitable for the buyers even when rating profiles can be erased. Although ostracism by bad reputation is not feasible trustworthy sellers still have the opportunity to distinguish themselves by a positive reputation. Monetary and non-monetary rewards to a positive reputation seem to be deemed sufficiently high to prevent wide spread opportunistic behavior.

In the final chapter of this dissertation we return to the labor market. Again we study a setup without explicit pay-for-performance contracts. Instead, agents compete for a promotion to a higher position in the hierarchy. The prospect of being promoted to a better-paid job creates incentives to work hard, even if current income is not tied to performance. Promotion tournaments therefore play an important role for the provision of incentives in firms, in particular when measuring absolute performance is much more difficult than measuring relative performance. In most companies, or hierarchical organizations in general, those who get promoted compete again for subsequent promotions. The incentive effect of these multi-stage elimination tournaments depends on two important variables. Performance incentives are stronger the higher the prize spread, i.e., the higher the immediate wage increase for an agent who gets promoted. But in multi-level hierarchies, the promoted agents do not only receive the immediate wage increase but also gain the chance of being promoted even further. This option value of further competition is the second source of incentives and distinguishes multi-stage from simple one-stage tournaments.

Despite the undisputed practical importance of multi-stage elimination tournaments clean evidence on this second—dynamic—incentive effect is scarce. Chapter 4 aims at providing a step towards filling this lacuna. In particular, we ask whether people take future stages of a tournament into account when deciding on their current effort and whether multi-stage elimination setups are behaviorally different from simple one-stage tournaments. To answer these questions we first compare a two-stage elimination tournament (TS) with a strategically equivalent one-stage tournament (OS). In addition, we study how the wage structure in multi-stage tournaments affects the provision of effort by comparing two-stage tournaments with a different degree of convexity in their wage structures (*TS* and *TSC*).

In the two-stage tournaments (*TS* and *TSC*) four subjects compete for being promoted to the second stage. The two promoted subjects compete against each other again for the highest position. Promotion depends on individual (costly) effort and an individual noise term. *TSC* is identical to *TS* with the exception that the wage structure is more convex, i.e., the intermediate wage is smaller in *TSC*. In the one-stage tournament (*OS*) four subjects compete once for two top positions. Wages for the promoted subjects are chosen such that the one-stage tournament is strategically equivalent to the first stage of *TS*. This means that promoted subjects in *OS* earn the sum of the intermediate wage in *TS* *and* the monetary equivalent of the second-stage option value in that tournament, implying that equilibrium effort is the same in both treatments. Comparing *OS* and *TS* thus allows testing whether strategic equivalence translates into behavioral equivalence.

Our results show that this is not the case. While average behavior in the one-stage tournament is remarkably close to the predictions of tournament theory (and previous findings), behavior in the first stage of *TS* differs strongly both from the one-stage treatment and from theoretical predictions. Subjects exert significantly higher efforts in the first stage of the two-stage tournament. The results of the *TSC* treatment confirm the occurrence of excess effort provision in the first stage of the two-stage tournaments. Subjects react only weakly to the change in the wage structure, implying that first-stage excess effort is even higher in *TSC*. Behavior in both two-stage tournaments also shows that the basic logic of multi-stage elimination tournaments indeed works. People do not only respond to differences in prizes, or wages, but are also motivated by the option value entailed in future promotion possibilities.

Chapter 4 provides insights with regard to the question whether one-stage tournaments are behaviorally equivalent to multi-stage designs.[4] Adding more stages seems to make a fundamental difference, as people tend to exert excess effort in early stages of the tournament. This shows that one cannot necessarily draw infer-

[4]Chapter 4 is joint work with Steffen Altmann and Armin Falk and was circulated under the title "Promotions and Incentives: The Case of Multi-Stage Elimination Tournaments".

ences from simple one-stage setups to more complex tournaments. Our findings may ultimately also help to explain why firms rely so heavily on promotion based incentive schemes even if other means of compensation are available. The excess effort could make a tournament incentive scheme "cheaper" compared to other incentive schemes like piece-rates.

Chapter 1

Do the Reciprocal Trust Less?

1.1 Introduction

By now there seems to be broad agreement that trust and reciprocity are conducive to economic performance and efficiency (e.g., Knack and Keefer 1997). Mutual trust between trading parties facilitates the realization of gains from trade, for instance by reducing contracting costs. Reciprocity can also enhance performance in many areas of economic life, for example by mitigating moral hazard problems in labor relations (Fehr et al. 1997). In order to better understand the economic implications of reciprocity, several formal models of social preferences have been developed (e.g., Fehr and Schmidt 1999, Falk and Fischbacher 2006).

In spite of the importance of trust and reciprocity surprisingly little is known about their relationship on an intrapersonal level. In other words, do reciprocal persons trust more or less than selfish ones? In this chapter we address precisely this question with the help of a controlled laboratory experiment, employing a variant of the trust game that allows us to measure both variables for each individual.[1]

[1]Several studies (e.g., Cox 2004, Ashraf et al. 2006) have analyzed behavior across games in order to disentangle subjects' unconditional kindness or altruism from trust and reciprocity, but these studies do not look at the direct link between a person's (own) reciprocity and trust. In addition, it is not clear to what extent inferences can be made from behavior in non-strategic environments (e.g., the dictator game) to players' motives in strategic interactions. See Fehr and Schmidt (2006) for a discussion of this point.

We find a strong and positive relationship between a person's reciprocity and her trusting behavior. Reciprocal players exhibit much higher levels of trust than more selfish ones, even when personal characteristics and preferences such as gender or risk attitudes are controlled for. This finding is also interesting from a theoretical perspective because theories of social preferences typically assume—at least implicitly—a connection between trust and reciprocity. In particular, the observed positive relation between the two raises important questions about theories which predict that *ceteris paribus* "fairer" players trust *less*.

The remainder of this chapter is organized as follows. The next section describes the design of our experiment, Section 1.3 presents the empirical results. Section 1.4 concludes by discussing the implications of our findings for modelling social preferences.

1.2 Experimental Design

In our experiment, subjects were anonymously matched in pairs and played a modified version of the trust game (Berg et al. 1995). Both players received an endowment of 120 points. The first mover (the *sender*) could send any amount $t \in \{0, 20, 40, 60, 80, 100, 120\}$ to the second mover (the *receiver*). The amount sent was tripled by the experimenter. Then, the second mover could send back any amount between zero and 480 points. The crucial feature that distinguishes our design from the original version of the trust game is the use of the strategy method to elicit each subject's trust and reciprocal inclination. In our experiment, subjects made decisions both in the role of the sender and the receiver. In the role of the receiver subjects had to decide how much to send back for any possible amount received. This procedure allows us to measure both the level of trust and the level of reciprocity for each subject in the same strategic environment.[2]

To give subjects the monetary incentives to take all decisions seriously while

[2]Other studies have employed the strategy method in trust games in which subjects play only one role (e.g., Bellemare and Kröger 2007, Falk and Zehnder 2007). Burks et al. (2003) have subjects play both roles but do not use the strategy method.

at the same time avoiding potential confounds if subjects interact repeatedly in different roles, we employed the following incentive-compatible procedure. After all decisions had been made, a random mechanism determined which player of a given pair actually had the role of the sender and which player had the receiver role. Then, players' decisions were implemented and subjects were paid accordingly.

The experiment was programmed with the software z-Tree (Fischbacher 2007) and conducted at the BonnEconLab. Twenty subjects participated in each of the 12 sessions that we ran so that we observe the choices of 240 different subjects. The trust game was part of a sequence of tasks (see Dohmen and Falk 2006 for a detailed description). Before subjects played the trust game they had to solve math problems under different monetary incentives.[3]

After the trust game, we elicited subjects' risk attitudes using a series of 15 choices between a safe payment and a lottery. The lottery was the same across choices (400 points or 0 points, each with probability 0.5) while the safe option increased from 25 points to 375 points in increments of 25. If subjects have monotonous preferences, they prefer the lottery up to a certain level of the safe option, and then switch to preferring the safe option in all subsequent choices. After a subject had made decisions for all 15 choices, it was randomly determined which choice became relevant for payment.[4] Together, the trust game and the lottery choice task lasted about 20-25 minutes and subjects earned 6.87 Euro on average.

1.3 Results

We measure *trust* by the amount that a subject sends as a first mover. Our measure of *reciprocity* (also denoted "r") is derived as follows: for each subject, we used the

[3] In 4 of the 12 sessions subjects worked under purely individual incentives (fixed wages and piece rates). In the remaining 8 sessions they could select into an incentive scheme (team or tournament) which involved anonymous interaction with another player. All subjects were randomly rematched in the trust game. In view of our results we are confident that neither solving math problems nor the different incentive schemes systematically affect behavior in the trust game (see below).

[4] The experimental instructions for the trust game can be found in Appendix A.1. Except for the payoff parameters, the lottery procedure was identical to the one described in Appendix D.3.

decisions as a second mover and ran an OLS-regression of the amounts sent back on the (hypothetical) amounts sent by the opponent, forcing the slope through the origin. The slope coefficient gives us a measure of a subject's willingness to reward kind actions of an opponent by own kind behavior, i.e., positive reciprocity. If a receiver, for example, always matches his final payoff with that of the sender, his reciprocity coefficient is $r = 2$.

In order to graphically present our main result, we classify subjects according to their behavior as second movers. We call subjects with a reciprocity parameter $r > 1$ "reciprocal", and subjects with a slope parameter $r = 0$ "selfish". Reciprocal types leave their opponent with a positive return to trust, sending back more than the amount sent to them by the sender. 64.6% of our subjects fall into this category. Selfish types, who make up 12.5% of subjects, never send back anything, irrespective of the first mover's behavior. The remaining 22.9% of subjects whose slope parameter is positive, but small ($r \leq 1$) are categorized as "intermediate" types.

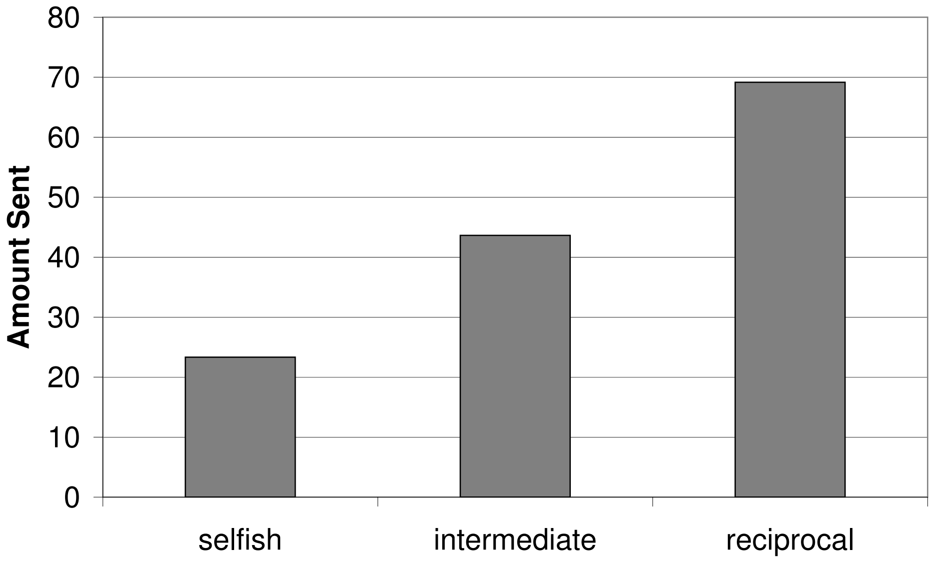

Figure 1.1: *Average amount sent by selfish, intermediate, and reciprocal players.*

Figure 1.1 plots the amount that the three types of subjects send on average in the trust game. Reciprocal types clearly send most (69.2 points on average), and selfish types send least (23.3 points). Subjects in the intermediate category send 43.6 points. Pairwise Mann-Whitney-U-tests indicate that all differences between the groups are highly statistically significant ($p < 0.01$ for all pairwise tests). Subjects who always "split the pie equally" as a responder (i.e., subjects with $r = 2$) trust most (83.5 points). The result depicted in Figure 1.1 is robust to different classifications of types using a finer "grid".

OLS-regressions of individuals' trust, measured by the amount sent in the trust game, on their reciprocal inclination, measured by the slope parameter r described above, confirm that reciprocal individuals trust more: an increase of one unit in the reciprocity measure is associated with sending 16.1 points more in the trust game (see Column (1) of Table 1.1).[5]

	Dependent variable: Amount sent		
	(1)	(2)	(3)
Reciprocity	16.091***	17.234***	17.761***
	(2.235)	(2.199)	(2.316)
1 if male		16.301***	15.167***
		(4.401)	(4.688)
Certainty equivalent			0.100**
			(0.044)
Constant	34.081***	24.329***	3.665
	(3.957)	(4.668)	(10.151)
R^2 adj.	0.175	0.217	0.230
Observations	240	240	221

Table 1.1: *Trust-Regressions. OLS estimates (standard errors in parentheses).* "Certainty Equivalent" *indicates the switch from the risky lottery to the safe option (=0 if subject is strongly risk averse, ..., =400 if subject is strongly risk loving). Significance at the 10%, 5% and 1% level is denoted by *, **, and ***, respectively.*

This key result is also robust to controlling for gender (cf. Column (2) of Table 1.1) and subjects' risk attitudes (cf. Column (3) in Table 1.1).[6] The positive influence

[5] This result still holds if we restrict the sample to the 4 sessions in which there was no interaction between subjects in the tasks preceding the trust game. In addition, we find the significant positive correlation between trust and reciprocity in the remaining 8 sessions irrespective of the chosen incentive scheme.

[6] The certainty equivalent cannot be determined unambiguously for 19 subjects because they

of reciprocity on trust is highly significant and also quantitatively very similar in all specifications. The effects of gender and risk attitudes are consistent with the findings in the literature. Men trust more than women (cf. Bohnet and Zeckhauser 2004), sending about 15 points more than female participants in our sample. Our results also confirm the importance of risk attitudes for trusting behavior: subjects who are more willing to take risks send significantly more (cf. Dohmen et al. 2006).

1.4 Discussion and Concluding Remarks

The strong, positive relationship between a person's reciprocal inclination and her trusting behavior has important implications for the evaluation and advancement of theories that incorporate social preferences. Some of the most prominent models (e.g., Fehr and Schmidt 1999, Bolton and Ockenfels 2000, Falk and Fischbacher 2006) predict that individuals who are more reciprocal (or inequity averse) ceteris paribus trust *less* than others in the trust game. The intuition for this result is that a selfish sender just suffers from the loss of her investment if the receiver sends back too little, whereas a fair-minded sender experiences additional disutility because his trust has been exploited.[7]

Our results show, however, that the relationship between trust and reciprocity may be more complex than captured by most models. The finding that people trust more the more reciprocal they are allows at least two different preliminary interpretations—one based on norm adherence and the other on systematic differences in beliefs. The idea of the former is that some people value adherence to a certain moral norm in itself. If these people follow a norm that, e.g., dictates cooperative behavior in either role, this could account for our main finding. Such norm-guided behavior could also help to explain why some senders in trust games send positive amounts despite expecting to get back less than they send (cf. Dufwen-

switched more than once between the safe option and the lottery. These subjects were excluded from the regression in Column (3). Including them with the lowest or highest switching point from the lottery to the safe option does not change the results.

[7]Along these lines, Fehr et al. (2007) have argued that "fairness preferences inhibit trusting behavior because trust typically involves a risk of being cheated."

berg and Gneezy 2000, Ashraf et al. 2006).

A different interpretation is that fair and selfish types have fundamentally different beliefs regarding the behavior of others. Such differences in beliefs might be the result of a "false consensus effect" (Kelley and Stahelski 1970). As an extreme example, assume that reciprocal players expect all others to behave reciprocally, and that a selfish subject expects all others to be selfish as well. In this case reciprocal types will send positive amounts and expect a positive return, while selfish types will never send anything since they expect that the receiver will not send anything back. Such systematic differences in beliefs would have interesting implications for the modelling of social preferences as they require giving up the widely used common-prior assumption. They potentially also have important practical implications as they could lead different types of players to select into different institutional settings. This could help to explain why environments with different degrees of exogenous enforcement coexist, e.g., in the labor market. Which of the two interpretations is more relevant cannot be answered with our data but remains an important question for future research.

Chapter 2

When Equality is Unfair

"To treat people fairly you have to treat people differently."
Roy Roberts, at that time VP of General Motors[1]

2.1 Introduction

In recent years, a vast body of literature has stressed the importance of gift exchange for mitigating moral-hazard problems of incomplete contracts: since many agents repay a gift in the form of higher wages by providing higher efforts, effort can be elicited under contractual incompleteness even in one-shot situations where no future gains can be expected (e.g., Akerlof 1982, Fehr et al. 1997, Maximiano et al. 2007). The potential of gift exchange as a contract enforcement device, however, is likely to depend on the institutions that shape the employment relation, above all the mode of payment. Yet little is known about the interaction of different payment modes with gift exchange. Exploring this interaction is crucial in order to understand under which conditions the efficiency-enhancing effects of gift exchange develop their full power. A key question in this context is how to treat agents relative to each other as this affects the perceived fairness of a pay scheme. In this chapter, we study this question by focusing on two important fairness principles: horizontal equality and equity.

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

On the one hand, it has been argued that *horizontal equality* is crucial for a wage scheme to be considered as fair. Differential pay of co-workers could cause resentment and envy within the workforce, and ultimately lower performance (e.g., Pfeffer and Langton 1993, Bewley 1999). Wage equality is also often referred to in employer-union bargaining as being a cornerstone of a fair wage scheme and is one of the most prevalent payment modes (see, e.g., Medoff and Abraham 1980, Baker et al. 1988). If workers care foremost about equality, a wage scheme that guarantees equal wages for co-workers should lead to an efficiency-enhancing gift exchange relation. On the other hand, the importance of the *equity principle* has long been discussed in social psychology, personnel management, and economics (e.g., Homans 1961, Fehr and Schmidt 1999, Konow 2003). In a work environment, the equity principle (or "equity norm") demands that a person who exerts higher effort should receive a higher wage compared to his co-worker. Only when performance of co-workers is the same, equity and equality coincide. However, in real-life work relations this is likely to be the exception rather than the rule. Whenever workers differ in their performance, horizontal wage equality violates the equity principle since a higher effort is not rewarded with a higher wage. In other words, if equity is important, the often-heard slogan "equal pay for equal work" implies "unequal pay for unequal work".[2]

Ideally, our research question would be examined in work environments that differ only with respect to the payment mode. To come close to this ideal world, we introduce a simple and parsimonious laboratory experiment that allows us to analyze the interaction between the institution of wage equality and gift exchange. In the experiment, one principal is matched with two agents. In a first stage the agents exert costly effort. After observing their efforts, the principal pays them a wage. In the main treatment he can choose the level of the wage but he is obliged to pay the same wage to both agents (*equal wage treatment* or EWT). In

[2]Lazear (1989) neatly summarizes this discussion (p. 561): "It is common for both management and worker groups such as labor unions to express a desire for homogeneous wage treatment. The desire for similar treatment is frequently articulated as an attempt to preserve worker unity, to maintain good morale, and to create a cooperative work environment. But it is far from obvious that pay equality has these effects."

the control treatment, the principal can wage discriminate between the two agents (*individual wage treatment* or IWT). In both treatments, neither efforts nor wages are contractible. Note that principals in the individual wage treatment are free to pay the same wage to both agents, i.e., the EWT is a special case of the IWT. If agents care foremost about wage equality, there should thus be no treatment difference; if equity considerations are more important, we should find that the EWT elicits lower effort levels than the IWT.

The main findings of the experiment are as follows. First, performance differs substantially between the EWT and the IWT: agents who are paid equal wages exert significantly lower efforts than agents who are paid individually. Effort levels are nearly twice as high under individual wages and efforts decline over time when equal wages are paid. Second, this strong treatment effect cannot be explained by differences in monetary incentives. The actual wage choices of principals imply that providing high effort levels is profitable for agents in both treatments. From a purely monetary viewpoint agents' behavior in both treatments should thus be similar. Third, we show that the frequent violation of the equity principle in the equal wage treatment can explain the effort differences between the treatments. In both treatments, agents who exert a higher effort and earn a lower payoff than their co-worker strongly decrease their effort in the next period. However, the norm of equity is violated much more frequently under equal wages. Principals in the IWT understand the mechanisms of equity quite well. When efforts differ they do pay different wages, rewarding the harder-working agent with a higher payoff in most cases. Agents' reactions cause completely different dynamics in the two main treatments. Under equal wages, initially hard-working agents get discouraged and reduce their effort to the level of their low-performing co-workers. By contrast, in the individual wage treatment the high performers keep exerting high efforts while the low performers change their behavior and strongly increase their effort levels.

Note that principals in the IWT can set two wages instead of one in the EWT. This opens the possibility that agents attribute a different degree of intentionality to principals' wage choices. It could be that this additional moment of discretion has a direct impact on the treatment difference. To rule out this potential confound, we conduct an additional control treatment where principals can again set only one wage as in the EWT. The second wage is set *exogenously* such that the equity principle is always fulfilled. Effort levels in the control treatment are similar to those of the IWT and much higher compared to the EWT. This confirms that the difference between our two main treatments is indeed driven by agents' desire for wages that are in line with the equity principle.

Our results suggest a psychological rationale for using individual wages. Subjects perceive equal wages for unequal performance as unfair and reduce their effort subsequently. The traditional literature on incentive provision in groups comes to a similar conclusion though for a different reason. It is usually argued that the inefficiency of equal wages stems from the fact that marginal products and wages are not aligned. This can lead to free-riding among selfish agents (e.g., Holmström 1982, Erev et al. 1993). We enlarge the scope of this critical view on wage equality: interestingly, in our setup it is precisely the presence of fair-minded agents and not their absence that calls for the use of individual rewards.

An earlier literature in social psychology also studies the consequences of equity in social exchanges (Homans 1961, Adams 1963, Adams 1965, Andrews 1967). In his influential equity theory, Adams (1965) operationalizes the general equity principle in an "equity formula", which states that the ratio of outcomes to inputs should be the same for every individual.[3] If this is not the case an individual experiences distress and seeks to reestablish equity. Our study complements this literature in several ways. As Mowday (1991) notes, interpreting the existing empirical evidence can often be difficult because important aspects such as the costs of effort or the relevant reference group are ambiguous. Our economic laboratory experiment offers a high level of control over these aspects. In addition, violations of the equity norm arise from the interaction of principals and agents in our study whereas they are induced by the experimenter in most earlier experiments, e.g., by making subjects believe they are over- or underqualified for the job (e.g., Adams 1963 or Lawler 1967).

[3]The idea of proportionality dates back to at least Aristotle's Nicomachean Ethics.

Since agents in our experiment compare their payoff with the payoff of their co-worker, our results also inform the literature analyzing the influence of relative income on satisfaction and performance. It has been shown that relative income affects people's well-being (e.g., Clark and Oswald 1996, Easterlin 2001, Fließbach et al. 2007). However, it is less clear how this influences performance, i.e., whether low relative income leads to frustration and reduced performance (as in Clark et al. 2006 and Torgler et al. 2006) or to an increase in performance due to a "positional arms race" (Neumark and Postlewaite 1998, Layard 2005, Bowles and Park 2005). The controlled laboratory environment of our experiment allows us to reconcile these differing views. Our results indicate that the comparison process goes beyond a one-dimensional comparison of income and also includes a comparison of effort. In particular, they suggest that receiving a lower income while exerting a *higher* effort leads to reduced performance as this conflicts with the equity principle. By contrast, a lower income that is generated by a *lower* effort leads to a (small) increase in performance.

There are only a few experimental studies that analyze the interaction of payment modes and social preferences (e.g., Bandiera et al. 2005, Fehr et al. 2007, Falk, Huffman and MacLeod 2008). Most closely related to this chapter is the work of Charness and Kuhn (2007). Here, one principal is matched with two agents differing in productivity; like in our study, wages and efforts are not contractible. In contrast to our results, they find that co-workers' wages do *not* matter much for agents' decisions. However, their design differs from ours in several important points. While Charness and Kuhn focus on heterogeneity in productivity, we look at the effect of actual output differences between agents. Furthermore, we allow for richer comparisons between the agents, as in their design agents are not aware of the magnitude and direction of the productivity differences. The different results underline the importance of information for determining the reference group: Charness and Kuhn's results rather apply to groups of workers that are loosely related and know little about each other, while our focus is on close co-workers who have a good understanding about their peers' abilities and efforts.

Regarding compensation practice in firms, our findings highlight the importance of taking the concerns for co-workers' wages into account. However, doing so by paying equal wages to a group of agents may actually do more harm than good. As soon as agents differ in their performance, equal wages which seem to be a fair institution at first sight might be considered very unfair. While the discouraging effect of equal wages on hard-working agents has long been informally discussed (e.g., Milgrom and Roberts 1992, p. 418f) we provide controlled evidence in favor of this intuition. Moreover, our findings suggest that it is the violation of the norm of equity that causes the discouragement and low performance. Our results should not be interpreted as arguments against wage equality in general. They rather point to limits of equal wages.[4] Wage equality is potentially a good choice in occupations where, e.g., due to technological reasons, workers' performance differs only slightly or where performance differences are due to random influences. In addition, the transparency of co-workers' work efforts and wages might have an influence on the optimal choice of the pay scheme.

The remainder of this chapter is structured as follows. In the next section we describe the experimental design and discuss theoretical predictions. In Section 2.3 we present and discuss our results and Section 2.4 concludes.

2.2 Experimental Setup

2.2.1 Design and Procedures

In the experiment, one principal is matched with two agents. The subjects play a two-stage game. In the first stage, agents decide simultaneously and independently how much effort they want to provide. Exerting effort is costly for the agents. Effort choices range from 1 to 10 and are associated with a convex cost function displayed in Table 2.1. The principal reaps the benefits of production: every unit of effort increases his payoff by 10.

[4] Independent of equity-equality trade-offs, equal wages might be beneficial for the principal because they could increase peer monitoring (Knez and Simester 2001) and lower transaction costs since contracts do not have to be negotiated with every worker individually (e.g., Prendergast 1999).

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

Table 2.1: *Cost of effort.*

In the second stage, after observing the effort decisions of his agents, the principal decides on wages for the two agents. The wages have to be between 0 and 100. Neither efforts nor wages are contractible. The only difference between treatments is the mode of payment. In our main treatment the principal can only choose one wage w that is paid to each of the agents (*equal wage treatment* or EWT). In the control treatment he can discriminate between the two agents by choosing wages w_1 and w_2 for agent 1 and 2, respectively (*individual wage treatment* or IWT). The EWT is thus a special case of the IWT. At the end of each period, the two agents and the principal are informed about efforts, wage(s), and the resulting payoffs for all three players. The payoff functions for the players are summarized in Table 2.2.

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

Table 2.2: *Payoffs of players.*

This game is played for twelve periods. We implemented a stranger design to abstract from confounding reputation effects, i.e., at the beginning of each period principals and agents were rematched anonymously and randomly within a matching group. A matching group consisted of three principals and six agents. The subjects kept their roles throughout the entire experiment. After the last period, subjects answered a short post-experimental questionnaire. The experiment was conducted in a labor market framing, i.e., principals were called “employers” and agents were called “employees”.[5]

[5]An English translation of the instructions can be found in Appendix B.1.

Our setup is related to the gift-exchange game introduced by Fehr et al. (1993) but differs in two important ways. First, in our experiment agents move first while in Fehr et al.'s setup the principal moves first. Our move order allows the principal to base his wage decision on the actually exerted effort. More importantly, a principal in our experiment is matched with two agents instead of one. This is an essential prerequisite to analyze the interaction between gift exchange and payment modes. It allows us to study the impact of relative wages on the perceived fairness of the wage scheme and agents' behavior.

All participants started the experiment with an initial endowment of 400 points that also served as their show-up fee. Points earned were converted at an exchange rate of 0.01 Euro/point. The experiment was conducted at the BonnEconLab at the University of Bonn in April 2005 using z-Tree (Fischbacher 2007). For each treatment, we ran four sessions with a total of 8 matching groups (144 participants). The experiment lasted approximately 70 minutes. On average subjects earned 8.30 Euro.

2.2.2 Behavioral Predictions

Efficiency is determined by agents' effort choices. It is maximized if both agents exert the highest possible effort of 10. However, if all players are rational and selfish the principal will not pay anything to the agents since wage payments only reduce his monetary payoff. Anticipating this, both agents will provide the minimal effort of one in the first stage. The finite repetition of the game in randomly rematched groups does not change this prediction. This subgame perfect equilibrium is the same for both payment modes. If all players were selfish we should therefore expect no difference between treatments.

By contrast, in laboratory experiments studying labor relations with incomplete contracts, one typically observes that efforts and wages exceed the smallest possible value. Moreover, wages and efforts are positively correlated (e.g., Fehr and Gächter 2000). These findings illustrate the potential of reciprocal gift exchange in enforcing incomplete contracts, as postulated in Akerlof and Yellen's fair wage-effort hypothesis (Akerlof and Yellen 1990). A fundamental prerequisite for the functioning of gift-exchange relations is that workers perceive their wage as fair. The fairness of a wage payment, however, may not only be evaluated in absolute terms, but also *relative* to the wages of other members in a worker's reference group.[6] This is not important for the special case of bilateral gift-exchange relationships where only one agent interacts with one principal (e.g., Fehr et al. 1997). However, horizontal fairness considerations potentially play a crucial role in our setup where workers can compare to co-workers.

How do the behavioral predictions depend on which horizontal fairness principle is most important? If agents in the experiment care foremost about *wage equality*, the EWT—which guarantees equal wages by design—should lead to efficient gift exchange between firms and workers. Additionally, we should expect no behavioral differences between treatments since firms in the IWT can pay their workers equal wages, too. Given that firms in the IWT recognize workers' desire for equal treatment, they will decide to do so. Thus, the wage-effort relationship and average effort levels should not differ across treatments. If some firms do nevertheless wage discriminate between workers, the IWT should lead to less efficient outcomes than the EWT.

By contrast, if workers consider *equity* to be more important than equality, we should expect differences in behavior between treatments. The equity principle demands that a person who exerts higher effort than his co-worker should receive a higher wage and payoff. Our experimental treatments differ in the extent to which the equity principle can be fulfilled by principals. Under the equal wage institution, the equity norm is violated whenever agents differ in their performance. Since both workers receive the same wage but have to bear the cost of effort provision, the worker who exerts more effort receives a lower monetary payoff. Under individual wages, principals' behavior determines endogenously whether the equity norm is violated or not. By differentiating wages in accordance to effort differences, principals

[6]Potentially many variables influence a worker's fairness perception of his wage, e.g., the unemployment rate, unemployment benefits, the prevailing market wage, etc. (see Akerlof 1982, Akerlof and Yellen 1990). These factors are ruled out by our experimental design, allowing us to isolate the influence of co-workers' wages on fairness perception and effort provision.

can adhere to the norm. If we assume that at least some principals do so, we expect to see less norm violations in IWT than in EWT.

What are the behavioral consequences of such differences in norm fulfillment? Agents who value equitable treatment should suffer from norm violations, feel dissatisfied and subsequently try to restore equity by adjusting their behavior. Equity theory proposes several possible reactions of agents after norm violations, such as altering own or others' efforts or payoffs, changing one's reference group or quitting the relationship (see Adams 1965). The virtue of our experimental design is that we can clearly identify agents' reactions, because the only variable that an agent can change after experiencing a norm violation is his work effort. An agent who faces a disadvantageous norm violation (i.e., relative underpayment) should lower his effort in the following period. An agent who experiences an advantageous norm violation (i.e., relative overpayment) should increase his effort. Note that a norm violation always includes one agent facing a disadvantageous violation and one agent facing an advantageous violation. Dissatisfaction and the resulting strength of reactions, however, is likely to depend on the direction of the norm violation. Previous evidence suggests that the decrease of effort after a disadvantageous norm violation will be stronger than the increase of effort after an advantageous violation (Loewenstein et al. 1989, Mowday 1991, Thöni and Gächter 2008). Consequently, a violation of the equity norm should lead to an *overall decrease* of efforts in the subsequent period.

If workers care about equitable payment in the sense of the postulated equity norm, aggregate effort in the EWT should thus be lower compared to the IWT since we expect to observe less norm violations in the latter.

2.3 Results

In this section we present the results of the experiment and discuss possible explanations for the observed behavior. We first analyze efficiency implications of the two payment schemes by comparing the effort choices of agents. We then demonstrate that the difference in agents' performance obtains even though monetary incentives—implied by principals' wage setting—should lead to similar effort choices in both treatments. Subsequently, we show that workers' behavior is strongly affected by the equity principle, which is more frequently violated in the EWT. Finally, we report the results of an additional control experiment. They demonstrate that the higher efficiency of the IWT is not driven by the fact that principals can set two wages instead of one (as in the EWT) but by the fact that principals set wages that are in line with the equity principle.

2.3.1 Effort Choices and Efficiency

Figure 2.1 shows the development of average efforts over time. Under equal wages, efforts are lower already in the first period (Mann-Whitney test: $p = 0.03$)[7] and decrease over time. Efforts under individual wages stay constant (Wilcoxon test for periods 1–6 against 7–12: IWT, $p = 0.56$; EWT, $p < 0.01$). This results in a strong overall treatment difference: average efforts are almost twice as high in the IWT compared to the EWT (8.21 vs. 4.40; Mann-Whitney test: $p < 0.01$). The treatment difference is also present when individual matching groups are considered: the highest average effort of an EWT matching group (5.88) is still lower than the lowest average effort of an IWT matching group (7.47).

The difference in agents' behavior can also be seen in the histogram of effort choices (Figure 2.2). In the individual wage treatment agents choose the maximum effort of 10 in 49% of the cases, 84% of the choices are higher than 6. Under equal wages, agents choose an effort higher than 6 in only 26% of all cases. The effort decisions are more spread out in the EWT, the minimal effort of 1 being the modal choice with 24% of the choices. Since higher efforts increase production and since the marginal product of effort always exceeds its marginal cost, the differences in effort provision directly translate into differences in efficiency.

[7]The comparison of first period effort choices is based on individual observations. Unless otherwise noted, all other tests use matching group averages as independent observations. Reported p-values are always two-sided.

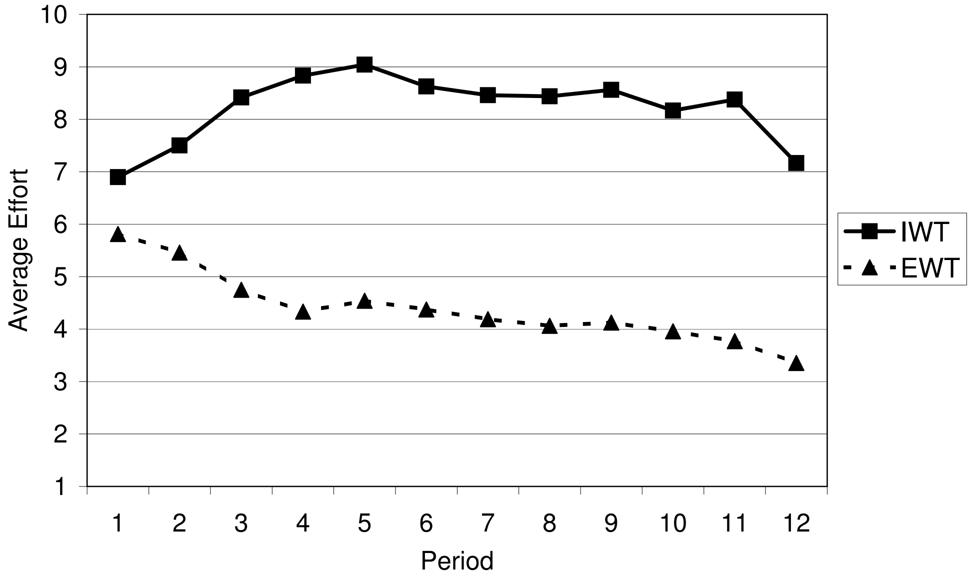

Figure 2.1: *Average effort per period. The effort is aggregated per period over all matching groups.*

> **Result 1:** *The two payment modes exhibit strong differences with respect to the performance they elicit: agents who are paid equal wages exert significantly lower efforts than agents who are paid individually. This results in a much higher efficiency under individual wages.*

Both, the agents and the principals benefit from the increase in efficiency. The average profit per period of a principal is 56 in the EWT compared to 100 in the IWT (Mann-Whitney test: $p < 0.01$), while an agent on average earns 10 under equal wages vs. 17 under individual wages (Mann-Whitney test: $p < 0.01$).

2.3.2 Wage Setting and Monetary Incentives

The strong difference in effort choices suggests that the degree to which gift exchange can mitigate moral-hazard problems depends on the payment mode that is used. Wage equality hampers efficiency, and we hypothesized above that this might be due to horizontal fairness concerns. However, performance differences might also be driven by differing monetary incentives across treatments. To rule this out, we now take a closer look at principals' wage setting and the resulting monetary incentives for the agents.

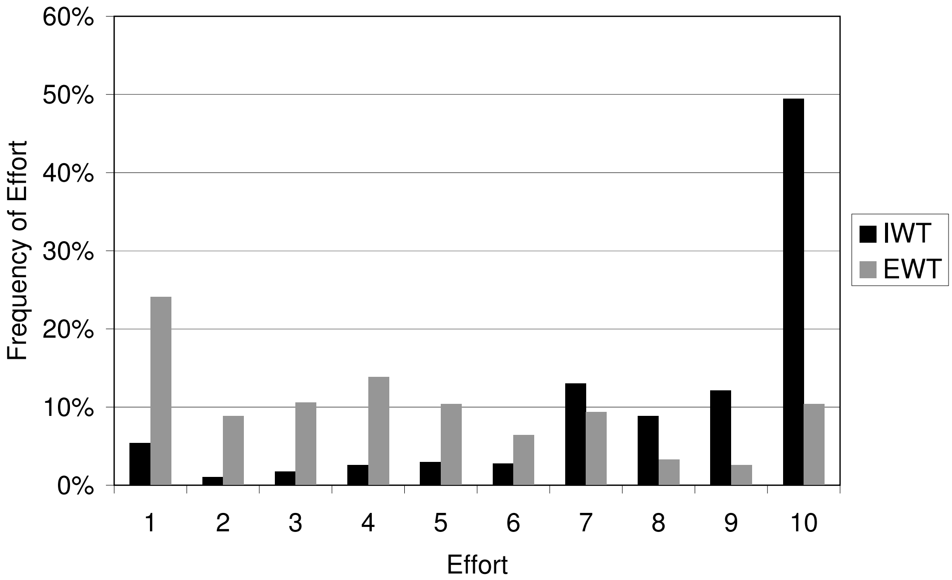

Figure 2.2: *Frequency of effort choices.*

Figure 2.3 plots the average wage per effort level in the two treatments. For both treatments we take the wage paid by the principal for each individual effort decision and calculate averages for a given effort level. The graph exhibits the upward sloping effort-wage relation of many gift-exchange experiments. For example, an agent in the equal wage treatment who exerts an effort of 1 receives on average a wage of 6.3 while an agent exerting an effort of 10 receives an average wage of 30.3. In the individual wage treatment, the corresponding wages are 1.7 and 39.5.[8] The effort-wage relation indicates that gift exchange indeed occurs between principals and agents. In both treatments, higher effort levels are reciprocated with higher wages.

[8]Since principals in the EWT have to pay the same wage to both agents, an interesting question concerns how they choose this wage when confronted with a low and a high effort. To answer this question, we assume that the wage-effort relation from IWT reflects the "true" wage-setting preferences of principals because wage choices were not constrained in this treatment. We regress wages on effort in the IWT and calculate predicted wages for all possible levels of effort. We then calculate the differences between actual wages paid in EWT and these predicted wages. This analysis shows that the actual wage in EWT is very close to the average between the predicted wage for the higher and lower effort in IWT (the detailed analysis is available upon request). This means that principals in the EWT weight the higher and lower effort about equally when deciding on the wage payment.

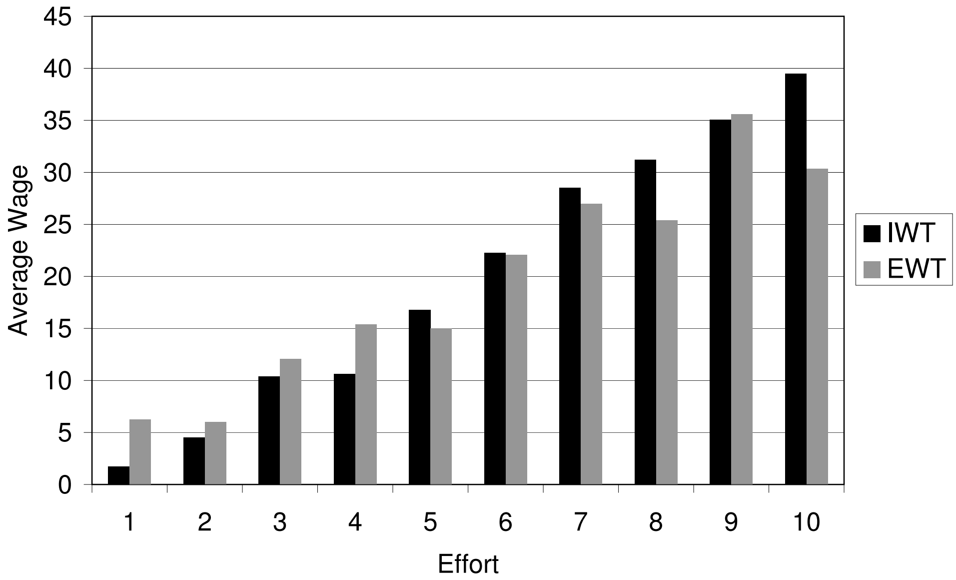

Figure 2.3: *Average wage for a given effort.*

Result 2: *Principals reward higher effort levels with higher wages in both treatments.*

Reciprocal behavior of the principals generates monetary incentives for the agents. In order to calculate the monetary incentives entailed in principals' wage decisions, one has to take into account agents' cost of effort exertion (see Table 2.1). Qualitatively, this does not change the picture of the effort-wage relation: higher effort levels seem to lead not only to higher wages, but also to higher profits for the agents. To check this in more detail, we estimate an OLS-model where we regress the agent's profit per period π_{A_i} on his effort level e_i and a constant. To account for potential differences between treatments we include a treatment dummy IWT, and an interaction term of the treatment dummy and the agent's effort. IWT equals 1 for the individual wage treatment and 0 for the equal wage treatment. Reported robust standard errors are adjusted for clustering within matching groups. The estimation results are shown in Column 1 of Table 2.3. The coefficients indicate that the effort-profit relation is indeed positive in both treatments. On average, an additional unit of effort increases the agent's profit under equal wages by 1.031 points. This coefficient is weakly significant. In the individual wage treatment the effort-profit relation is slightly steeper: an effort increase of 1 leads to an increase in agent's profit of 1.804 points (1.031 + 0.773). The difference between treatments, however, is not significant.

Dep. Variable	π_{A_i}	π_{A_i}
e_i	1.031*	0.854**
	(0.535)	(0.348)
$IWT \times e_i$	0.773	0.995*
	(0.615)	(0.469)
cons	5.927**	-5.815***
	(2.614)	(1.523)
IWT	-3.744	11.004***
	(3.235)	(3.274)
e_j		2.774***
		(0.280)
$IWT \times e_j$		-3.178***
		(0.403)
N. Obs.	576	576
R^2	0.100	0.238

Table 2.3: *Profit regressions. Robust standard errors are adjusted for clustered matching groups and are given in parentheses. For each firm, one observation per period is included in the analysis. The dummy "IWT" is equal to 1 for the individual wage treatment. Significance at the 10%, 5% and 1% level is denoted by *, **, and ***, respectively.*

We also estimate a second model where we control for the co-worker's effort e_j (see Column 2 of Table 2.3). The results indicate that the co-worker's effort choice has a substantial influence on an agent's profit under wage equality while it has a negligible influence if individual wages are paid. An increase in agent j's effort increases agent i's profit in a given period by 2.774 points in the EWT, while the (insignificant) influence in the IWT is -0.404 ($= 2.774 - 3.178$). However, it is still individually profitable for the agents to exert high efforts in the EWT. An additional unit of (own) effort increases the agent's profit by 0.854 points.[9] Our

[9] One could object that subjects in the experiment did not have access to the analyses we just presented, because these are "ex-post" examinations while subjects only observed behavior and

findings concerning agents' monetary incentives can thus be summarized as follows.

> **Result 3:** *The wages paid by principals imply similar monetary incentives in both treatments. A higher effort level leads to a higher profit in both treatments.*

2.3.3 The Importance of Equity

In light of the previous result, the strong differences in actual efforts and especially the low effort levels under equal wages are remarkable and stress the significance of non-pecuniary motivations for agents' performance. Agents under equal wages predominantly choose low efforts, thereby foregoing considerable profits. Apparently, equal wages are not reconcilable with agents' horizontal fairness considerations. On the other hand, agents under individual wages provide very high effort levels. Thus, *aggregated* behavior is consistent with the predictions of equity-concerned agents. We therefore focus our analysis of non-monetary motivations on the question whether *individual* behavior is in line with a concern for the fulfillment of the norm of equity.

Agents' Reactions to Norm Violations

We first analyze how agents react to a violation of the norm of equity. Equity theory argues that agents experience distress from inequity and take action to reduce it—which in our setup means to increase or decrease the individual effort. The direction of the effort adjustment should depend on the type of norm violation. An equity-concerned agent who works more but does not receive a higher payoff than his co-worker faces a *disadvantageous norm violation*. To restore equity, he can only

---

outcomes of their previous groups. We therefore calculate the profit-maximizing effort level for each agent in each period based on the information this subject actually has. If we assume that agents choose the effort level that was on average the most profitable of all effort levels they have observed so far, the calculations show that agents in the EWT could have increased their efforts and profits considerably even by using only their limited information. In the last period, the average profit-maximizing effort level exceeds the average actual level in that period by 61%. By contrast, the average actual effort levels of subjects in the IWT are very close to the profit-maximizing levels.

	Effort Down	Effort Constant	Effort Up	N. Obs.
EWT				
No Violation	19.1 %	54.4 %	26.5 %	68
Adv. Violation	12.2 %	43.5 %	44.3 %	230
Disadv. Violation	52.6 %	33.9 %	13.5 %	230
Total	30.7 %	40.7 %	28.6 %	528
IWT				
No Violation	19.2 %	51.8 %	29.0 %	448
Adv. Violation	45.0 %	27.5 %	27.5 %	40
Disadv. Violation	35.0 %	57.5 %	7.5 %	40
Total	22.3 %	50.4 %	27.3 %	528

Table 2.4: *Frequency of effort reactions.*

decrease his effort. Analogously, his co-worker who exerts a lower effort and earns a higher profit faces an *advantageous norm violation* and should increase his effort.[10]

Table 2.4 shows how often agents decrease, increase or do not change their effort from period t to $t+1$ after they experienced no, an advantageous or a disadvantageous norm violation in period t. The top panel of Table 2.4 reports data for the equal wage treatment. When the norm is fulfilled, most agents keep their effort constant (54%) and slightly more agents increase their effort than decrease it. After experiencing an advantageous norm violation, agents tend to increase their effort (44%) and only few reduce it (12%). The opposite is true after a disadvantageous norm violation: the majority of agents decrease their effort (53%) and only few increase their effort in the following period (14%). In line with equity theory these numbers suggest that agents change their effort provision in the direction that makes a violation less likely to occur in the next period.

Behavior in the individual wage treatment (bottom panel) is very similar to behavior in the EWT for the cases of no violation and disadvantageous violations.

[10]More precisely, an advantageous norm violation comprises all cases when efforts are equal but profit is higher, or when effort is lower but profit is not. A disadvantageous norm violation occurs if efforts are equal but profit is lower, or if effort is higher but profit is not.

When the norm is not violated agents mostly keep their effort unchanged. After a disadvantageous norm violation efforts are decreased rather than increased, as in the EWT. The only difference between the treatments is observed when agents experience an advantageous norm violation: agents in the IWT tend to decrease their effort while the EWT agents tend to increase it in this case.[11]

The pattern of individual reactions to norm violations indicates that agents care about equity; we therefore check next how often norm violations occur in the two treatments. We expected to see more norm violations in the EWT than in the IWT, because the equal wage institution forces principals to set wages that are not in line with the norm of equity whenever agents exert different efforts. This is indeed what we observe. While the norm is violated in 87% of all cases (460 out of 528) in the EWT, the figure for the IWT is only 15% of all cases (80 out of 528). Thus, even if individual reactions in a given situation are similar, agents in the EWT are far more often exposed to norm violations than agents in the IWT. Principals in the IWT seem to understand quite well that agents care about equity and use the possibility to set different wages in a sophisticated way. If efforts differ, they reward the more hard-working agent with a higher wage in 90% of these cases. If agents exert the same effort, principals pay equal wages in 90% of the cases.

> **Result 4:** *Agents mostly react to disadvantageous violations of the norm of equity by reducing their effort and by increasing it after an advantageous norm violation. The norm of equity is far more often violated in the equal wage treatment.*

So far we have seen that agents' reactions are largely in line with the hypotheses of equity theory and that treatments differ with respect to the frequency of equity-norm violations. Yet, this is not sufficient to explain the treatment effect, since a

[11] We checked the robustness of the reaction patterns in several ways. For example, it could be that agents react differently to norm violations if they are paid very high or low absolute wages. However, performing the analysis only for agents receiving a wage out of the top or bottom quartile of the ex-post wage distribution does not alter the result. An implicit assumption of our analysis is that the gift-exchange relation is generally intact between principal and agent, i.e., that agents exert a non-minimal effort and that principals pay a positive wage. The results do not change if one restricts the analysis to these cases. Also if one defines gift exchange as requiring the agent's profit to be positive, i.e. $w_i > c(e_i)$ instead of $w_i > 0$, the results are very similar.

norm violation is always advantageous for one agent and at the same time disadvantageous for the other one. If both agents adjust their effort in a similar way but in opposite directions the adjustments will cancel out. However, previous evidence suggests that reactions to a disadvantageous norm violation are stronger than reactions to an advantageous one (e.g., Loewenstein et al. 1989, Mowday 1991, Thöni and Gächter 2008). If this is the case, norm violations could explain the downward trend in the EWT and the treatment difference in effort provision.

Figure 2.4 shows the average magnitude of changes in effort provision from period t to period $t+1$ after an agent experienced no norm violation, a disadvantageous or an advantageous norm violation in period t. The width of the bars corresponds to the number of observations in the respective category (cf. last column of Table 2.4). When the equity-norm is not violated agents tend to keep their effort constant or even slightly increase it. After a disadvantageous norm violation, agents in the EWT react strongly. They decrease their effort by 1.30. Their co-worker, experiencing an advantageous norm violation, increases his effort but not as strong. He raises his effort by only 0.75. The difference is statistically significant (Wilcoxon test of the absolute values: $p = 0.01$). In the IWT, both groups of agents experiencing a norm violation decrease their effort. The strength of reactions indicate that agents suffer more from a disadvantageous norm violation than from an advantageous one. This results in an overall decrease of efforts after a norm violation.

> **Result 5:** *Agents' reactions to a violation of the norm of equity are asymmetric: the negative reaction of the disadvantaged agents is stronger than the reaction of the advantaged agents. This asymmetry in agents' reactions results in an overall negative time trend in efforts for the EWT and in the strong treatment difference in effort.*

The analyses above suggests that agents care about equity and experience the equal wage scheme as unfair. Interestingly, even the principals consider the equal wage scheme as less fair. In the post-experimental questionnaire, principals were presented three hypothetical game situations that included effort choices, wage choices, and the resulting payoffs for all players. They were asked whether they considered the resulting allocation as just. One of the three situations reflected their own aver-

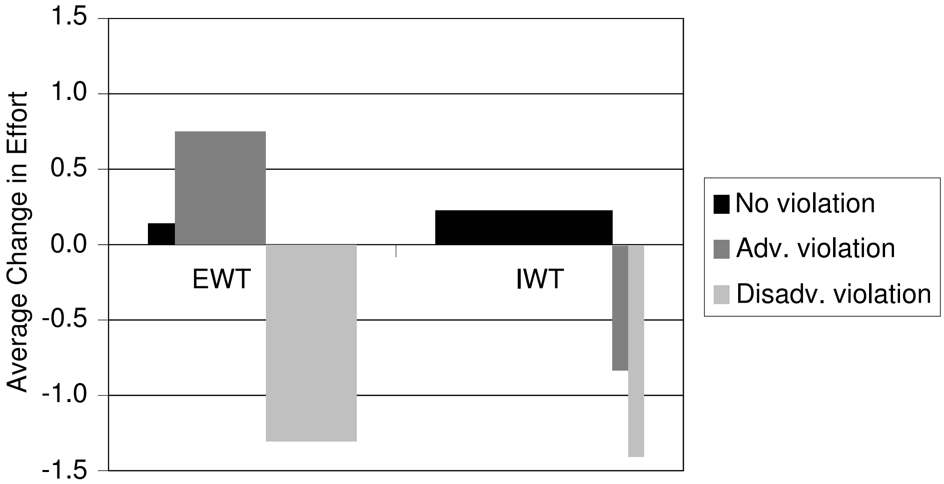

Figure 2.4: *Magnitude of effort reactions. The average change in effort from period t to period $t+1$ is shown given that the agent experienced no norm violation, an advantageous violation or a disadvantageous norm violation in period t. The width of the bars corresponds to the number of observations.*

age behavior in the experiment.[12] The principals did not know that they were facing their own past decisions when answering the question. 63% of the principals in the IWT considered their own decisions fair while only 38% of the principals in the EWT shared this view (Mann-Whitney test on matching group shares: $p = 0.03$).

Simulation with Equity-Concerned Agents

We demonstrated above that horizontal fairness concerns shape agents' behavior under the two payment schemes. In combination with the frequent violations of the norm of equity in the EWT we are able to explain the performance differences across treatments. In order to further illustrate how institutions and equity-concerns interact, we take our previous findings on agents' period-to-period reactions and link them to the aggregate dynamics in the experiment. We do so with a simulation in which all agents are assumed to derive utility from money, but to also suffer whenever

[12]This situation was constructed as follows: We calculated the average effort of the higher-effort and of the lower-effort providers that the principals actually faced during the experiment. We then took the average of the wages the principals paid to the two groups. Finally, we calculated hypothetical payoffs for all three "average" players by considering the costs of the average efforts.

the equity principle is not met. When deciding about their effort in a given period, the simulated agents compare their effort and profit in the previous period with the effort and profit of their co-worker in that period. According to the comparison along these two dimensions, four reactions can be distinguished for the simulated agents. (i) For an agent who had a higher effort and a higher profit, the norm of equity is fulfilled and the pecuniary comparison is also advantageous for him, so he keeps his effort constant. (ii) For an agent who exerted a lower effort and got a lower profit, the norm is satisfied but profit maximization is not, thus he partly adjusts his effort in the direction of his co-worker's effort, i.e., he chooses an effort $(e_{i,t} + e_{j,t})/2$. (iii) An agent with higher effort and lower profit feels distressed as he suffers from a disadvantageous norm violation. He adjusts his effort fully and chooses $e_{i,t+1} = e_{j,t}$. (iv) Finally, for an agent with lower effort and higher profit the norm violation is advantageous, thus the resulting utility is higher than in case (iii). He chooses an effort $(e_{i,t} + e_{j,t})/2$. The reactions in cases (i) to (iv) are in line with the period-to-period reactions presented in Table 2.4 and Figure 2.4.

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

Figure 2.5 shows how effort choices evolve over time in the experimental data and in the simulations. The simulations *'EWT sim'* and *'IWT sim'* trace the real data very well and are able to reproduce the large effort difference between treatments. In the individual wage simulation, efforts increase like the real efforts although the slight downward trend in the second half of the experiment cannot be reproduced. Efforts in the equal wage simulation constantly decrease down to an effort level slightly above 3 in the final period. This pattern is very similar to the dynamics in the real data.

Note that the pivotal agent is different between the simulated treatments: in the equal wage simulation the norm of equity is violated when agents choose different

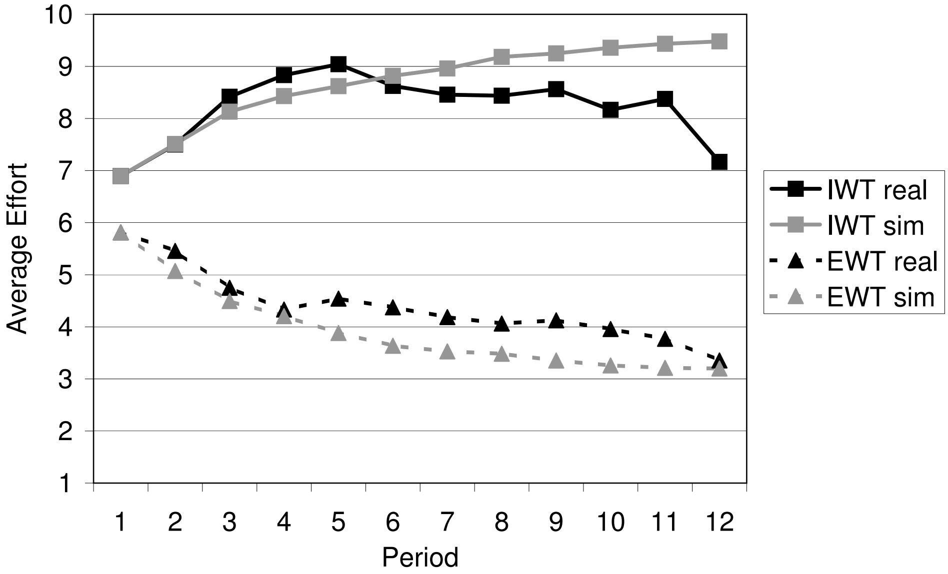

Figure 2.5: *Simulated efforts of agents adapting to equity-norm violations.*

effort levels. In these cases, the agent with the higher effort will fully adjust his effort in the direction of his co-worker's effort while the co-worker will increase his effort level only to the average effort of the last period. In the EWT simulation, the average effort therefore converges to the lowest first period effort as agents are subsequently re-matched: the low-effort providers are pivotal. By contrast, in the IWT the high-effort providers have the decisive impact on the overall outcome. The norm of equity is mostly fulfilled in the IWT. Thus, the agent with the higher effort keeps his effort constant while his co-worker adjusts his effort. The average effort therefore converges to the highest first period effort. We will analyze this point in more detail in the next section.

> **Result 6:** *Simple simulations based on agents who have preferences for money and equitable treatment are in line with the efforts observed in the experiment and are able to reproduce the observed treatment effect.*

2.3.4 Dynamics of High-Effort and Low-Effort Providers

As already seen in Figure 2.2, subjects exhibit a substantial degree of heterogeneity with respect to effort provision. In the following, we analyze if the agents who are most or least willing to exert effort are affected differently by the two payment modes at hand. A common informal argument claims that equal wages will be especially detrimental to the motivation of high performers but clean empirical evidence is scarce. Furthermore, it is unclear how weakly motivated agents react to equal or individual wages. We also address the question whether high and low performers impact the overall results differently in the two treatments. The simulations presented in the previous section suggest that this could indeed be the case: in the EWT simulation, the low-effort providers are decisive for the final outcome while it is the high-effort providers in the IWT simulation.

To analyze these questions in the experimental data we classify agents according to their effort decision in the *first period*. We define the agent with the highest first-period effort in each matching group as "high-effort provider" and the agent with the lowest effort as "low-effort provider". This type definition is chosen because when agents decide on their effort in the first period, they do not have any information about the behavior of other subjects and all learning and coordination processes occur after this initial effort choice. Thus first-period effort is likely to be a good proxy for the intrinsic willingness of a specific agent to exert effort. If some of the subjects are intrinsically inclined to exert high efforts they should show up in the group of high-effort providers. In contrast, if some of the subjects are intrinsically inclined to exert low efforts they should show up in the group of low-effort providers.

In Figure 2.6 we follow the high-effort providers and low-effort providers in both treatments and show their effort decisions over time. In the first period, the groups of high-effort providers and the groups of low-effort providers are close together across treatments.[13] This changes completely over the course of the 12 periods. In the individual wage treatment, high-effort providers continue to provide high effort levels. Low-effort providers increase their efforts dramatically up to the level of the high-effort providers and even higher in the last periods. In the equal wage

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

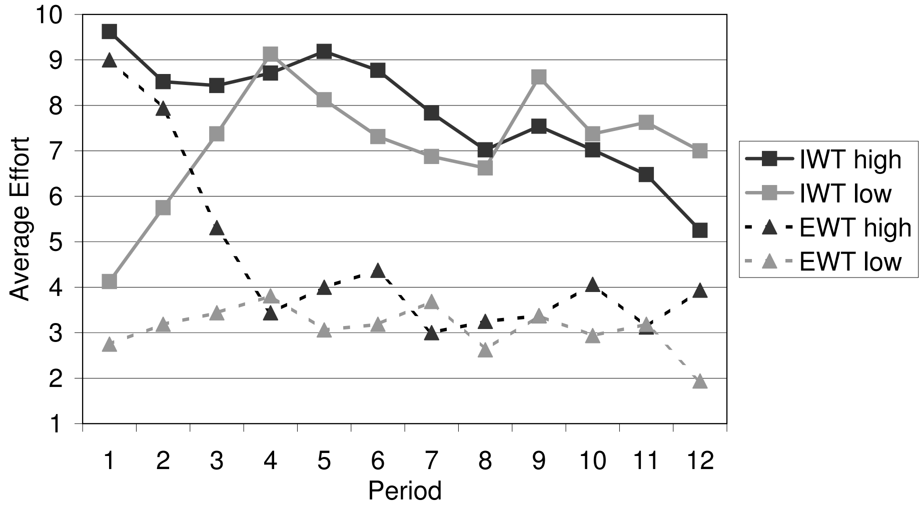

Figure 2.6: *Effort decisions of high-effort and low-effort providers. In each matching group, the agent with the highest (lowest) effort in the first period is defined as the high (low)-effort provider.*

treatment, the dynamics are reversed. Here, the low-effort providers keep their effort provision constant and the high-effort providers reduce their efforts to the level of the low-effort providers. In the last six periods, effort levels are not different within treatments (Wilcoxon signed rank test: $p = 0.67$ (IWT), $p = 0.78$ (EWT)) while they differ between treatments (Mann-Whitney test: $p < 0.01$ (high-effort providers), $p < 0.01$ (low-effort providers)). Put simply, the "good" agents push the "bad" agents up under individual wages while under equal wages the "bad" ones pull the "good" ones down.

These dynamics underline the importance of the different non-monetary motives induced by the two wage setting institutions. Remember that agents face similar monetary incentives in both treatments, but wage equality often violates the norm of equity. Agents in this treatment who are in principle willing to exert high levels of effort get frustrated and lower their efforts. On the contrary, under individual wages where the norm of equity is intact, good performance spreads. These results suggest that choosing a wage scheme also influences the social dynamics between the agents. In our experiment, individual wages lead to positive dynamics since agents orientate themselves by the most hard-working agents. In contrast, the equal wage scheme focuses agents' attention on the least motivated agents.

> **Result 7:** *The pivotal agent is different between treatments: in the IWT, agents who initially provide low effort align with the high-effort providers over time. In the EWT, agents who initially provide high effort align with the low-effort providers over time.*

2.3.5 The Role of Intentions

So far, we interpret our results as supporting the notion that subjects care about the norm of equity. However, by design our treatments necessarily differ in the number of instruments that a principal has at hand. In the EWT, principals only choose a single wage—whereas principals in the IWT decide on two wages and consequently can tailor reactions individually to agents' preceding choices. Therefore, agents might attribute a different degree of intentionality to principals' decisions: In the EWT, the role of intentions is limited to the level of the wage. The IWT contains an additional element of intentionality because principals also decide on relative wages and consequently whether the equity norm is fulfilled or violated. In light of the literature that stresses the behavioral importance of intentions in situations of reciprocal interaction (e.g., Dufwenberg and Kirchsteiger 2004, Falk et al. 2005), there is thus a potential alternative explanation for our treatment effect. In other words, one might speculate that the difference is not caused by the different frequency of norm fulfillment *per se*, but rather by the additional element of intentionality.[14]

To test this alternative explanation, we conducted an additional control treatment (*wage level treatment* or WLT) that clearly isolates the effect of norm fulfillment on agents' effort choices. As in the EWT, principals in the WLT only choose a single wage. The other agents' wage is then exogenously set by a computer program such that the equity norm is always fulfilled, i.e., agents who exerted a higher effort than their co-worker automatically receive a higher payoff. Importantly, this implies that the fulfillment of the equity norm is not attributable to principals' decisions. Except for this change of the wage-setting institution, the instructions and the ex-

[14] We thank Patrick Bolton and an anonymous referee for pointing this out.

perimental design were identical to the previous treatments. The 72 subjects who participated in the four additional sessions had not previously taken part in the IWT or the EWT.

The specific equity norm implemented in the WLT experiments dictates proportionality between agents' monetary payoffs and efforts. We chose this "equity formula" as it is probably the most prominent formulation of the equity principle (see Section 2.2.2). Given a principal's decision for the low-effort agent, the wage for the high-performing agent is exogenously fixed such that both agents receive the same payoff per unit of effort provided, i.e., $\frac{\pi_{\text{low}}}{e_{\text{low}}} = \frac{\pi_{\text{high}}}{e_{\text{high}}}$ holds. For example, if the principal observes efforts of 2 and 6 and sets the wage for the low-effort provider to be 5, the payoff of this agent is $5 - c(2) = 4$ (compare Table 1). Following the equity formula, the payoff of the high-effort provider will then automatically be set to $\frac{4}{2} \cdot 6 = 12$; which implies a wage of 20 after taking the cost of providing 6 units of effort into account.[15]

The wage-setting institution in the WLT is not meant to be an analog of institutions found in actual labor markets, as it is the case for the IWT and the EWT. It exogenously implements the incentive structure that is endogenously created by principals in the IWT.[16] If we observe similar efforts in the WLT as in the IWT we can rule out intentions as an explanation for the difference between our two main treatments, IWT and EWT.

Figure 2.7 compares agents' mean effort choices over time for all three treatments.

[15]The equity formula leads to counterintuitive implications whenever negative values for the inputs or outcomes are possible. Therefore, if in our experiment the efforts differ and the principal's choice of w_{low} implies $\pi_{\text{low}} \leq 0$, the other agent's wage is instead set such that $\pi_{\text{high}} = \pi_{\text{low}} + 5$. This guarantees that a norm of equity is fulfilled for all possible wage-effort combinations. Nevertheless, the high-effort agent still faces the risk of making losses whenever the low-effort agent gets a negative payoff.

[16]As shown in Result 3, the monetary incentives in the IWT imply that profit-maximizing agents should provide non-minimal effort levels. As a consequence of exogenously implementing these implicit incentives in the WLT, new subgame-perfect Nash equilibria necessarily arise. Our focus of interest, however, rests on the comparison of the observed behavior across treatments rather than on comparing behavior to the game-theoretical equilibrium predictions. For a similar approach of "exogenizing" endogenous incentives to test for the impact of intentions, compare for example Blount (1995), Charness (2004), or Cox (2004).

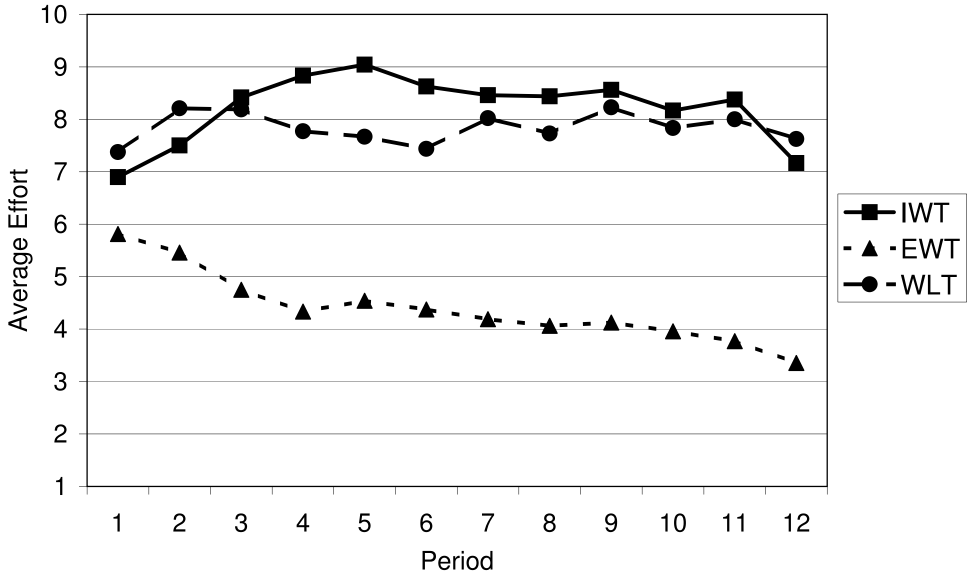

Figure 2.7: *Average effort per treatment.*

As can be seen, the exogenous implementation of the equity norm suffices to elicit high efforts from the agents. The average effort difference between the WLT and the IWT of only 0.36 is insignificant (Mann-Whitney test: $p = 0.83$). Compared to the EWT, efforts are on average 3.44 units higher in the wage-level treatment (Mann-Whitney test: $p < 0.01$). As in the IWT, efforts do not decrease over time in the WLT (Wilcoxon test for periods 1–6 against 7–12: $p = 0.44$). Also the distribution of efforts in the WLT closely resembles the one in the IWT. Under both treatments, the modal choice is the provision of maximum effort. In the WLT, an effort level of 10 is chosen in 46.5% of all cases, compared to 49% in the IWT.

> **Result 8:** *The wage level treatment shows that the treatment difference between the IWT and the EWT is not caused by the fact that principals can set two wages instead of one* per se. *Differences in equity norm fulfillment—independent of intentionality—are the driving force behind agents' performance.*

Taken together, the results from the additional control treatment corroborate our previous findings. They suggest that the observed performance differences are not driven by the differing degree of intentionality across treatments. This, of course, does not imply that intentions are unimportant in general; however in our setup, treatment differences are almost exclusively caused by equity considerations.

2.4 Concluding Remarks

In this chapter, we studied the interaction of gift exchange with different payment modes; more specifically, we analyzed how horizontal fairness concerns of employees affect the effectiveness of gift exchange as a contract enforcement device. In our experiment, one principal is matched with two agents. The principal pays equal wages in one treatment and can set individual wages in the other. The use of equal wages elicits substantially lower efforts and efficiency in spite of similar monetary incentives: exerting high effort pays off under both wage schemes. The strong treatment difference is driven by subjects' preferences for horizontal equity and the fact that the equity principle is frequently violated in the equal wage treatment. This is not the case in the individual wage treatment, as principals set wages almost always in line with the norm of equity. The results of a control treatment show that indeed norm fulfillment *per se* and not different degrees of intentionality are the driving force behind agents' behavior.

Our results have a number of implications, both for the advancement of existing theories and for the design of wage schemes in practice. First of all, while it is well-known that equal wages can distort monetary incentives, in our experiment they are efficiency decreasing *even though* individuals' monetary incentives are qualitatively not affected. Rather, equal wages oftentimes lead to situations which are considered as unfair by the workforce. This holds in particular because agents are heterogeneous and equal wages violate the equity principle whenever workers differ in their performance. It may thus be oversimplifying to argue that equal wages lead to less envy and therefore higher work morale, as it is frequently done in the political discussion.

In this regard, it is doubtful whether strict wage equality can be reconciled with the use of reciprocal gift exchange to enforce incomplete contracts. Our findings suggest that adherence to the norm of equity is a necessary prerequisite for a successful gift-exchange relation. Consequently, the wage setting institution must provide principals with means to account for possible differences in agents' behavior, e.g., to individually reward agents who outperform their co-workers. The performance of agents in the individual wage treatment and in the wage level treatment shows how effective gift exchange can be, as long as horizontal equity concerns are respected: although explicit contract enforcement is absent, 80% of the possible efficiency gains are realized.

In practice, the discretion to fulfill the norm of equity does not have to be in monetary terms. Perks and non-monetary benefits like extra vacation or awards can be useful devices to motivate workers in this context. These instruments become especially important when it is not possible to wage discriminate on a given hierarchical level, e.g., because the firm's internal pay structure, agreements with a union or legistlation dictate wage equality.

The results in this chapter should not be interpreted as arguments against wage equality in general. They rather suggest that equal wages come at a cost that has to be weighed against their potential benefits. For example, equal wages are easier to implement than individual wages, and they may encourage peer monitoring and collaboration. The relative importance of these costs and benefits (and also the impact of the workforce's social preferences more generally) is likely to depend on the details of the institutional setting. These include the production technology, the information structure, and the organizational design of the firm. In this chapter we presented results for one such setting. Our design provides a simple and parsimonious framework that can successively be enriched to study these aspects in future research.

Chapter 3

Identity Changes and the Efficiency of Online Reputation Systems

3.1 Introduction

Online trade usually takes place between strangers, payment occurs before the good is shipped, and legal enforcement of an agreement may be prohibitively costly. This particular constellation generates moral hazard and adverse selection problems, especially on the seller side. Sellers have an incentive to ship a good of lower quality than promised or not to ship at all. Reputation systems such as eBay's Feedback Forum are the most important tool in e-business to address these problems and to induce honest behavior among users. They disseminate information about past conduct of a user to the whole community. After a trade both parties involved can express their satisfaction with the outcome of the transaction by rating their transaction partner. Reputation systems store this information in the user's profile and make it available to all potential transaction partners. In principle, a reputation profile thus allows buyers to distinguish honest sellers from dishonest ones and to interact only with the former.

The successful use of reputation systems in traditional markets with similar moral hazard problems can be traced back at least to the beginnings of long-distance trade in the Middle Ages (Greif 1989, Milgrom et al. 1990). Online reputation in most settings, however, differs from reputation in traditional markets in a very important way. On platforms like eBay or Amazon, reputation is only connected to the virtual identity of a person, i.e., the user name, and not the person itself. After a bad rating a dishonest seller can comparatively easily abandon his old virtual identity and create a new account under a new user name with no reputation attached to it. The implications of this distinctive feature of online reputation for the efficiency of reputation systems are not well understood.

At first sight identity changes seem to severely weaken the disciplinary power of a reputation system. Dishonest behavior can still be punished with a negative rating but these ratings may lose their edge when they can easily be shed by creating a new identity. Real newcomers and sellers who have changed identity after a bad rating become indistinguishable in this context. This argument suggests a high frequency of identity changes accompanied by an increase in opportunistic behavior and lower buyer trust. Theory on the other hand also provides some guidance on why reputation systems might still work effectively (Friedman and Resnick 2001, Ockenfels 2003). If buyers anticipate that dishonest sellers start over as new players they will not trust newcomers or only interact with them at very unfavorable conditions, e.g. low sale prices. Starting as a new seller may then become so costly that cheating and creating a new identity is not profitable anymore. In this case we would observe a high level of seller trustworthiness and no identity changes. Buyer trust would be high in transactions with experienced sellers and low towards new sellers.

In this paper we examine empirically how the option to change one's virtual identity affects the efficiency of markets and the performance of reputation systems in inducing trust and trustworthiness. We study two experimental markets in which buyers play a trust game (Berg et al. 1995) with varying sellers. In this game the buyer can send an investment to the seller which is tripled by the experimenter on the way. The seller can then decide how much of the tripled amount he wants to return to the buyer. Buyers can rate sellers after each transaction; and before deciding whether to interact with a seller, a buyer can see the rating profile of this seller. New players enter the market over time in both treatments. The only difference is that sellers in one market can change their identity (*change treatment*), i.e., erase their rating profile and start over as new players, while in the other market this is not possible (*no-change treatment*). In view of the discussion above we organize our investigation around the following questions: Do sellers use the opportunity to change their identity? If so, does this go along with an increase in opportunistic behavior? How is buyer trust affected? In particular, are new members treated differently when identity changes are possible?

These questions are of high practical relevance given widespread reports that dishonest behavior remains an important problem for online trade. For example, in the 2005 reports prepared by the National Consumer League[1] and the Internet Crime Complaint Center[2] online auction fraud is by far the largest source of complaints with 42% and 63% respectively of the complaints received on internet fraud in general. Recent studies (Dellarocas and Wood 2008, Bolton et al. 2008, Jin and Kato 2006) also indicate that the low percentages of neutral and negative ratings cited for eBay grossly understate the true extent of problematic transactions since users may be reluctant to give negative rating for fear of retaliatory ratings. Dellarocas and Wood (2008) estimate the true figure of mildly or very dissatisfied buyers at up to 21%. Dishonest behavior is also an important obstacle to the continued growth of e-business. According to a survey commissioned in 2003 by the European Commission 21% of the EU 15 population abstain from buying online because they do not trust the internet.[3]

It is difficult to determine whether these problems (also) arise because sellers can shed a negative reputation comparatively easily. So far there is little evidence on the effects of identity changes, mainly because identity changes are not directly observable in field data. It is also difficult to observe to which extent consumers do not buy online for fear of being cheated. In addition, even if one could observe identity changes, comparing the same reputation system with and without the possibility to change one's identity would be extremely difficult in the field. We therefore

[1] http://www.fraud.org/2005_Internet_Fraud_Report.pdf

[2] http://www.ic3.gov/media/annualreport/2005_IC3Report.pdf

[3] http://ec.europa.eu/public_opinion/archives/ebs/ebs_201_executive_summary.pdf

use a controlled laboratory experiment to address our research questions. Though a modest step, we nevertheless believe it is an important one in understanding how reputation affects the emergence and functioning of markets given the observability problems with field data.

The main findings can be summarized as follows. First, the reputation system in the no-change treatment is remarkably successful in inducing seller trustworthiness and buyer trust. Buyers condition their investment on seller reputation and weigh negative ratings more strongly than positive ones. On average buyers make a substantial profit on their investment. Opportunistic sellers, who return little or nothing to the buyer, do not make a higher profit than those returning significant amounts to the buyer. Second, when sellers have the opportunity to erase their rating profile they do use it. Identity changes mainly occur when the number of negative ratings becomes larger than the number of positive ratings. Third, sellers in the change treatment seem to choose between two kinds of behavior - being trustworthy and maintaining a fixed identity, and behaving opportunistically and changing identity multiple times. As a consequence, identity changes go along with a lower overall degree of trustworthiness. Interestingly though, trustworthiness is still high enough to make investing profitable for the buyers even when rating profiles can be erased. Fourth, buyer trust is also significantly lower in the change treatment. We also observe a greater incidence of buyers dropping out of the market by not investing at all. As suggested by theory the decrease in trust is driven by lower trust in new sellers. Finally, opportunistic players nevertheless earn more than their counterparts who share more equally with the buyer in the change treatment. A substantial fraction of the latter players display behavior which is consistent with an intrinsic preference for trustworthy behavior.

We believe that this work contributes on several fronts. Our results for the no-change treatment complement the findings of previous laboratory experiments on reputation systems without identity changes. Keser (2002) shows that providing more information can improve the efficiency of a reputation system, Bolton et al. (2004) studies whether the source of information matters by comparing a market with a reputation system to one with long-term relationships. Both studies find that a reputation system leads to high levels of trust and trustworthiness. We extend this finding to an environment in which new players enter the market over time. The mere fact that there are new players need not affect the efficiency of a reputation system as long as buyers can be sure that a new seller really is a person who has just entered the market.

The results for the change treatment demonstrate that the efficiency of reputation systems in inducing trust and trustworthiness is reduced if sellers have the opportunity to change their identity. While the negative effects of identity changes have long been discussed theoretically in the literature (e.g. Dellarocas and Wood 2003) this paper provides controlled evidence and establishes a *causal link* between identity changes and a higher incidence of dishonest behavior. This is the main contribution of our paper.

In addition, the controlled environment of the laboratory allows us to take a closer look at the mechanisms through which identity changes influence market outcomes. Several of our specific findings which we can causally attribute to potential identity changes are consistent with observations from online platforms. For example, Ockenfels (2003) finds a very low percentage of sellers with more negative than positive ratings in a sample from the platform half.com. In the same sample new sellers also ask for lower prices than experienced sellers. In our experiment the differential treatment of new sellers in the two treatments due to higher uncertainty about their trustworthiness in the change treatment is one of the main driving forces of lower buyer investment. Our results also show that a substantial part of the negative effects of identity changes may arise from buyers dropping out of the market or not entering it in the first place and thus complement survey evidence indicating that large efficiency losses occur because consumers do not trust online trade(rs). Finally, lower overall trustworthiness in our change treatment is due to a group of sellers who follow a "take the money and change your identity" strategy. Jin and Kato (2006) find evidence for such behavior as a byproduct of their investigation on the relation between price, (claimed) quality and reputation in the eBay submarket for unrated baseball cards. They observe several sellers who built up their reputation first and then fail to deliver on a large number of parallel auctions.

While identity changes increase the incidence of opportunistic behavior our results also demonstrate the ability of reputation systems to make investments worthwhile even when identity changes are possible. Although ostracism by bad reputation is not feasible trustworthy sellers still have the opportunity to distinguish themselves by a positive reputation. Monetary and non-monetary rewards to a positive reputation seem to be deemed sufficiently high to prevent wide spread opportunistic behavior.

The rest of the paper is structured as follows. Section 2 describes the experimental design. Section 3 reports and discusses our results. Section 4 concludes.

3.2 Experimental Design

3.2.1 Trust Game and Reputation System

We use the trust game introduced by Berg et al. (1995) to mimic the essential features of online trade in our experiment. The trust game is played by two players, which we refer to as buyer and seller in our setup. Both players have an endowment of 10. In the first stage the buyer can send an amount between 0 and 10—the *investment*—to the seller. On the way to the seller this amount is tripled by the experimenter. In a second stage the seller can then decide how much (if anything at all) of the tripled amount he wants to send back to the buyer. The trust game nicely captures the moral hazard problem of online trading. There are efficiency gains to be realized because the amount sent is tripled. However, this requires that the buyer sends money without a guarantee of receiving anything in return from the seller. The buyer's (risky) investment can therefore be taken as a measure of his trust while the amount that the seller sends back (as a share of what he received) reflects the latter's trustworthiness. More precisely, our measure of trustworthiness is *return on (buyer) investment* which we define as (Amount Received/Amount Sent) - 1.

In our experiment subjects play the trust game described above, half of them in the role of buyers and the other half in the role of sellers. Every participant keeps his role throughout the entire experiment which lasts 20 rounds. Buyers and sellers are randomly rematched after every round. The matching mechanism ensures that players do not play with the same person in two consecutive rounds and this is common knowledge.

Based on Keser (2002) we enhance the basic trust game with a reputation system in the following way. At the end of each round after the seller has decided how much to return to the buyer, the latter rates this decision. He can give the seller a positive, neutral or negative rating. If there was no interaction, i.e., the buyer did not send anything to the seller, no rating can be made. The rating is stored in the seller's rating profile.[4]

At the beginning of each round, the seller sees his rating profile. It lists the last rating and the total number of positive, neutral and negative ratings. In the no-change treatment every seller keeps his rating profile for the whole experiment and therefore is simply informed about his current profile. Then the buyer sees the seller's rating profile and decides how much money to send to the seller. In the change treatment the seller can decide *before* the buyer sees his rating profile whether he wants to erase the profile and start as a new player. This profile erasure is what we refer to as an identity change. If the seller keeps his rating profile it is then shown to the buyer. If the seller *erases* his profile, the buyer will only be given the information that he is matched with a *new* player before making his investment decision.

If buyers in real online markets face a new seller they are unable to tell whether this person really is a newcomer or a (dishonest) seller who has abandoned his old account and created a new one. To achieve this crucial feature in our laboratory experiment we let new subjects enter the experiment over time. Only six subjects (3 buyers, 3 sellers) start the experiment right from the beginning in every session.

[4]When we conducted this experiment sellers could also rate buyers on *eBay*. We abstracted from this option since there is no scope for opportunistic buyer behavior in our setup (and only to a very limited degree on *eBay*). Our experiment was ahead of its time in the sense that *eBay* later on changed its rules due to wide-spread complaints about (threats of) retaliatory ratings. While buyers now can still rate sellers, sellers can only give positive ratings to buyers. Seller complaints, e.g., about slow payment are now handled via a non-public mechanism.

The others enter in the course of the experiment.[5] Subjects know that new players enter the experiment but not when and how many. When faced with a new seller buyers thus cannot discern whether the seller really is a new player or has just erased his rating profile.[6]

The experiment was programmed using z-Tree (Fischbacher 2007) and conducted at the BonnEconLab at the University of Bonn. Eight sessions with 12 subjects each were played of each treatment yielding a total of 192 participants. At the beginning of each session participants were seated in separate cubicles. Before the experiment started the experimenter read out the instructions and answered all remaining questions privately in a low voice.[7] All subjects were undergraduates from different fields recruited via announcements posted on campus. The experiment lasted approximately one hour and subjects earned 13.16 Euro on average.

3.2.2 Behavioral predictions

In the trust game presented above, the seller does not send anything back to the buyer when agents only care about material payoffs. In anticipation of this the buyer does not send anything to the seller in the first place. In the Nash equilibrium the potential efficiency gains are thus not realized. A substantial literature in economics studies reputation building and its potential to improve efficiency in similar situations.[8] Independent from whether identity changes are possible or not the introduction of a reputation mechanism does not change this result as long as the game is finitely repeated. Based on this model we would therefore expect no

[5]The timing of the entry of new players is the same for all sessions. One buyer and one seller each enter in rounds 4, 8 and 13, respectively.

[6]Players who entered the experiment at a later stage had to do a paid real effort task which consisted of finding errors in a short text. We introduced the real effort task to keep subjects busy and to avoid income effects. Real effort payments were calibrated such that earnings from the real effort experiment were not significantly different from those of the participants who played the trust game right from the beginning.

[7]An translated version of the instructions can be found in the appendix.

[8]See Dellarocas and Wood (2003) for an excellent survey of this literature and its relevance for online reputation systems. Bolton et al. (2004) provide a very intuitive account of different modeling approaches and their consequences.

differences between our treatments.

Previous experiments with finite repetitions of a trust game (Keser 2002 and Bolton et al. 2004) nevertheless observe that a reputation system without identity changes substantially reduces seller moral hazard and improves efficiency. This finding can be rationalized by models which assume different types of players (Kreps et al. 1982).[9] If, for example, some sellers are intrinsically trustworthy it can be profitable for the opportunistic sellers to camouflage as trustworthy at least until the final rounds of play. Another potential explanation for the success of reputation systems in previous experiments is that subjects are myopic in the sense that they fail to fully account for the finite nature of the experiment.[10] For an infinite time horizon, Kandori (1992) and Okuno-Fujiwara and Postlewaite (1995) demonstrate that the seller moral hazard problem described above can be completely solved with the help of a reputation system - even in large communities in which agents only interact infrequently and agents enter and exit over time. The simple intuition behind this result is that sellers who behave opportunistically can be punished later on by other members of the community if past conduct is common knowledge.

The only models which explicitly consider the effects of identity changes (Friedman and Resnick 2001 and Ockenfels 2003) stand in the tradition of this earlier, infinite time horizon work. With identity changes direct punishment of opportunistic players is not feasible anymore since players with a bad reputation can simply reenter the game as new players with no reputation at all. To prevent defection and subsequent identity changes these have to become unprofitable. This can be achieved by making new players "pay their dues", i.e., all new players receive unfavorable treatment from experienced players. In Friedman and Resnick (2001), for example, which uses the prisoner's dilemma to model the moral hazard problem, new players pay their dues by cooperating while the experienced player defects. In an online auction setup, sellers could pay their dues by achieving a lower sale price when they are new. Both models show that "pay your dues" strategies lead to a

[9]Brown et al. (2004) provide an example in which some players are inequity averse.

[10]Formally, assuming that some players are intrinsically trustworthy is similar to assuming that these players have an infinite time horizon.

welfare maximizing (second-best) equilibrium with high levels of cooperation when there are costless identities changes.

Even though these models differ from our setup with respect to the time horizon and the exact formalization of the moral hazard problem, the basic mechanisms they describe may nevertheless serve as an orientation for our hypotheses, in particular, in view of previous experimental results. For our setup we expect the following. First, there will be no sellers with a reputation worse than that of a new seller when identity changes are possible. Second, new sellers in the change treatment will pay their dues by receiving lower investments from the buyers compared to the no-change treatment. As a consequence, overall investment and thus efficiency will be lower in the change treatment. Third, the models predict that identity changes are unprofitable and therefore never occur in equilibrium. However, if some sellers do deviate from this equilibrium strategy by changing their identity, this off-equilibrium behavior will be characterized by opportunistic behavior preceding the identity change. Trustworthiness will therefore be equal or lower in the identity change treatment.

3.3 Results

In this section we first present the results for the no-change treatment to demonstrate the basic mechanisms of a reputation system at work. We then analyze the effect of identity changes by contrasting the results for the change treatment with those for the no-change treatment. The results of this treatment comparison are organized around three main questions. First, we check whether sellers use the opportunity to create a new identity at all and if so under which circumstances. Second, we analyze the effects of identity changes on seller trustworthiness. Finally, we discuss the resulting differences in buyer trust between the two markets.

3.3.1 A reputation system at work-the no-change treatment

A reputation system's efficiency can be determined by two related measures. Ultimately, it should induce buyers to trust and sellers to be trustworthy. The degree to which these two conditions are fulfilled determines to which degree potential gains from trade will be realized. As described above, we use buyer *investment* to measure trust and *return on (buyer) investment* as a measure of sellers' trustworthiness. Note that a return of zero means that the buyer got back exactly the amount that she sent, a positive (negative) return implies that the buyer made a profit (loss) on her investment. A great advantage of using the trust game for our study is that there is a substantial amount of data on returns on investment in a variety of settings. We can thus compare the no-change treatment to the benchmark case of a one shot trust game and to previous results for trust games with a reputation system without identity changes. In his review of the literature, Camerer concludes that return on investment is usually around zero in one shot trust games (Camerer 2003). On the other hand, previous experiments on reputation systems without identity changes and without new players entering over time (Bolton et al. 2004 and Keser 2002) find that the introduction of a reputation system raises trustworthiness to a level which makes investing profitable. Our no-change treatment replicates these findings, buyers on average make a profit of 61% on their investment. In more than 80% of the cases in which buyers invest they make a profit (596 out of 720).

The reputation system is also successful in inducing buyer trust. Remember that buyers could send an amount between 0 and 10 to the seller. Average investment is 7.04 and thus substantially higher than Nash equilibrium predictions and investments typically observed in one-shot trust games. Keser (2002), for example, observes an average investment of 3.91 in a one-shot trust game played with changing partners.[11] Investing the entire endowment of 10 is by far the most frequent choice with 35% of all investment decisions. Since investment is tripled the high average investment also generates a substantial surplus. However, this surplus is

[11]Note, however, that a high return on investment is not sufficient to warrant high investments. The key prerequisite for high investments is that buyers *believe* that sending money to the seller is profitable, i.e., that the reputation system successfully prevents dishonest seller behavior.

not distributed equally. In line with the trust game literature buyers gain 14.49 on average, while sellers make 19.58 (two-sided Wilcoxon test p=0.012). Both buyers and sellers are nevertheless better off compared to their initial endowment.

Why do we observe such a high degree of efficiency in the no-change treatment? Reputation systems operate through two basic mechanisms. First, dishonest seller behavior should translate into a bad reputation. Second, a bad reputation should lead to lower buyer trust. The success of the reputation system in our no-change treatment can be traced back to these mechanisms. To analyze the first mechanism in our setup we calculate the Spearman rank order correlation between the return on investment and the rating given for this return. We code a good rating as 1, a neutral rating as 0, and a negative rating as -1. This coding is also used on eBay to calculate aggregate reputation. The correlation is substantial (r=0.77) and highly significant (p<0.001). Lower returns thus lead to lower ratings for sellers.

To check whether a bad reputation in turn leads to lower trust we analyze how the reputation profile a buyer sees at the beginning of a period influences his investment decision. This profile shows the number of good, neutral and bad ratings, or whether a seller is new. Column 1 in Table 3.1 shows a random effects panel regression of buyer investment on the percentage of good and bad ratings of the seller, a dummy variable for new sellers and a time trend. Since all players were new in the first period and had no rating profile we exclude this period from the regression. Standard errors are adjusted for clustering at session level. The results of the regression indicate that buyers take a seller's reputation into account when deciding how much to invest. The coefficients for the percentages of good and bad ratings are significant and have the expected sign. Bad ratings, however, seem to have a stronger impact. On average, a 10% increase in the percentage of positive ratings increases investment by 0.17 whereas a comparable increase in the percentage of negative ratings reduces it by 0.31. A Wald-test for the equality of coefficients confirms that the absolute difference is significant (p<0.001). This is in line with findings for laboratory data and field data from eBay (e.g. Bolton et al. 2004, Lucking-Reiley et al. 2007).[12] In

[12]See Bajari and Hortaçsu (2004) or Bolton et al. (2008) for more references on the growing number of field studies on the influence of reputation information on outcomes in online trade.

Buyer Investment	Coeff./Rob.se
percentage good	1.721***
	(0.636)
percentage bad	-3.123***
	(0.470)
new	-0.177
	(0.642)
period	-0.081**
	(0.039)
Obs.	760
R^2	0.224

Table 3.1: *Investment regression. Observations from the first period were not included in the analysis. Robust standard errors adjusted for clustering at the session level are given in parentheses. "new" and "change" are dummies equal to 1 for new sellers and observations from the change treatment respectively. Significance at the 10%, 5% and 1% level is denoted by *, **, and ***.*

addition, we observe a significantly negative time trend. The new player dummy is not significant implying that new players are not treated differently than players with a completely neutral reputation.

Ideally, a reputation system prevents that opportunistic seller make a higher profit than honest ones in the long run. Even if opportunistic sellers receive lower ratings and therefore lower investments as in our no-change treatment it is not obvious that this will be the case. It might still be more profitable for the sellers to return little or nothing of a smaller investment than to send back a larger share of a larger investment. To prevent this, the negative effect of a bad reputation on investments must be large enough to outweigh the benefits of returning less to the buyer. Our final analysis for the no-change treatment is therefore a simple test of whether bad behavior is profitable. We classify a seller as "bad" ("good") if his aggregate reputation at the end of the experiment, i.e., the number of positive ratings minus the number of negative ratings, is negative (positive). On average good sellers earn 19.32 and bad sellers earn 20.04, but the difference is not significant (two-sided

Wilcoxon test p=0.21).[13] The reputation system in the no-change treatment thus not only induces high levels of trust and trustworthiness but also achieves another important property, namely that bad behavior does not pay off. In the following we study to which degree the findings of this section have to be modified if sellers can erase their reputation.

3.3.2 Understanding the impact of identity changes

Analyzing change behavior

Our data suggest that the mechanisms which could prevent sellers from changing their identity are not strong enough in our setup. Sellers do use the opportunity to change their identity. On average they erase their reputation profile every 3.5 rounds, i.e., in 28.4% of the rounds (209 out of 736).[14] This numbers hides a considerable degree of heterogeneity in change behavior. On a session level the average frequencies with which sellers erased their reputation profile range from 4% to 53%, on an individual level from 0% to 95%. However, behavior is far from arbitrary, rather subjects seem to follow one of two distinct strategies. The majority of subjects can be classified into two groups - sellers who maintain their identity ('keepers') and those who use multiple identities ('changers').[15] 44% of sellers are keepers and 25% are changers.

These numbers suggest that subjects use identity changes strategically. This impression is corroborated by an analysis of *when* sellers erase their reputation. A strategic seller erases his reputation once it becomes worse than the reputation of a new seller. Arguably this is the case when the aggregate reputation becomes zero or

[13]Unless otherwise noted all non-parametric tests use session averages as independent observations. One-sided tests are used for directed hypotheses (see previous section).

[14]In total 98 rounds were played in each role in each session (3 players in each role played 20 rounds, 1 player each played 17, 13, or 8 rounds respectively). For each role there are thus (60+17+13+8)*8= 784 rounds per treatment. For the analysis of identity change behavior we exclude the first round of each seller since identity could not be changed in this round. This leaves us with 736 rounds.

[15]We classify all sellers with a maximum of one ("tryout") ID change in the entire experiment as keepers and those who change their identity at least every second round as changers.

negative, i.e., the number of negative ratings is greater than the number of positive ratings. 63% of the pseudonym changes occur when the aggregate reputation is -1, another 25% when it is 0. As a consequence buyers in the change treatment only meet a seller with a negative aggregate reputation in 6% of the rounds while their counterparts in the no-change treatment make this experience in 32% of the rounds.

Effects on seller trustworthiness and ratings

We hypothesized above that trustworthiness in the change treatment will be lower if some sellers change their identity. Figure 3.1 displays the cumulative distributions of return on buyer investment for the two treatments. From the graphs it is obvious that return on investment is indeed lower in the change treatment. A one-sided U-test confirms this impression ($p = 0.029$). Returns on investment of 100% and above are more frequent in the no-change treatment, for lower returns the opposite holds. The most extreme case in which a buyer makes an investment but the seller does not return anything occurs 63 times (8.8% of all investments made) in the no-change treatment but 102 times (15.3% of all investments made) in the change treatment. The buyer risk of making a loss on an investment is 36% in the change treatment and 17% in the no-change treatment.

The reputation system is thus less successful in inducing trustworthy behavior in the market in which participants can erase their reputation. A key question, however, is whether the system is still *successful enough* in the sense that it can curb opportunistic behavior at least to the degree that interacting is still profitable for buyers. In the no-change treatment buyers on average make a profit of 61% on their investment. Interestingly, this desirable trait can be sustained even when identity changes are possible. Average return on investment in the change treatment is 36% and thus still substantially above the zero returns typically observed when there is no reputation system at all.

Are sellers who behave opportunistically and then change their identity driving the lower trustworthiness in the change treatment as suggested by the models? For a first impression Figure 3.2 plots the relative frequency of identity changes for

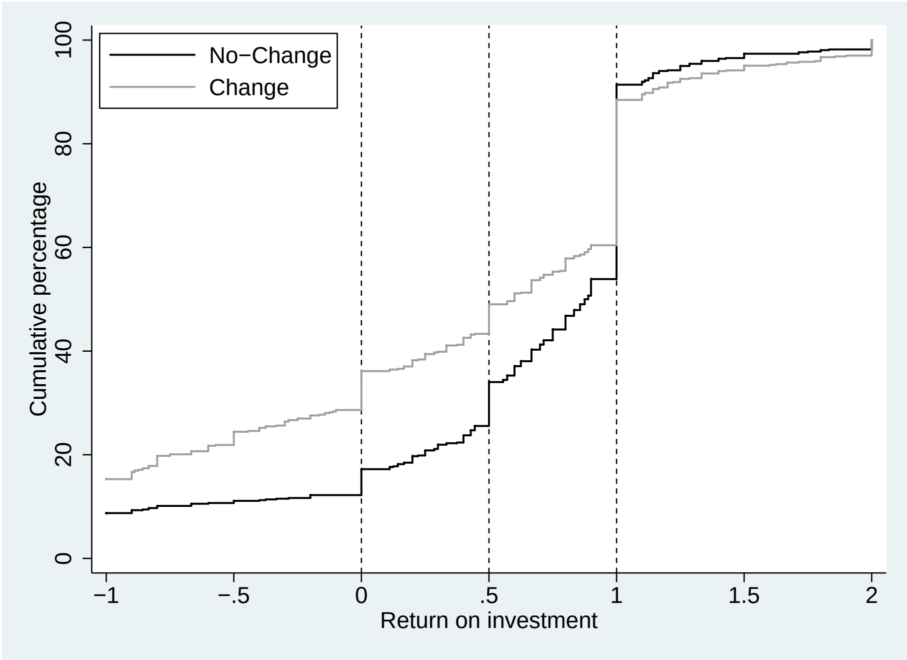

Figure 3.1: *Cumulative distribution of returns on investment*

the different sellers on the X-axis, while the Y-axis displays the average return on investment by each seller. Sellers with a high frequency of changes return less than those who never or only rarely change their identity. If sellers use identity changes strategically the relative amount a seller returns should be particularly low in the round before an identity change. A seller who intends to erase his reputation should be unconcerned about a negative rating and can therefore keep most or all of the buyer's investment for himself. This is borne out by the data. In the round before a seller changes his identity the average return on buyer investment is actually negative (-45%), whereas it is positive (+46%) when the rating profile is not erased.

The cumulative distribution in figure 3.1 also shows that, besides the selfish choice of returning nothing (return on investment=-1), three values of return on investment — 0, 0.5, and 1 (marked by the dashed lines) — are chosen particularly frequently. Interestingly, these returns correspond to different degrees of fairness or fairness norms. A return of 0 implements what could be called a minimal fairness norm; the seller keeps all the surplus but the buyer at least does not make a loss on his investment. The second salient choice is an equal split of the tripled amount sent without taking the endowments into account, implying a return on investment

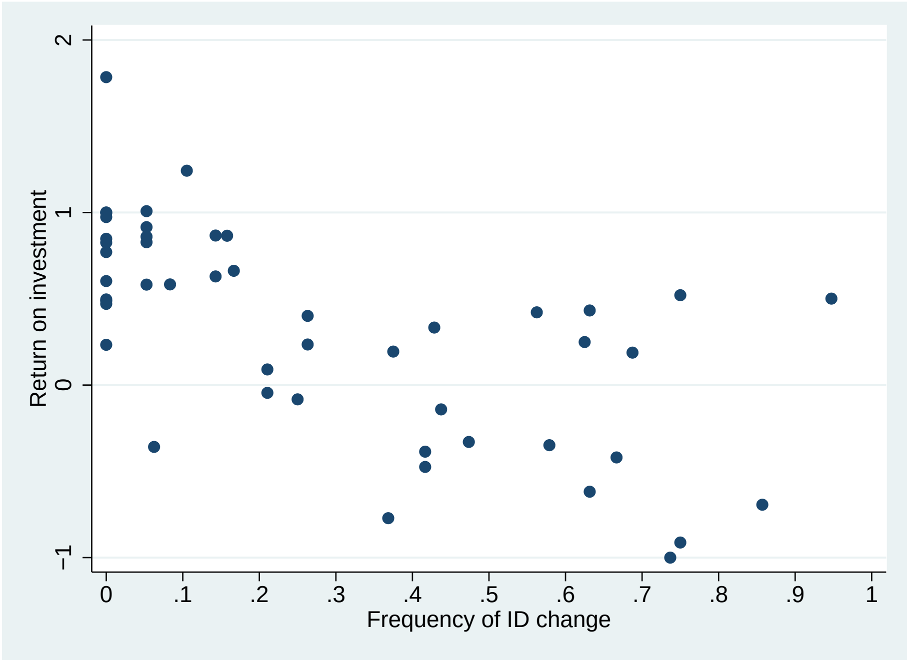

Figure 3.2: *Return on investment and frequency of identity changes*

of 50%. Finally, a return of 100% which implies equal payoffs is the most common return decision in both treatments. For most sellers in both treatments an equal split of payoffs seems to be the natural fairness norm.

Figure 3.3 suggests that the same is true for buyers. The figure depicts the relative frequency of good, neutral, and negative ratings for returns on investment of -1 (selfish), 0 (no loss for the buyer), 0.5 (equal split of tripled amount) and 1 (equal payoffs) for both treatments. While the ratings for the former three return decisions are overwhelmingly negative or neutral, implementing equal payoffs is rewarded with a good rating in about 80% of the cases in the no-change and the change treatment. Overall, rating standards do not seem to differ between treatments.[16]

The figure also suggests that relative return seems to have a strong influence on the rating given. This impression is confirmed by the significant Spearman rank order correlation between the two ($r=0.77$, $p<0.001$). The correlation coefficient has the same value as in the no-change treatment. Given that lower returns on investment translate into lower ratings in both treatments it is not surprising that

[16]This is confirmed by an ordered probit regression of ratings on relative returns and a term for the interaction of treatment and return on investment (not reported here).

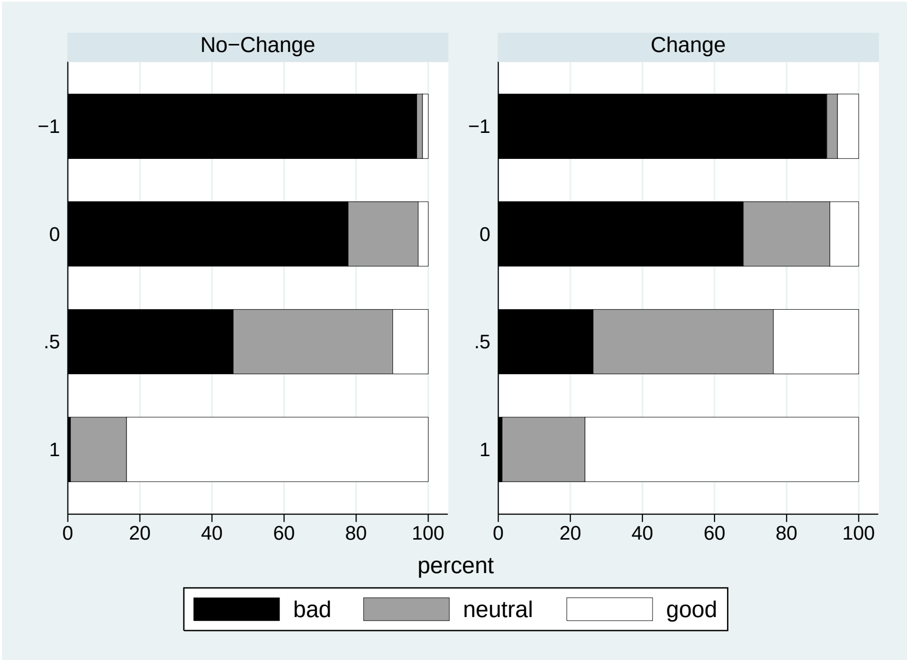

Figure 3.3: *Frequency of good, neutral, and bad ratings for selected returns on investment*

we on average observe lower ratings in the change treatment. While in the no-change treatment negative ratings account for 30% of all ratings, this figure is 37% in the change treatment. The opposite is the case for positive ratings (47% vs 41%). The average rating in the no-change treatment is 0.17, and 0.04 in the change treatment.

The analyses above show that for sellers who keep their identity the rating profile is informative about their past behavior in both treatments. However, in the change treatment rating profiles can be erased. Making inferences about new sellers therefore becomes much more difficult for the buyers. Assuming that a new seller will behave like the average experienced seller seems justified in the no-change treatment where entry into the experiment is determined exogenously. For example, return on investment does not differ significantly between sellers with a rating profile (61%) and new sellers (46%, two-sided Wilcoxon test p=0.40) and a loss for the buyer even occurs more frequently when facing an experienced seller—in 9% of the cases versus 4% for new sellers.[17] Sellers with the label "new" in the change treatment, however,

[17]Given that all players were new players at the start of the experiment and the label could thus not carry any information we exclude the first period from the analyses in this paragraph. The qualitative results are robust to including these observations.

are a mix of truly new sellers and (potentially opportunistic) sellers who erased their rating profile. Buyers in our experiment as well as in real life cannot distinguish between these two groups, but the controlled environment of the laboratory allows us to disentangle them. In fact, only around 10% of the new sellers in the change treatment are subjects who just enter the experiment if we exclude the first round. Average return on investment by these real new sellers is 40%. In contrast, the remaining 90% of the sellers labeled as "new", i.e., those who changed their identity, display an average return on investment of -10%. As a consequence we observe a significant difference in the behavior of sellers labeled as "new" as a whole and those with a reputation profile. Average return on investment by the former is -4%, i.e., buyers on average make a loss when they interact with new sellers in the change treatment. Interactions with sellers with a reputation on the other hand yield a profit of 52%. This difference is significant (two-sided Wilcoxon test p=0.017). The average frequency with which buyers make a loss is also twice as high when they face a new seller compared to one with a reputation profile (22% vs 46%).

Consequences for buyer trust

Above we hypothesized that lower investment towards new sellers in the change treatment would lead to lower overall investment and efficiency. Such behavior would be warranted empirically on an ex-post basis given that the possibility to change one's identity leads to lower trustworthiness among sellers with the label "new". Investments towards new sellers yield a negative return, even though, overall, investing is profitable in the change treatment. How is buyer trust affected? Figure 3.4 shows average buyer investment over time for the two treatments, including those instances in which the buyer did not invest anything at all. Investment is lower in the change treatment right from the beginning but the difference becomes more pronounced in the second half of the experiment. On average investment in the change treatment is 15% lower than in the no-change treatment (6.00 vs 7.04, one-sided U-test p=0.0465). Since the socially optimal level of investment is 10 the lower investments in the change treatment imply a loss of efficiency. In both treatments we also observe a typical end-game effect, i.e., a decline in investment

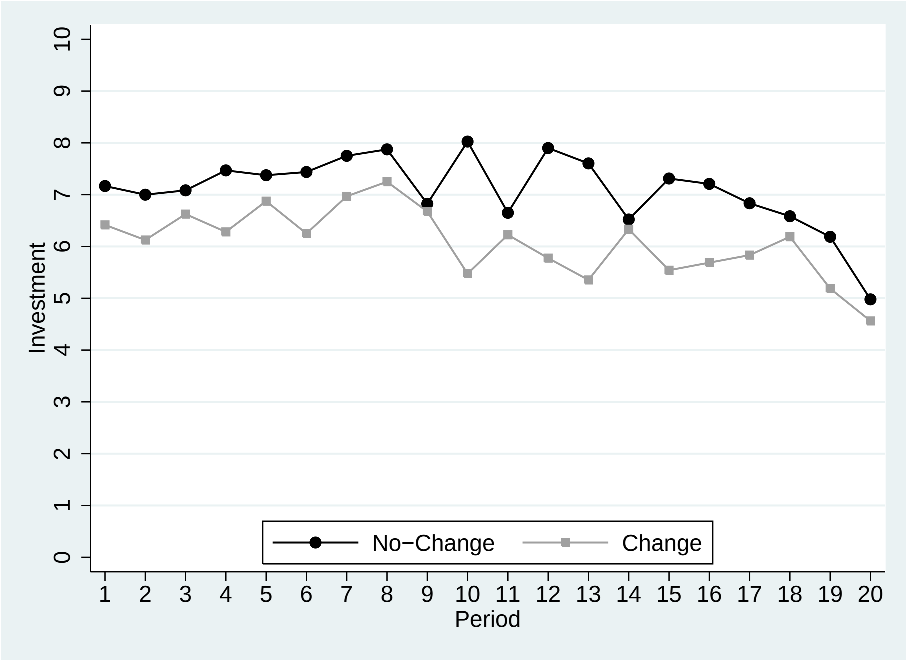

Figure 3.4: *Investment over time*

towards the end of the experiment.

Figure 3.5 displays the relative frequencies of the different investments for the two treatments. Investments from 1 to 5 are more frequent in the change treatment, amounts from 6 to 10 are invested more frequently in the no-change treatment. However, for those cases in which a positive investment was made the difference between the treatments is smaller than overall (7.66 vs 7.05). Survey evidence suggests another important source of inefficiency. Lower trustworthiness may drive buyers out of the market or prevent them from entering it in the first place. These decisions are usually not observable in field data. In our setup we can quantify the frequency of zero investments and whether potential identity changes increase this share. Buyers in the no-change invest zero in 128 cases or 8% of all rounds. Strikingly, buyers in the change treatment do not invest anything almost twice as often, namely in 234 or 15% of all rounds (one-sided U-test p=0.029). Our results suggest that these rather indirect costs of identity changes make up a large share of the associated efficiency losses.

To refine our understanding of why we observe lower investments in the change treatment we plot investment towards new sellers and those with a reputation profile

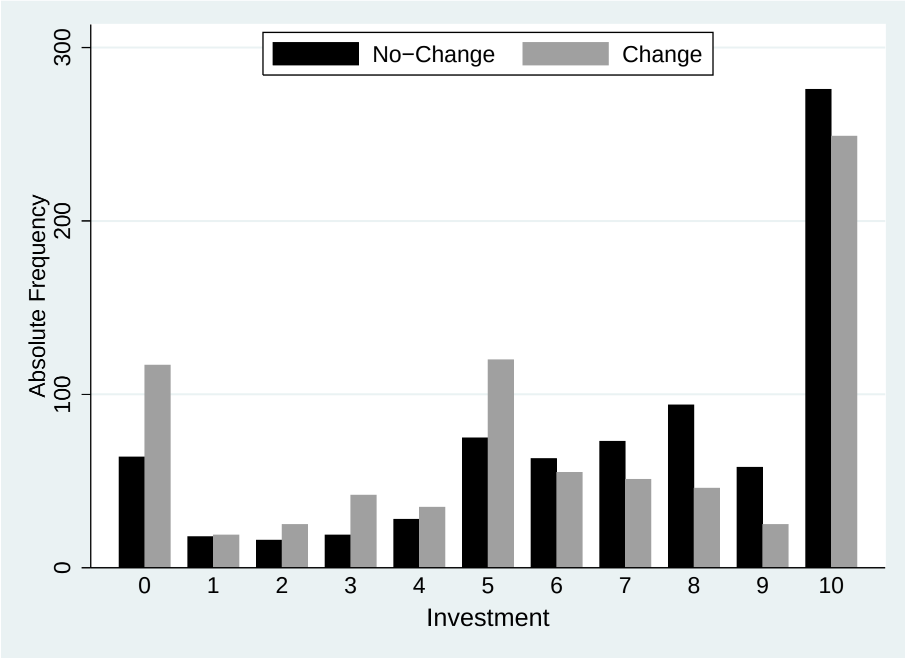

Figure 3.5: *Frequency of investment levels*

in Figure 3.6. While there is no significant difference between investment towards new sellers and the investment towards sellers with a reputation in the no-change treatment (7.38 vs 7.02), new sellers receive significantly less than players with a reputation in the change treatment (4.88 vs 6.48, two-sided Wilcoxon test p=0.027).[18] Interestingly, there is no significant difference between investment towards experienced sellers across treatments. New sellers, however, receive on average 34% less in the change treatment (two-sided U-test p=0.001). These differences suggest that the perception of newcomers does indeed change when identity changes are possible. Because truly new sellers and those who have switched identity after a bad rating become indistinguishable in this context sellers with the label"new player" are met with distrust in the change treatment. They have to pay their dues by receiving lower investments.

In the models no player changes his identity in equilibrium because the discrimination of players with a bad reputation and new players is sufficiently strong to make cheating and creating a new identity unprofitable. While we do observe identity changes in the change treatment the question remains whether they are a

[18]We again exclude the first period from this analysis.

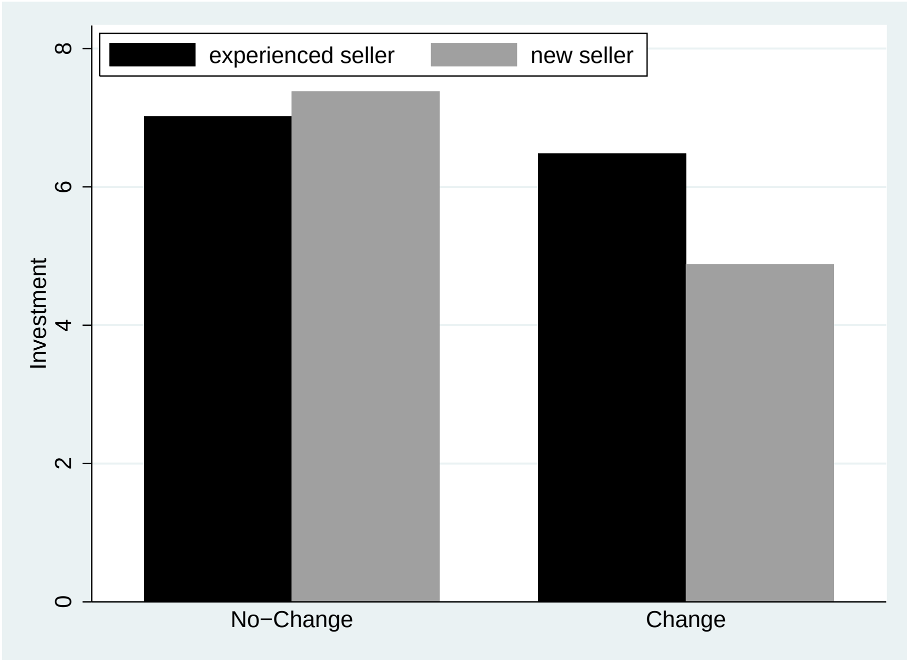

Figure 3.6: *Investment towards new and experienced sellers*

profitable strategy. After analyzing returns on investment and investment behavior we can now compare payoffs to the two strategies described earlier, i.e., maintaining a fixed identity ('keepers') or changing identity frequently ('changers'). Above we showed that identity changes are associated with low trustworthiness. Keepers on average provide a return on investment of 74%, changers only return -14%. Keepers earn 18.48 and changers earn 20.00 overall. We have observations for both groups in six out of eight sessions—in two sessions there are no changers. In the sessions for which we observe both groups changers earn significantly more than keepers (20.15 vs 18.43, two-sided Wilcoxon test p=0.028).[19] The reputation system in the change treatment thus does not succeed in making good behavior at least as profitable as bad behavior. Since opportunistic players camouflage as new players, all new players receive lower investments. But the discount for new sellers is not large enough to offset the lower returns by opportunistic sellers. As a consequence there is no

[19]Classifying sellers who acquired more (less) positive than negative ratings in the course of the experiment as good (bad) sellers and conducting the same analysis as for the no-change treatment leads to similar results. In this case we have observations for both groups in all 8 sessions. Good sellers earn 18.29 on average, bad sellers make 20.86. In contrast to the no-change treatment this difference is significant (two-sided Wilcoxon test p=0.02).

significant difference between seller profits in the two treatments (19.58 vs 19.67, two-sided U-test p=0.92). The efficiency losses arising from lower investments in the change treatment are entirely borne by the buyers who earn significantly less in the change treatment (14.49 vs 12.33, two-sided U-test p=0.035).

Given the significant payoff difference between keepers and changers in the change treatment one might ask why not more subjects choose to behave opportunistically and change their identity afterwards. Analyzing behavior in the last round suggests that a substantial fraction of sellers does not act purely out of strategic concerns. In the last round reputational concerns do not play a role anymore. A purely strategic seller will therefore not return anything. Nevertheless 11—all of them keepers—out of the 34 sellers who receive an investment in the last period provide a return on investment for the buyer of 50% or more. At least half of the keepers thus seem to have an intrinsic preference for trustworthy behavior. By themselves these are not enough to make investment worthwhile, return on investment is -23% in the last period. However, the reputation system succeeds in creating additional strategic incentives for sellers which are sufficiently strong to make investment on average profitable for buyers.

3.4 Conclusions

In this paper we studied the influence of identity changes on the efficiency of reputation systems by comparing two markets. In both markets new buyers and sellers enter over time, but in one market sellers cannot erase their reputation while in the other market they can. We find reduced seller trustworthiness and buyer trust in the latter market. A substantial fraction of sellers uses the opportunity to behave opportunistically and change their identity afterwards. Buyers cannot distinguish between these players and those who are really new and react with lower investments towards players with the label "new player". This reduction, however, is not large enough in the sense that opportunistic behavior with frequent identity changes still is a profitable strategy in the change treatment. Nevertheless the reputation system succeeds to keep trustworthiness at a level which implies positive returns on investment for the buyers even when identity changes are possible.

An important implication of our findings for online market design is that any measure which reduces the uncertainty connected with the label "new player" caused by identity changes and thus increases the ability to distinguish between good and bad sellers can increase efficiency.[20] An obvious suggestion is to introduce fixed identifiers, another less invasive measure is to give sellers the opportunity to have their identity verified by a third party.[21] However, Friedman and Resnick (2001) also point out that users may value a certain degree of anonymity. The results in this paper thus need not be interpreted as arguments against allowing identity changes in general. They rather suggest that doing so comes at a cost which has to be weighed against the potential benefits of increased anonymity. These are likely to depend on the environment in which the reputation system is to be used.

Our finding of positive returns on investment in the change treatment demonstrates the tremendous potential reputation systems offer even in environments in which identities can be changed. An interesting question for future research is whether introducing partner choice increases the efficiency of reputation systems in this setup. On the seller side competition is likely to create stronger incentives for trustworthy behavior. Competition on the buyer side could lead to a greater buyer willingness to take risks with sellers whose reputation is not entirely positive. This seems to hold in particular for new buyers (Jin and Kato 2006). Which of these effects is stronger is ultimately an empirical question. Our simple and parsimonious framework can successively be enriched to study partner choice and other measures such as verified IDs in future research.

[20]Friedman and Resnick (2001) draw the same conclusion from their model and provide a discussion of different measures to achieve this.

[21]This service is called ID Verify on eBay. Sellers who have their identity verified can display a little logo beside their user name.

Chapter 4

Behavior in Multi-Stage Elimination Tournaments

4.1 Introduction

Promotions play an important role for the provision of incentives in firms and other hierarchical organizations. According to Lazear and Gibbs (2008), they are probably the most important source of extrinsic motivation for middle managers in most firms. The prospect of being promoted to a better-paid job creates incentives to work hard, even if current income is not tied to performance. In virtually all firms those who get promoted compete again for subsequent promotions: in many companies, there are up to a dozen hierarchical levels between the CEO and entry-level management (Belzil and Bognanno 2008, Lazear and Gibbs 2008). The prevalence of internal labor markets is also underlined by the fact that hiring is concentrated on lower levels of the hierarchy and that positions on higher ranks are filled primarily through promotion (Baker et al. 1994, Bognanno 2001). In a seminal contribution, Rosen (1986) has modelled the competition for promotion in such hierarchies as a multi-stage elimination tournament where in each stage fewer agents are selected for the next step of the career ladder. Incentives generated in such tournaments depend on two important components of the organizational structure: the immediate wage increase for an agent who gets promoted and the option value of competing in further stages of the tournament, i.e., having the chance to earn even higher wages.

Although the importance of multi-stage elimination tournaments is undisputed, stringent empirical tests of their incentive effects are scarce. In this chapter, we provide a step towards closing this gap with the help of a controlled laboratory experiment. Our main questions are: do people take future stages of a tournament into account when deciding on current effort? Are multi-stage elimination setups behaviorally different from simple one-stage tournaments? How does the wage structure in multi-stage tournaments affect the provision of effort?

We study these questions by comparing three treatments. Our main treatment is a two-stage tournament (*TS*) in which four subjects compete for being promoted to the second stage. Promotion depends on subjects' output which is a function of costly effort and an individual noise term. The two subjects with the lowest output levels in the first stage are eliminated from further competition and receive a low wage. The two subjects with the highest output levels in the first stage are promoted, i.e., they are allowed to take part in the second stage where they compete against each other again. The subject with the highest second-stage output receives a high wage, whereas the other finalist is paid an intermediate wage. Parameters in this treatment are chosen to make the tournament incentive maintaining in the sense that equilibrium effort is identical in both stages (Rosen 1986).

We compare this treatment to a one-stage tournament (*OS*) in which four subjects compete once for two top positions. Wages for the promoted subjects are chosen such that the one-stage tournament is strategically equivalent to the first stage of *TS*. This means that promoted subjects in *OS* earn the sum of the intermediate wage in *TS* *and* the monetary equivalent of the second-stage option value in that tournament, implying that equilibrium effort is the same in both treatments. Comparing *OS* and *TS* thus allows testing whether strategic equivalence translates into behavioral equivalence. Our third treatment (*TSC*) is identical to the *TS* treatment with the exception that the wage structure is more convex, i.e., the intermediate wage is smaller in *TSC*. It is designed to study how subjects react to differences in the wage structure. In particular, we can test whether subjects—as predicted by theory—exert lower efforts in the first and higher efforts in the second stage under the more convex wage structure of *TSC* compared to *TS*.

Our findings can be summarized as follows: First, average behavior is remarkably close to the predictions of tournament theory in our one-stage treatment. This parallels findings of previous experiments on symmetric one-stage tournaments (e.g., Bull et al. 1987, Orrison et al. 2004). Second, behavior in the *TS* treatment indicates that subjects take the option value of future promotion possibilities into account when deciding on their work effort in multi-stage tournaments. Third, we also observe important departures from theoretical predictions in the *TS* treatment. Behavior in the first stage of *TS* differs strongly both from the one-stage treatment and from theoretical predictions. Subjects exert significantly higher efforts in the first stage of the two-stage tournament. Fourth, the results of the *TSC* treatment confirm the occurrence of excess effort provision in the first stage of the tournament. Subjects react only weakly to the change in the wage structure, implying that first-stage excess effort is even higher in *TSC*. Finally, the *TS* treatment is incentive maintaining in the sense that efforts in the second stage are not significantly different from first-stage efforts.

Most promotion tournaments that we observe in firms and other hierarchical organizations have multiple stages. Our results indicate that the mechanisms of incentive provision in multi-stage tournaments largely operate as suggested by theory. People do not only respond to differences in prizes, or wages, but are also motivated by the option value generated by future promotion possibilities. Our findings also provide insights with regard to the question whether one-stage tournaments are behaviorally equivalent to multi-stage designs. Adding one or more stages seems to make a fundamental difference, as people tend to exert excess effort in early stages of the tournament. This shows that one cannot necessarily draw inferences from simple one-stage setups to more complex tournaments.

The finding that people tend to exert excess effort in early stages of a multi-stage competition also has interesting organizational implications. If—as our findings suggest—multi-stage competitions provoke excess effort exertion by employees, this may offer a possible explanation why firms rely heavily on promotion based incentive schemes even if more direct means of performance assessment and compensation are available. Excess effort makes a promotion tournament "cheaper" for principals compared to other incentive schemes such as piece rates or bonus contracts. Moreover, excess effort influences the "optimal architecture" of promotion tournaments and other contests (see Moldovanu and Sela 2006).

The chapter complements previous studies which have used field data from executive compensation, sports, or agricultural production to evaluate predictions of tournament theory (e.g., Ehrenberg and Bognanno 1990, Knoeber and Thurman 1994, Bognanno 2001). It has been found that many implications of tournament theory are consistent with empirical observations: for instance, higher prizes tend to increase performance (Ehrenberg and Bognanno 1990), wage profiles in the field are convexly increasing (Eriksson 1999), and winner prizes increase with the number of competitors (Bognanno 2001). An advantage of testing tournament theory with laboratory experiments is the possibility to directly test *pure incentive effects* of tournaments. For instance, one can measure agents' behavioral reactions to changes in the wage structure without being concerned that the composition of the workforce might change due to the modified incentive scheme (Lazear 2000, Dohmen and Falk 2006). Experimental data also allows to rule out influences of unobservable variables that might affect actual promotion decisions in the field, like agents' soft skills or supervisor favoritism.

A number of previous papers have studied various aspects of tournaments in experiments, such as the effects of different prize spreads (Bull et al. 1987, Harbring and Irlenbusch 2003), sabotage activities (Falk, Fehr and Huffman 2008, Harbring and Irlenbusch 2008) or asymmetries in promotion chances (Schotter and Weigelt 1992). In contrast to our work, this literature concentrates on one-stage tournaments. In view of our findings it is not clear to what extent these previous findings translate to multi-stage setups. For instance, more competitive behavior (excess effort) in early stages might also lead to an increase in sabotage activities in multi-stage tournaments.

The remainder of the chapter is organized as follows: the next section presents a simple model of multi-stage elimination tournaments on which our experiment is based. Section 4.3 discusses our experimental design and derives hypotheses. Section 4.4 shows the main results and Section 4.5 concludes.

4.2 A Simple Model of Multi-Stage Elimination Tournaments

We consider a simple elimination tournament in which four identical agents compete for promotion.[1] The promotion decision depends on relative output produced by the agents. Competition consists of two stages: in the first stage, all four agents compete against each other. The two agents with the lowest output levels in that stage receive a wage w_{low} and are eliminated from further competition. The two agents with the highest output levels in the first stage are promoted, i.e., they are allowed to take part in the second stage (or "final") where they compete against each other again. The agent who produces more output in the second stage receives a wage w_{high}, whereas the other finalist gets an intermediate wage w_{med}. Note that the decision of who receives w_{high} or w_{med} does not depend on the first-stage output of the finalists.

This two-stage elimination tournament can be modeled as follows. In the first stage of the tournament four agents $i = 1, 2, 3, 4$ compete against each other. Agents who participate in stage $k \in \{1, 2\}$ individually produce output $y_{i,k}$ according to the production function

$$y_{i,k} = e_{i,k} + \epsilon_{i,k}$$

where $e_{i,k}$ denotes the effort level that agent i exerts in stage k. $\epsilon_{i,k}$ is a random shock faced by agent i in stage k. Shocks are assumed to be drawn independently for each agent in each stage. For simplicity, we assume that $\epsilon_{i,k}$ is uniformly distributed on the interval $[-q, q]$.[2] Agent i's output in stage k does not depend on previous effort or output and the production technology is identical for all agents i in all stages k. Agents bear the cost of effort exertion. We assume the cost function to be

[1]Most of the assumptions below follow the classic (one-stage) tournament model introduced by Lazear and Rosen (1981).

[2]Virtually all tournament experiments use the uniform distribution, primarily because its concept is easy to understand for experimental subjects. The predictions of the model, however, can be generalized to other distributions of shocks. Lazear and Rosen (1981) discuss which conditions have to be fulfilled for the existence of an equilibrium in pure strategies.

of the following form:

$$C(e_{i,1}, e_{i,2}) = \frac{e_{i,1}^2}{c} + \frac{e_{i,2}^2}{c}$$

Note that this specification implies separability of costs across stages, i.e., in line with Rosen (1986) there is also no carry-over of costs between stages. Furthermore, we assume that agents are identical and risk-neutral with utility functions which are additively separable in wages and effort costs

$$U_i(w, e_{i,1}, e_{i,2}) = w - C(e_{i,1}, e_{i,2})$$

For the derivation of equilibrium predictions we restrict our attention to the set of symmetric subgame perfect Nash equilibria. The two-stage tournament can be solved by backward induction. Because (i) the decision who wins the second stage solely depends on the output of the finalists in this stage, (ii) there is no cost carry-over between stages and (iii) the random terms are independently distributed both across stages and agents, the final of our two-stage tournament is equivalent to a simple one-stage tournament in which two participants compete for a promotion. Given that agent i has reached the second stage where two agents compete for one winner prize w_{high} and one loser prize w_{med}, he chooses stage-two effort $e_{i,2}$ in order to maximize an expected utility function of the following form:[3]

$$EU_i(w_{high}, w_{med}, e_{i,2}, e_{j,2}) = \pi(y_{i,2} > y_{j,2})w_{high} + [1 - \pi(y_{i,2} > y_{j,2})]w_{med} - C(e_{i,1}, e_{i,2})$$

$\pi(y_{i,2} > y_{j,2})$ denotes the probability that i's output in stage two is greater than the output of agent j. With our assumptions regarding the production function and random terms, this expression can be rewritten as follows:

$$EU_i(w_{high}, w_{med}, e_{i,2}, e_{j,2}) = F_{\epsilon_{j,2}-\epsilon_{i,2}}[e_{i,2} - e_{j,2}](w_{high} - w_{med}) + w_{med} - C(e_{i,1}, e_{i,2})$$

where $F_{\epsilon_{j,2}-\epsilon_{i,2}}[\cdot]$ denotes the cdf of the difference between random terms $\epsilon_{j,2}$, $\epsilon_{i,2}$. Maximizing $EU_i(\cdot)$ over $e_{i,2}$ yields the following first-order condition:

$$f_{\epsilon_{j,2}-\epsilon_{i,2}}(e_{i,2} - e_{j,2})(w_{high} - w_{med}) = \frac{\partial C(\cdot)}{\partial e_{i,2}}$$

[3]The parameters chosen in our experimental treatments (see below) ensure that all participation constraints are fulfilled. We therefore do not explicitly consider them here.

Assuming symmetry yields $f_{\epsilon_{j,k}-\epsilon_{i,k}}(0) = \frac{1}{2q}$ for $\epsilon_{i,k}, \epsilon_{j,k} \sim U[-q, q]$. The symmetric subgame perfect Nash equilibrium of the two-stage tournament thus entails the following second-stage effort level $e_{i,2}^{*,TS}$ for the two agents who participate in the final:

$$e_{i,2}^{*,TS} = \frac{(w_{high} - w_{med})c}{4q}$$

Given that both finalists play this equilibrium, the expected utility gain in the final, i.e., the continuation value for an agent in the first stage is given as follows

$$EV_{i,2} = w_{med} + \frac{1}{2}[w_{high} - w_{med}] - \frac{\left(e_{i,2}^{*,TS}\right)^2}{c}$$

An agent who reaches the final earns a wage of w_{med} for sure. By exerting stage-two effort $e_{i,2}^{*,TS}$, he has the chance to receive the higher wage w_{high} instead. In the symmetric equilibrium, this occurs with probability $\frac{1}{2}$. Moreover, he has to pay the cost of effort exertion in the second stage.

Turning to the analysis of the tournament's first stage, it is obvious that (in expected values) this stage can be modeled as a one-stage tournament between four agents with two winner prizes $EV_{i,2}$ and two loser prizes w_{low} (see Rosen 1986). The derivation of equilibrium effort for such a tournament follows the same steps as above. Alternatively we can apply a result from Orrison et al. (2004) who show that equilibria of fully symmetric one-stage tournaments are not affected by "organizational replication" for our specification of the production function, cost function, and random terms. This implies that an equilibrium in a tournament with two identical participants and one winner prize is also an equilibrium in a tournament with four identical participants and two winner prizes.[4]

As a shortcut we can therefore use the solution for $e_{i,2}^{*,TS}$ and simply replace w_{high} and w_{med} with $EV_{i,2}$ and w_{low} to obtain the equilibrium effort level for the first stage:

$$e_{i,1}^{*,TS} = \frac{(EV_{i,2} - w_{low})c}{4q} = \frac{\left(w_{med} - w_{low} + \frac{1}{2}\left[w_{high} - w_{med}\right] - \frac{\left(e_{i,2}^{*,TS}\right)^2}{c}\right)c}{4q}$$

[4]More generally Orrison et al. (2004) show that an equilibrium in a tournament with n participants and $\frac{1}{2}n$ winner prizes is still an equilibrium in a tournament with mn participants and $\frac{m}{2}n$ winner prizes.

This expression illustrates the two components of incentive provision in multi-stage tournaments. By winning the first stage and qualifying for the final, an agent receives an *immediate wage gain* $(w_{med} - w_{low})$, but additionally has the option to compete in the final and win the top prize w_{high}. The *value of this option* is

$$\frac{1}{2}[w_{high} - w_{med}] - \frac{\left(e_{i,2}^{*,TS}\right)^2}{c}.$$

Several aspects of the model deserve special emphasis. First, our design closely follows the original model of elimination tournaments by Rosen (1986) with one notable exception: instead of having two semifinals with two participants each, who compete for one slot in the final, we analyze a setup with four participants competing for two slots in the final. While both variants are theoretically equivalent for symmetric agents, we employ the latter because it allows us to design a one-stage tournament which is procedurally as close as possible to our main treatment.[5] Note also that we abstract from heterogeneous abilities of participants in order to keep the design as simple and parsimonious as possible. This allows us to focus on the incentive aspect of elimination tournaments by ruling out selection of more able individuals into higher positions of a hierarchy. Including the selection aspect of tournaments would be an interesting follow-up to our study.

4.3 Experimental Design

4.3.1 Treatments and Hypotheses

Our experiment comprises of three treatments that allow us to study behavior in multi-stage tournaments from different angles. For all treatments our benchmark is the prediction of the symmetric subgame perfect Nash equilibrium. The main treatment *TS* is a two-stage elimination tournament with four participants competing for w_{high}, w_{med}, and w_{low} as discussed in the previous section. We compare this treatment to a one-stage tournament (*OS*) in which four subjects compete for two

[5]In particular—as will become clear in the next section—both tournaments have the same number of participants and subjects compete for the same number of promotions.

top positions. The two subjects with the highest output levels receive a wage w'_{med} in *OS* while the two losers of the competition receive a wage w_{low}. The *OS* treatment fulfills several purposes. First, it serves as a validity check for our results given that a number of studies on one-stage tournaments already exists. In particular, our parametrization of this treatment is very close to a treatment from Orrison et al. (2004).

More importantly, however, the *OS* treatment allows us to investigate whether one-stage tournaments are behaviorally different from multi-stage ones. To investigate this question we design *OS* such that it is strategically equivalent to the first stage of the two-stage tournament *TS*. As discussed in the previous section the first stage of a two-stage tournament can be interpreted as a one-stage tournament in which agents compete for the expected value of participating in the second stage. Strategic equivalence between *OS* and the first stage of *TS* is thus achieved by keeping w_{low} constant and choosing:

$$w'_{med} = EV_{i,2} = w_{med} + \frac{1}{2}[w_{high} - w_{med}] - \frac{\left(e_{i,2}^{*,TS}\right)^2}{c}$$

In other words, the wage for the promoted agents in the one-stage tournament (w'_{med}) is equivalent to the wage w_{med} from *TS plus* the option value of participating in the final of *TS*. This choice implies that equilibrium effort levels in the *OS* treatment and in the first stage of the *TS* treatment are the same. We can therefore formulate the *behavioral equivalence hypothesis*:

$$e_{i,1}^{OS} \quad = \quad e_{i,1}^{TS}$$

In addition to comparing behavior in tournaments with different numbers of stages, we are interested in how a change in the wage structure influences behavior in multi-stage tournaments. This comparative statics exercise is of practical interest. Wages are one of the variables which (within certain bounds) are most amenable to manipulation in organizational design. We therefore compare our main treatment to a second two-stage tournament with a more convex wage profile (*TSC*). It is identical to the *TS* treatment with the exception that the intermediate wage w''_{med} in this treatment is smaller than in *TS*. This implies weaker incentives (lower equilibrium effort) in the first stage and a higher wage spread and higher equilibrium effort in the second stage (*wage structure hypothesis*):

$$e_{i,1}^{TSC} < e_{i,1}^{TS}$$
$$e_{i,2}^{TSC} > e_{i,2}^{TS}$$

Finally, we analyze behavior in our main treatment across stages. In the *TS* treatment wages are chosen such that equilibrium efforts in the first and second stage are equal. Remember that the two elements of incentives in multi-stage tournaments are the wage spread and the option value of competing for further promotions. In the final stage, the option value is zero because there are no further promotions beyond that stage. To make the tournament in the *TS* treatment incentive maintaining in the sense of Rosen (1986), this decrease in the option value in the second stage is offset by an appropriate increase in the wage spread ($w_{high} - w_{med} > w_{med} - w_{low}$). Comparing behavior across stages thus allows to test the *incentive maintenance hypothesis*:

$$e_{i,1}^{TS} = e_{i,2}^{TS}$$

Experimental parameters and the resulting equilibrium efforts for all treatments are shown in Table 4.1. When deciding on their efforts, subjects could choose any integer $e_{i,k} \in \{0, 1, ..., 125\}$. The effort costs in each stage are given by

$$C(e_{i,k}) = \frac{e_{i,k}^2}{2250}$$

and $\epsilon_{i,k} \sim U[-60, 60]$.[6] The parameters chosen imply equilibrium efforts of 74 in both stages of *TS* and in *OS*, the lower intermediate wage in *TSC* changes equilibrium efforts to 42 in the first stage and 100 in the second stage of this treatment.

4.3.2 Experimental Procedures

The experiment was conducted at the BonnEconLab of the University of Bonn. A total of 96 subjects in six sessions were divided into groups of four. We employed

[6]Note that, while in equilibrium all players make positive profits, the range of feasible efforts and the specification of the cost function imply that in principle subjects could make losses. In the few cases where this occurred, losses were deducted from the showup-fee.

Treatment	*OS*	*TS*	*TSC*
c	2250	2250	2250
q	60	60	60
w_{high}	–	20	20
w'_{med} / w_{med} / w''_{med}	13.62	12.11	9.33
w_{low}	5.73	5.73	5.73
$e^*_{i,1}$	74	74	42
$e^*_{i,2}$	–	74	100

Table 4.1: *Experimental parameters and resulting equilibrium predictions.*

a one-shot between-subjects design, i.e., subjects participated either in the *OS*, the *TS* or the *TSC* treatment. The experiment was programmed and conducted with the software z-Tree (Fischbacher 2007).

Before the tournament started, subjects received detailed written instructions on the respective treatment they took part in (*OS*, *TS* or *TSC*). These were neutrally framed and did not contain potentially value-laden terms like "tournament", "final", "winner", etc.[7] After reading the instructions subjects completed several control questions. The experiment started only after all participants had answered all control questions correctly. During the tournament, subjects simultaneously entered their effort decision and were then asked to state their expectations about the other participants' efforts on the next screen. This question was not announced beforehand. After the first stage participants in *TS* and *TSC* were only informed about the realization of their own random draw and about whether they had been promoted to the second stage. The finalists then again made an effort choice and entered their expectation about their opponent's effort. At the end of the tournament subjects in all treatments were informed about their earnings and asked to fill in a questionnaire. The structure of the experimental session ensured that subjects' decisions in the tournament can be treated as independent observations. The whole experimental session lasted on average 100 minutes and subjects earned an average of 18.25 Euro (1 Euro = 1.26 USD at the time of experiment), including a showup-fee

[7]Instructions for the *TS* treatment and the schedule of effort costs handed out to participants can be found in Appendix D.1 and Appendix D.2, respectively.

of 4 Euro and a fixed payment of 3 Euro for completing the questionnaire.

Note that our experimental procedures differ from previous tournament experiments in that we implement a one-shot interaction structure. Previous experiments have typically used repeated interactions. The advantage of the latter is that it allows for learning, which is potentially important given the non-trivial decision environment in tournaments. A potential downside, however, is that repeated game structures question the validity of static equilibrium predictions. Since we are explicitly interested in testing theoretical predictions, we decided to use a one-shot design. This has the additional advantage that stakes in the one-shot interaction are relatively high. We check the regularity of our results by comparing the outcome of our *OS* treatment with those of similar repeated tournaments. Finding similar results would make us confident that our main findings are not driven by the one-shot character of our set-up, but instead by treatment differences.

4.4 Results

In this section we first test whether the results for the one-stage tournament replicate earlier findings from similar tournaments. We then study the dynamic aspect of multi-stage tournaments by comparing the one-stage tournament *OS* to the strategically equivalent first stage of the two-stage tournament *TS*. In a third step, we investigate the effects of differences in the wage structure on effort provision by comparing the two-stage tournaments *TS* and *TSC*. Finally, we address the question whether the *TS* treatment is incentive maintaining by analyzing behavior in the first and second stage of the tournament.

4.4.1 Behavior in the One-Stage Tournament

Table 4.2 reports effort decisions in the *OS* treatment (Column 1). Two points are worth noting. First, efforts are on average very close to the theoretical predictions. While the average effort of 69.9 is slightly below the Nash prediction of 74, median effort coincides exactly with the predicted effort level. Second, there is substantial heterogeneity in subjects' behavior (see also Figure 4.1).

Both observations are in line with previous findings from symmetric one-stage designs (e.g., Bull et al. 1987, Eriksson et al. 2006). In particular, our results replicate those found by Orrison et al. (2004) who observe an average effort of 73.3 for an almost identical tournament which was repeated 20 times using lower stakes. The similarity of our results to those of Orrison et al. (2004) show that one of the most important findings in the experimental literature on symmetric promotion tournaments—average effort being close to Nash predictions—is quite robust with respect to using one-shot vs. repeated interactions and with respect to increased stake sizes.

> **Result 1:** *Average behavior in the one-stage tournament is close to the predictions of the symmetric Nash equilibrium.*

4.4.2 Testing Behavioral Equivalence

Our one-stage tournament and the first stage of the two-stage tournament *TS* are strategically equivalent in the sense that the wage w'_{med} in the one-stage tournament includes the equilibrium option value of participating in the second stage of the two-stage tournament. A comparison of $e_{i,1}^{OS}$ and $e_{i,1}^{TS}$ therefore serves as a test of how subjects in the two-stage tournament perceive this option value. If subjects, for instance, do not take the option value into account when deciding on their first stage effort in *TS*, efforts in this stage should be lower compared to the *OS* treatment. If subjects evaluate the option value correctly, efforts in both treatments should be identical. Columns 1 and 2 of Table 4.2 show that behavior differs strongly between the two treatments. Average effort in the first stage of the *TS* treatment is 89.2, while median effort is 91. Thus, subjects behave much more competitively in the multi-stage tournament, exerting efforts which are more than 20% higher than those of their counterparts in the *OS* treatment. A Mann-Whitney U-test confirms that this treatment difference is highly significant ($p = 0.005$, two-sided).

Comparing effort levels in the two treatments to the theoretical predictions derived in Section 4.3 indicates that it is excess effort in *TS* rather than "too low"

Treatment	*OS*	*TS*	*TSC*
Average Effort	69.9	89.2	82.4
Median Effort	74	91	83
min(Effort)	1	40	40
max(Effort)	125	125	125
Variance	913.3	359.7	605.4
e^*	74	74	42

Table 4.2: *First stage behavior in the* OS, TS, *and* TSC *treatment.*

effort in *OS* that drives the treatment difference. A t-test with the null hypothesis that efforts are equal to Nash predictions confirms this: the null hypothesis is rejected in the *TS* treatment ($p < 0.001$), but cannot be rejected in *OS* ($p = 0.453$).[8] Subjects' effort choices suggest that they are not naive in the sense that they ignore the second stage. Quite to the contrary, the two-stage elimination tournament seems to trigger especially competitive behavior in the first stage.

The treatment difference is not just driven by some subjects choosing extreme effort levels in the *TS* treatment. A closer look at the distributions of first stage efforts in Figure 4.1 reveals instead that the whole effort distribution is shifted to the right in the *TS* treatment. As a consequence, efforts are less dispersed (Levene's test for equality of variances, $p < 0.05$, two-sided). The effort distribution illustrates that exerting excess effort is quite widespread in *TS*: 84% of subjects choose efforts higher than the equilibrium effort level of 74. This compares to only 47% in the *OS* treatment. The strong difference between treatments is also reflected at the lower tail of the distribution. While the lowest effort in the *OS* treatment is 1, no subject

[8]Note that the derivation of equilibrium predictions is based on the assumption of risk neutrality. We check the validity of this assumption by eliciting subjects' risk preferences with an incentive compatible lottery procedure (a detailed description of the procedure can be found in Appendix D.3). It turns out that our experimental subject pool is close to risk neutrality: the median subject in all three treatments is risk neutral, and the certainty equivalent of more than 50% of subjects lies in a range of $+/-0.25$ Euro around the risk neutral certainty equivalent (equal to 2 Euro for the lottery that was used). In addition, effort levels in *TS* are above theoretical predictions irrespective of the subjects' degree of risk aversion. Thus excess effort is not driven by subjects' risk preferences.

exerts effort below 40 in the *TS* treatment.

Result 2: *Efforts in the first round of* TS *are significantly higher than in the* OS *treatment. This difference is driven by excess effort in* TS.

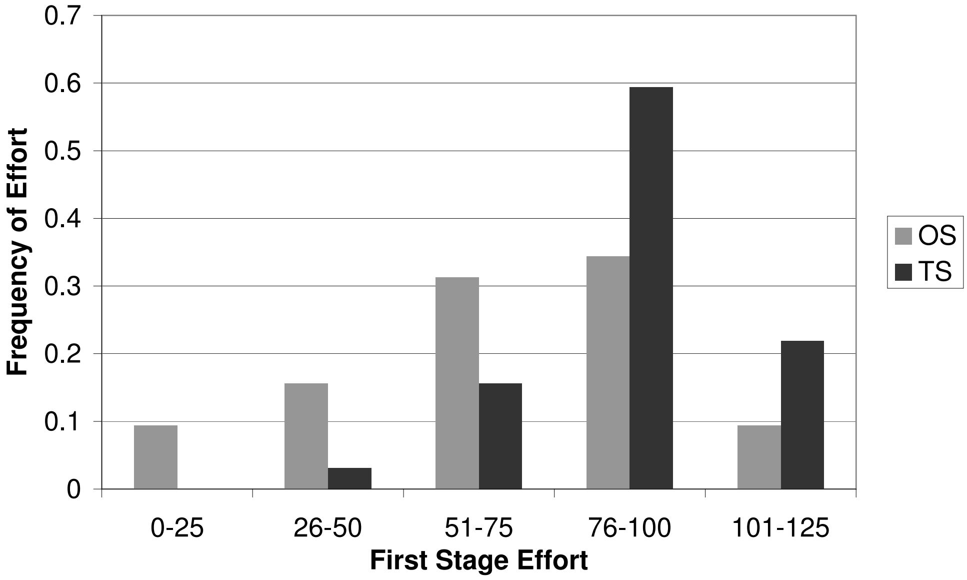

Figure 4.1: *Frequency of effort choices in the* OS *treatment and the first stage of the* TS *treatment.*

4.4.3 Wage Structures in Two-Stage Tournaments

In our next treatment comparison we investigate the behavioral effects of different wage structures in two-stage tournaments. Our main interest concerns the question whether first stage excess effort is a robust phenomenon. Is excess effort in the first stage an artefact of the specific wage structure we used in the *TS* treatment, e.g., is it specific to the incentive maintaining wage spreads? Or do we observe excess effort even if we reduce the incentive to provide effort in the first stage? To analyze these questions we implemented the *TSC* treatment, which is identical to the *TS* treatment except that the subject who produces less output in the second stage receives only $w''_{med} = 9.33$ Euro instead of $w_{med} = 12.11$ Euro. This more convex wage profile has the following theoretical implications. Incentives to provide effort in the first stage are weakened—the equilibrium effort level in the first stage of *TSC* is only 42 instead of 74 in the *TS* treatment. In the second stage, equilibrium effort increases from 74 to 100 (see Table 4.1).

Column 3 of Table 4.2 summarizes behavior in the first stage of the *TSC* treatment. Efforts in the first stage are much higher than theoretically predicted. The average effort level is 82.4 points, about 40 points higher than the equilibrium effort level of 42. A t-test with the null hypothesis that efforts are equal to the Nash prediction confirms that effort choices are significantly above the equilibrium prediction ($p < 0.001$). Indeed, about 88% of subjects choose efforts higher than 42. This suggests that excess effort in the first stage of our two-stage tournaments is robust to differences in the wage structure. Comparing effort choices in the *TSC* treatment to those observed in *TS* reveals that subjects react only weakly to the change in the wage structure between the two treatments. Average effort in *TSC* is only 7 points lower than in *TS*. This finding is striking given that there is a 32-point difference in equilibrium efforts. Despite the fact that first stage incentives are much weaker than in *TS*, effort choices do not differ significantly between the two treatments (Mann-Whitney U-test, $p = 0.245$, two-sided).

> **Result 3:** *The more convex wage structure in* TSC *induces even higher excess effort in the first stage. First stage efforts in* TS *and* TSC *are not significantly different, despite much weaker incentives in the* TSC *treatment.*

4.4.4 Testing Incentive Maintenance

Finally, we turn to behavior in the second stage of the *TS* treatment. Remember that parameters were chosen such that the tournament in *TS* is incentive maintaining, i.e., equilibrium effort levels are the same (equal to 74) in both stages. We know already that efforts are above the equilibrium prediction in the first stage of *TS*. In this sense, we can reject the hypothesis of *equilibrium* effort choices in both stage of *TS*. It remains to show, however, whether effort levels are the same in both stages of the tournament, or whether they differ. Does the two-stage character of *TS* induce above equilibrium effort also in the second stage or do players reduce efforts relative to their first stage behavior?

In turns out that the the evidence is mixed. On average, effort decreases when comparing behavior across stages in the *TS* treatment. Average effort in the first stage is 89, in the final it goes down to 82. The median effort choice decreases from 91 to 87.5. Note that average effort in the first stage includes the efforts of those who did not make it to the second stage. Since—by design of the promotion tournament—the latter usually exerted lower effort, the decrease from stage one to stage two is larger if we consider only finalists' behavior. Their average effort in the first stage is 96, implying that on average finalists decrease their effort by 14 points. However, these numbers hide considerable heterogeneity on an individual level: the fraction of finalists who decrease their effort in the second stage is only slightly higher than the fraction of subjects who increase it (50% and 44%, respectively). It is therefore not surprising that the overall decrease in finalists' efforts is insignificant (Wilcoxon signed rank test, p-value = 0.289). However, those who adjust their effort downwards on average do so much stronger than those who raise their effort.

Does this mean that excess effort in the first stage just mirrors the lower efforts in the second stage? That is, do subjects *expect* low effort levels and thus low effort costs in the second stage and therefore increase their first stage effort due to a higher (perceived) option value? While this might be the case for some finalists, we can rule out that it accounts for first stage behavior on a more general level as, on average, effort choices in the second stage are still above the equilibrium prediction. We can also use the observed first stage effort choice of a subject to calculate the option value implicitly underlying her decision. Using this option value, we can then construct the (hypothetical) second stage effort level which would rationalize the first stage effort choice of the subject at hand. For instance, if a subject beliefs that all players will exert zero effort in the second stage, her subjective option value increases to $\frac{1}{2}[w_{high} - w_{med}]$ since $\frac{\left(e_{i,2}^{TS}\right)^2}{c} = 0$ (see Section 4.2). Assuming this option value instead of the equilibrium option value rationalizes a first stage effort of 97 (using the wages and cost parameters of *TS*). Conducting the calculation for the first stage effort choice of the median subject in *TS* (equal to 91) yields an implied second stage effort of 37. I.e., if the median subject had expected a second stage effort level of 37, the perceived option value would rationalize her observed first stage effort choice. This value is, however, far below the actual effort choices in the second stage (cp. Table 4.2).[9] The expectation of low second stage effort levels can thus not account for the observed excess effort in the first stage of *TS*.

> **Result 4:** *Efforts in the second stage of* TS *are lower, but not significantly different from efforts in the first stage. In this sense, the* TS *treatment is incentive maintaining.*

4.5 Concluding Remarks

Promotions in most hierarchical organizations take the form of multi-stage elimination tournaments. In this chapter we have studied behavior in such tournaments with simple laboratory experiments. Our results demonstrate the importance of carefully analyzing the incentive effects of promotions in multi-level hierarchies. They show that the basic logic of incentive provision in multi-stage elimination tournaments works in the sense that people take future promotion possibilities into account when deciding on current work effort. However, we also observe important departures from theoretical predictions. Subjects tend to exert excess effort in the first stage of our two-stage elimination tournament. By contrast, we do not observe this phenomenon in a strategically equivalent one-stage tournament. Under a more convex wage structure, the overprovision of effort is even more pronounced.

Our experiments suggest that behavior in multi-stage tournaments deviates from behavior in one-stage tournaments in a systematic way. Our data do, however, not allow us to give a definite answer on the precise mechanism that causes this change in behavior. Several factors may act in concert: it could be that subjects experience additional non-monetary "joy of winning" when being promoted (Parco et al. 2005, Kräkel 2008), which might be more pronounced when the hierarchy has more layers. An additional potential rationale for subjects' behavior are preferences for status (Moldovanu et al. 2007). Multi-stage tournaments with their more precise

[9] The same exercise for the *TSC* treatment yields a value of 14 while the average effort level actually observed in the second stage of this treatment is 82.9. Thus, although subjects on average choose efforts below the equilibrium value of 100 in the second stage of *TSC*, this cannot explain the excess efforts exerted in the first stage.

definition of hierarchical level (and status) might trigger especially competitive behavior of status concerned agents. Our data on subjects' expectations are consistent with these interpretations: about two thirds of subjects choose an effort above the second-highest effort level that they expect from their competitors, irrespective of the absolute level of the effort expectation.

The observed behavior could also help to rationalize why firms rely so strongly on promotions as incentive device, even in work environments where more direct performance pay is feasible. Excess effort in early stages of multi-stage tournaments makes this form of incentive provision comparatively "cheap" for the principal as it decreases the wage cost per unit of effort. For instance, in our *TS* tournament this cost is 9.81 Cent in theory, but only 8.36 Cent in practice. This implies that a principal who implemented the *TS* wage scheme had to pay less for every unit of effort (and production) than theoretically predicted. Which wage profile a principal or tournament designer actually prefers depends on his objectives. In some situations it might be sufficient to concentrate on the cost per effort. In other situations a tournament designer may, for example, put special emphasis on the performance of agents in higher stages of the tournament. The *TSC* treatment suggests that excess effort in early stages eventually might come at the cost of reduced performance in later stages if the wage structure becomes too convex. Independent of the specific objective function, a tournament designer should take into account that agents' behavior can ultimately change the optimal architecture of promotion tournaments in terms of wage profiles, promotion rates, etc.

Bibliography

Adams, J. S. (1963), 'Wage inequities, productivity and work quality', *Industrial Relations* **3**, 9–16.

Adams, J. S. (1965), Inequity in social exchange, *in* L. Berkowitz, ed., 'Advances in Experimental Social Psychology', Vol. 2, Academic Press, New York, pp. 267–299.

Akerlof, G. A. (1982), 'Labor contracts as partial gift exchange', *The Quarterly Journal of Economics* **97**, 543–569.

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

Altmann, S., Dohmen, T. and Wibral, M. (2008), 'Do the reciprocal trust less?', *Economics Letters* **99**(3), 454–457.

Andreoni, J. and Miller, J. (2002), 'Giving According to GARP: An Experimental Test of the Consistency of Preferences for Altruism', *Econometrica* **70**(2), 737–753.

Andrews, I. R. (1967), 'Wage inequity and job perfomance: An experimental study', *Journal of Applied Psychology* **51**, 39–45.

Ashraf, N., Bohnet, I. and Piankov, N. (2006), 'Decomposing trust and trustworthiness', *Experimental Economics* **9**, 193–208.

Bajari, P. and Hortaçsu, A. (2004), 'Economic Insights from Internet Auctions', *Journal of Economic Literature* **42**(2), 457–486.

Baker, G., Gibbs, M. and Holmström, B. (1994), 'The internal economics of the firm: Evidence from personnel data', *The Quarterly Journal of Economics* **109**(4), 881–919.

Baker, G. P., Jensen, M. C. and Murphy, K. J. (1988), 'Compensation and incentives: Practice vs. theory', *Journal of Finance* **43**(3), 593–616.

Bandiera, O., Barankay, I. and Rasul, I. (2005), 'Social preferences and the response to incentives: Evidence from personnel data', *The Quarterly Journal of Economics* **120**(3), 917–962.

Bellemare, C. and Kröger, S. (2007), 'On representative social capital', *European Economic Review* **51**, 183–202.

Belzil, C. and Bognanno, M. (2008), 'Promotions, demotions, halo effects, and the earnings dynamics of american executives', *Journal of Labor Economics* **26**(2), 287–310.

Berg, J. E., Dickhaut, J. and McCabe, K. (1995), 'Trust, reciprocity, and social history', *Games and Economic Behavior* **10**, 122–142.

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

Blount, S. (1995), 'When social outcomes aren't fair: The effect of causal attribution on preferences', *Organizational Behavior and Human Decision Processes* **63**, 131–144.

Bognanno, M. L. (2001), 'Corporate tournaments', *Journal of Labor Economics* **19**(2), 290–315.

Bohnet, I. and Zeckhauser, R. (2004), 'Trust, risk and betrayal', *Journal of Economic Behavior and Organization* **55**(4), 467–484.

Bolton, G. E., Katok, E. and Ockenfels, A. (2004), 'How effective are online reputation mechanisms? an experimental investigation', *Management Science* **50**(11), 1587–1602.

Bolton, G. E. and Ockenfels, A. (2000), 'Erc: A theory of equity, reciprocity, and competition', *American Economic Review* **90**(1), 166–193.

Bolton, G., Greiner, B. and Ockenfels, A. (2008), 'Engineering trust - reciprocity in the production of reputation information', *Mimeo* .

Bowles, S. and Park, Y. (2005), 'Emulation, inequality, and work hours: Was thorsten veblen right?', *The Economic Journal* **115**, F397–F412.

Brown, M., Falk, A. and Fehr, E. (2004), 'Relational contracts and the nature of market interactions', *Econometrica* **72**(3), 747–780.

Brown, M. and Zehnder, C. (2008), 'Credit reporting, relationship banking, and loan repayment', *Journal of Money, Credit, and Banking* **39**, 1883–1918.

Buchan, N. R., Croson, R. T. and Solnick, S. (2008), 'Trust and gender: An examination of behavior and beliefs in the investment game', *Journal of Economic Behavior and Organization* **68**, 466–476.

Bull, C., Schotter, A. and Weigelt, K. (1987), 'Tournaments and piece rates: An experimental study', *Journal of Political Economy* **95**(1), 1–33.

Burks, S. V., Carpenter, J. P. and Verhoogen, E. (2003), 'Playing both roles in the trust game', *Journal of Economic Behavior and Organization* **51**, 195–216.

Camerer, C. F. (2003), *Behavioral Game Theory*, Princeton University Press, Princeton, New Yersey.

Charness, G. (2004), 'Attribution and reciprocity in an experimental labor market', *Journal of Labor Economics* **22**(3), 665–688.

Charness, G. and Kuhn, P. (2007), 'Does pay inequality affect worker effort? experimental evidence', *Journal of Labor Economics* **25**, 693–723.

Clark, A. E. and Oswald, A. J. (1996), 'Satisfaction and comparison income', *Journal of Public Economics* **61**(3), 359–381.

Clark, A., Masclet, D. and Villeval, M.-C. (2006), Effort and comparison income: Survey and experimental evidence. GATE Working Paper 06-01.

Cox, J. C. (2004), 'How to identify trust and reciprocity', *Games and Economic Behavior* **46**(2), 260–281.

Dellarocas, C. and Wood, C. A. (2003), 'The Digitization of Word of Mouth: Promise and Challenges of Online Feedback Mechanisms', *Management Science* **49**(10), 1407.

Dellarocas, C. and Wood, C. A. (2008), 'The Sound of Silence in Online Feedback: Estimating Trading Risks in the Presence of Reporting Bias', *Management Science* **54**(3), 460.

Dohmen, T. and Falk, A. (2006), 'Performance pay and multi-dimensional sorting: Productivity, preferences and gender', *IZA Discussion Paper* **2001**.

Dohmen, T., Falk, A., Huffman, D. and Sunde, U. (2006), 'The intergenerational transmission of risk and trust attitudes', *IZA Discussion Paper* **2380**.

Dufwenberg, M. and Gneezy, U. (2000), 'Measuring beliefs in an experimental lost wallet game', *Games and Economic Behavior* **30**, 163–182.

Dufwenberg, M. and Kirchsteiger, G. (2004), 'A theory of sequential reciprocity', *Games and Economic Behavior* **47**, 268–298.

Easterlin, R. A. (2001), 'Income and happiness: Towards a unified theory', *The Economic Journal* **111**, 465–484.

Ehrenberg, R. G. and Bognanno, M. L. (1990), 'Do tournaments have incentive effects?', *Journal of Political Economy* **98**(6), 1307–1324.

Erev, I., Bornstein, G. and Galili, R. (1993), 'Constructive intergroup competition as a solution to the free rider problem: A field experiment', *Journal of Experimental Social Psychology* **29**, 463–478.

Eriksson, T. (1999), 'Executive compensation and tournament theory: Empirical tests on danish data', *Journal of Labor Economics* **17**(2), 262–280.

Eriksson, T., Teyssier, S. and Villeval, M.-C. (2006), 'Self-selection and the efficiency of tournaments', *IZA Discussion Paper* **1983**.

Falk, A., Fehr, E. and Fischbacher, U. (2003), 'On the nature of fair behavior', *Economic Inquiry* **41**(1), 20–26.

Falk, A., Fehr, E. and Fischbacher, U. (2005), 'Driving forces behind informal sanctions', *Econometrica* **73**(6), 2017–2030.

Falk, A., Fehr, E. and Huffman, D. (2008), 'The power and limits of tournament incentives', *Working Paper, University of Bonn* .

Falk, A. and Fischbacher, U. (2006), 'A theory of reciprocity', *Games and Economic Behavior* **54**(2), 293–315.

Falk, A., Huffman, D. and MacLeod, W. B. (2008), 'Institutions and contract enforcement', *IZA Discussion Paper* **3435**.

Falk, A. and Zehnder, C. (2007), 'Discrimination and in-group favoritism in a city-wide trust experiment', *IZA Discussion Paper* **2765**.

Fehr, E. and Gächter, S. (2000), 'Cooperation and punishment in public goods experiments', *American Economic Review* **90**, 980–994.

Fehr, E., Gächter, S. and Kirchsteiger, G. (1997), 'Reciprocity as a contract enforcement device: Experimental evidence', *Econometrica* **65**, 833–860.

Fehr, E., Kirchsteiger, G. and Riedl, A. (1993), 'Does fairness prevent market clearing? an experimental investigation', *Quarterly Journal of Economics* **108**, 437–460.

Fehr, E., Klein, A. and Schmidt, K. M. (2007), 'Fairness and contract design', *Econometrica* **75**(1), 121–154.

Fehr, E. and Schmidt, K. M. (1999), 'A theory of fairness, competition and cooperation', *Quarterly Journal of Economics* **114**, 817–868.

Fehr, E. and Schmidt, K. M. (2006), The economics of fairness, reciprocity and altruism – experimental evidence and new theories, *in* S.-C. Kolm and J. M. Ythier, eds, 'Handbook of the Economics of Giving, Altruism and Reciprocity', Elsevier.

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

Fließbach, K., Weber, B., Trautner, P., Dohmen, T., Sunde, U., Elger, C. and Falk, A. (2007), 'Social comparison affects reward-related brain activity in the human ventral striatum', *Science* **318**, 1305–1308.

Friedman, E. J. and Resnick, P. (2001), 'The social cost of cheap pseudonyms', *Journal of Economics & Management Strategy* **10**(2), 173–199.

Greif, A. (1989), 'Reputation and coalitions in medieval trade: Evidence on the maghribi traders', *Journal of Economic History* **49**(4), 857–882.

Harbring, C. and Irlenbusch, B. (2003), 'An experimental study on tournament design', *Labour Economics* **10**, 443–464.

Harbring, C. and Irlenbusch, B. (2008), 'How many winners are good to have? on tournaments with sabotage', *Journal of Economic Behavior and Organization* **65**, 682–702.

Holmström, B. (1982), 'Moral hazard in teams', *Bell Journal Of Economics* **13**(2), 324–340.

Homans, G. C. (1961), *Social Behavior: Its Elementary Forms*, Routledge & Kegan Paul, London.

Jin, G. Z. and Kato, A. (2006), 'Price, quality, and reputation: evidence from an online field experiment', *The RAND Journal of Economics* **37**(4), 983–1005.

Kandori, M. (1992), 'Social Norms and Community Enforcement', *Review of Economic Studies* **59**(1), 63–80.

Kelley, H. H. and Stahelski, A. J. (1970), 'Social interaction basis of cooperators' and competitors' beliefs about others', *Journal of Personality and Social Psychology* **16**, 66–91.

Keser, C. (2002), 'Trust and Reputation Building in E-Commerce', *CIRANO Working Paper 2002s-75* .

Knack, S. and Keefer, P. (1997), 'Does social capital have an economic payoff? a cross-country investigation', *Quarterly Journal of Economics* **112**(4), 1251–1288.

Knez, M. and Simester, D. (2001), 'Firm-wide incentives and mutual monitoring at continental airlines', *Journal of Labor Economics* **19**(4), 743–772.

Knoeber, C. R. and Thurman, W. N. (1994), 'Testing the theory of tournaments: An empirical analysis of broiler production', *Journal of Labor Economics* **12**(2), 155–179.

Konow, J. (2003), 'Which is the fairest one of all? a positive analysis of justice theories', *Journal of Economic Literature* **41**, 1188–1239.

Kreps, D., Milgrom, P., Roberts, J. and Wilson, R. (1982), 'Rational cooperation in the finitely repeated prisoners' dilemma', *Journal of Economic Theory* **27**(2), 245–52.

Kräkel, M. (2008), 'Emotions in tournaments', *Journal of Economic Behavior and Organization* **67**, 204–214.

Lawler, E. E. (1967), 'Effects of hourly overpayment on productivity and work quality', *Journal of Personality and Social Psychology* **10**, 306–313.

Layard, R. (2005), *Happiness: Lessons from a New Science*, 1 edn, Penguin, London, UK.

Lazear, E. P. (1989), 'Pay equality and industrial politics', *Journal of Political Economy* **97**(3), 561–80.

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

Lazear, E. P. and Gibbs, M. (2008), *Personnel Economics in Practice*, 2 edn, John Wiley & Sons.

Lazear, E. P. and Rosen, S. (1981), 'Rank-order tournaments as optimum labor contracts', *Journal of Political Economy* **89**(5), 841–864.

Loewenstein, G., Thompson, L. and Bazerman, M. (1989), 'Social utility and decision making in interpersonal contexts', *Journal of Personality and Social Psychology* **57**, 426–441.

Lucking-Reiley, D., Bryan, D., Prasad, N. and Reeves, D. (2007), 'Pennies from eBay: the Determinants of Price in Online Auctions', *Journal of Industrial Economics* **55**(2), 223–233.

Maximiano, S., Sloof, R. and Sonnemans, J. (2007), 'Gift exchange in a multi-worker firm', *The Economic Journal* **117**, 1025–1050.

McCabe, K. A., Rigdon, M. L. and Smith, V. L. (2003), 'Positive reciprocity and intentions in trust games', *Journal of Economic Behavior and Organization* **52**, 267–275.

Medoff, J. L. and Abraham, K. G. (1980), 'Experience, performance, and earnings', *The Quarterly Journal of Economics* **95**(4), 703–36.

Milgrom, P. R., North, D. C. and Weingast, B. R. (1990), 'The role of institutions in the revival of trade: The law merchant, private judges, and the champagne fairs', *Economics & Politics* **2**(1), 1–23.

Milgrom, P. and Roberts, J. (1992), *Economics, Organization and Management*, Prentice Hall, Upper Saddle River, New Jersey.

Moldovanu, B. and Sela, A. (2006), 'Contest architecture', *Journal of Economic Theory* **126**(1), 70–96.

Moldovanu, B., Sela, A. and Shi, X. (2007), 'Contests for status', *Journal of Political Economy* **115**(2), 338–363.

Mowday, R. T. (1991), Equity theory predictions of behavior in organizations, *in* R. M. Steers and L. W. Porter, eds, 'Motivation and Work Behavior', McGraw-Hill, New York, pp. 111–131.

Neumark, D. and Postlewaite, A. (1998), 'Relative income concerns and the rise in married women's employment', *Journal of Public Economics* **70**, 157–183.

Ockenfels, A. (2003), 'Reputationsmechanismen auf internet-marktplattformen: Theorie und empirie', *Zeitschrift für Betriebswirtschaft* **73**(3), 295–315.

Okuno-Fujiwara, M. and Postlewaite, A. (1995), 'Social norms and random matching games', *Games and Economic Behavior* **9**(1), 79–109.

Orrison, A., Schotter, A. and Weigelt, K. (2004), 'Multiperson tournaments: An experimental investigation', *Management Science* **50**(2), 268–279.

Parco, J. E., Rapoport, A. and Amaldoss, W. (2005), 'Two-stage contests with budget constraints: an experimental study', *Journal of Mathematical Psychology* **49**, 320–338.

Pfeffer, J. and Langton, N. (1993), 'The effect of wage dispersion on satisfaction, productivity, and working collaboratively: Evidence from college and university faculty', *Administrative Science Quarterly* **38**(3), 382–407.

Prendergast, C. (1999), 'The provision of incentives in firms', *Journal of Economic Perspectives* **37**, 7–63.

Rabin, M. (1993), 'Incorporating fairness into game theory and economics', *American Economic Review* **83**(5), 1281–1302.

Rosen, S. (1986), 'Prizes and incentives in elimination tournaments', *American Economic Review* **76**(4), 701–715.

Roth, A. E. (2002), 'The economist as engineer: Game theory, experimentation, and computation as tools for design economics', *Econometrica* **70**(4), 1341–1378.

Schotter, A. and Weigelt, K. (1992), 'Asymmetric tournaments, equal opportunity laws, and affirmative action: Some experimental results', *The Quarterly Journal of Economics* **107**(2), 511–539.

Thöni, C. and Gächter, S. (2008), ''Kinked conformism' in voluntary cooperation', *mimeo* .

Torgler, B., Schmidt, S. L. and Frey, B. S. (2006), 'Relative income position and performance: an empirical panel analysis', *IEW Working Paper* **268**.

Appendices

A.1 Instructions for Chapter 1

In this part of the experiment, your endowment, contributions, and earnings are calculated in points. At the end of the experiment the points you have earned will be converted at an exchange rate of **50 points = 1 Euro** and paid out to you in cash.

In this part of the experiment you are in a group together with one of the other 15 participants. This participant will be **assigned to you randomly by the computer**. In each group of two there are two roles (a sender and a receiver). One member of the group will be the sender, and one member of the group will be the receiver. Each participant receives an endowment of 120 points.

The experiment consists of two stages:

In the first stage the sender can make a transfer to the receiver. The transfer is a number between 0 and 120 in steps of 20 points, i.e., either 0, 20, 40, 60, 80, 100, or 120 points. The transfer is tripled by the experimenter.

Examples: In case the sender transfers 60 points, the receiver gets 180 points. In case the sender transfers 10 points, the receiver gets 30 points. In case nothing is transferred, the receiver gets 0 points etc.

This means that, at the end of the first stage, the amount of points available to the receiver is the sum of the endowment plus three times the transfer.

In the second stage the receiver can transfer any number of points back to the sender. The amount sent back is not tripled. The back transfer must be a number between 0 and 480. Once the receiver has decided on the back transfer the payoffs are determined.

The payoff functions are thus as follows:

For the sender: 120 – transfer + back transfer

For the receiver: 120 + 3*transfer – back transfer

An example: Let's assume the sender transfers 40 points. At the end of the first stage the sender then has 120 – 40 = 80 points, and the receiver has 120 + 3*40 = 240 points. In the second stage, the receiver chooses a back transfer of 50 points. The payoffs then are as follows: for the sender: 120 – 40 + 50 = 130 points. For the receiver: 120 + 3*40 – 50 = 190 points.

At the end of this part of the experiment, the computer randomly determines if you are a sender or a receiver. Because you do not know if you are a sender or a receiver by now, you have to make a decision for both roles. Therefore you have to make decisions on two separate computer screens: at first for the role of the receiver, then for the role of the sender.

Please press CONTINUE to make your decision.

B.1 Instructions for Chapter 2 (EWT)

Welcome to this decision-making experiment.

Please read the following instructions carefully. At the end of the instructions you will find several control questions. After correctly answering these questions the experiment will begin. The experiment will be conducted anonymously, that is to say you will not learn with whom of the other participants you are interacting. During the experiment you can earn taler. How much you earn depends on your decisions and the decisions of the other participants. At the end of the experiment these taler will be **converted to Euro at an exchange rate of**

1 Taler = 1 Cent

and be paid out to you. Please wait until you will be called to collect your earnings. Please turn in all the documents you received to conduct the experiment when you collect your earnings.

You will begin the experiment with an initial endowment of 400 talers (4 Euro). This amount of talers will increase or decrease respectively with your earnings from the different rounds. You can always rule out losses through your own decisions.

Please keep in mind that from now on and throughout the whole experiment you are not allowed to talk to the other participants. In case you do not respect this rule we are forced to terminate the experiment. **If you have any questions, please hold your hand out of the cubicle and we will come to you.**

As a participant in this experiment you either play the role of an **employer** or an **employee**. Your specific role will be randomly determined at the beginning of the experiment and it will be the same for all the following rounds.

The experiment will last for 12 rounds. At the beginning of each round, groups of 3 participants will be formed, each consisting of one employer and two employees (employee1 and employee2). **In every new round new groups of three participants will be randomly formed.** Your decisions will only be disclosed to the other two participants in your current group. All the other participants will not be informed about your decisions.

Each round consists of two stages. **In the first stage employee1 and employee2 independently choose their amounts of work.** There is a set of ten different possible amounts of work out of which they can select their amount of work. **The smallest amount of work is 1 and the highest is 10**. Every unit of work creates 10 taler for the employer. For example: If the selected amout of work is 1 then the employer receives 10 taler, if it is 2, he will receive 20 taler and so on. If the selected amount of work is 10 the employer will receive 100 taler. For the employee the selection of an amount of work is associated with costs. The higher the selected amount of work, the higher the associated costs for the employee. The employee's costs only depend on his selected amount of work. The amount of work selected by the other employees does not influence his costs. The costs regarding the amount of work for the employee are the following:

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

So, the amount 1 of work is free for the employee. If the amount of work is 2, it will cost the employee 1 taler and so on. If the amount of work is 10 the associated cost is 20. All employees have the same cost table. The cost table will not change during the 12 rounds.

In the second stage the employer will be informed about the amounts of work selected by the employees in his group. Subsequently the **employer will chose a wage w**, which he pays out to employee1 and employee2, that is he pays w taler to employee1 and w taler to employee2. The wage per capita must not be smaller than 0 taler and must not be greater than 100 taler.

While the employer chooses the wage-level w, each employee estimates the amount of work **the other employee** in his group has selected. If his estimation proofs correct he will receive 2 taler, if incorrect 0 taler. Other than that this estimation has **no further consequences**. This estimation will not be disclosed to any of the other participants.

At the final stage of the round both employees and the employer will be informed about the selected amounts of work, the payroll of each employee and the resulting earnings of the employer.

The earnings in taler per round are thus calculated in the following way:

Earnings of the employer = 10 x amount of work of employee1 + 10 x amount of work of employee2 - 2 x wage (w)

Earnings of employee 1 = wage (w) - cost of the amount of work of employee1

Earnings of employee 2 = wage (w) - cost of the amount of work of employee2

At the end of the experiment your total earnings, comprising the initial endowment, the sum of your earnings in the different rounds and your earnings for correctly estimating the amounts of work will be paid out to you. As mentioned above 1 taler equals 1 Cent.

Please note that you are to write down all data of your group (amount of work, wage w, earnings) on the result sheet at the end of each round.

C.1 Instructions for Chapter 3 (Change Treatment)

In what follows, we present a translation of the instructions for players in the change treatment. The instructions for the no-change treatment differed only in the paragraph regarding the possibility to erase one's rating profile.

Instructions

There are two types of players in this experiment, **player A** and **player B**. You will be assigned the role of either player A or player B, and you will keep your assigned role throughout the entire experiment.

The experiment lasts **20 rounds**. At the beginning of each round the number in the upper part of the screen indicates the current round. In each round a player A is matched with a player B. This matching process is random. However, you will **never** be matched **with the same player two rounds in a row**.

Structure of each round:

At the beginning of each round each player is endowed with 10 Taler.

Player A then decides how many Taler of his endowment he wants to send to player B. Player A can send any (integer) amount between 0 and 10 Taler.

The experimenter triples this amount, so that player B receives the tripled amount of what player A decided to send.

Player B then decides how many Taler he wants to send back to player A. He can send any (integer) amount between 0 Taler and the tripled amount transferred by player A. The amount of Taler that player B sends back to player A is **not tripled** by the experimenter.

Rating:

After both players made their decisions, player A can **rate** player B's decision as **"good", "neutral" or "bad"**. The rating is stored in the rating profile. There is no rating stage if player A sends zero Taler to player B. In that case the phrase "No rating, because no exchange took place" is displayed on the screen.

Before player A decides on how much to send to player B, the **rating profile** of the player B he is matched with is displayed on the screen. The rating profile includes the rating in the previous round as well as the total number of good, neutral and bad ratings that player B has received so far.

Deleting the rating profile:

At the beginning of each round player B can **delete his existing rating profile**, and start over as a **new player**.

New players:

During the course of the experiment new players enter the experiment. Players which enter the experiment in later rounds find paper and pen at their desks and have to work on a different task before starting as a player in the computerized experiment. These players will also be paid for working on the non-computerized task at the end of the experiment. They will find a more detailed description of the task at their desks. These players will start with their decisions in the computerized experiment as soon as the screen in their cubicle displays an entry mask.

The label "New player" is used for players who deleted their rating profile, as well as for players who have just entered the experiment.

Questionnaire:

After the last round of the experiment a short questionnaire is displayed on the screen. Please answer the questionnaire as precisely as possible. When you are done with the questionnaire please wait in your cubicle until we ask you to collect your payment.

Payment: After the experiment the amount of Taler you earned is exchanged into Euros at an **exchange rate of 3,5 Cent/Taler**.

D.1 Instructions for Chapter 4 (TS Treatment)

Brief overview

Now today's main experiment starts. You already know that your former decisions do not affect your chances at this experiment. Please read the following information carefully. At the end of the information you will find some sample questions. After all participants have answered the questions correctly, this part of today's session will start.

At the beginning of the experiment **groups consisting of 4 participants will be assigned randomly.** Your income in this experiment will depend on your own decisions and the decisions of the other three members of your groups. The experiment will remain anonymous, which means that you will not know at any time who the other group members are.

The experiment consists of **one round** divided into **two stages**. At each stage the participating group members choose a **decision number** independently from each other.

The choice of the decision number is associated with certain **costs.** The sum of your decision number and a number chosen randomly by the computer will determine your **result**. The result of the first stage will determine whether you participate in the second stage: **In the first stage all four group members participate; in the second stage only the two group members with the highest results participate.**

Your income will be calculated in the following way: The group member with the highest result in the second stage will receive the amount of **20.00 Euro** and has to pay the costs of his two decision numbers. The participant with the lower result in the second stage will receive the amount of **12.11 Euro** minus the costs of his two decision numbers. The two participants who did not participate in the second stage will receive the amount of **5.73 Euro** each minus the cost of their decision number in the first stage.

At the end of the second stage, each participant will be informed about his resulting income. After this notification the experiment is over. Your income will be paid to you in cash at the end of the session.

Detailed Information

During this experiment you are in a **group together with three other participants**, so each group consists of four persons. You will not know at any time who the other members of your group are. The experiment consists of exactly **one round** divided into **two stages**. Take your time to think about your decision(s).

Stage 1

In the first stage all participants choose a **decision number** independently from each other. This number can be **between 0 and 125**.

When choosing your decision number you will produce some costs, which you can see in the enclosed cost table. In general, it can be said that the higher the decision number, the higher the cost. The attached **table** contains all **eligible decision numbers and the respective costs**. For example, a decision number "0" causes costs of 0 Euro, a decision number of 50 causes costs of 1.11 Euro, etc.

Your result depends on the chosen **decision number**. Additionally, your result depends on a **random number**. This random number is a number **between -60 and 60** randomly chosen by the computer. Each number between -60 and 60 has the same probability to be selected. The computer will choose the random number **for each participant individually** and independently, which means that the random numbers of the individual participants can be different from each other.

Your decision number (chosen by yourself) and the random number (chosen by the computer) will determine your result for the first stage:

Your result (stage 1) = your decision number + your random number

Your result is higher when your decision number is higher. That means that the **probability to reach the second stage increases with a higher decision number**.

Additionally, your result increases with the random number. For example, if you choose a decision number of 30 and the computer chooses a random number of 15, your result is 45. If you choose a decision number of 93 and the random number is -34, your result is 59.

At the end of the first stage, all participants will be notified about their random number and their result. The result of the first stage determines whether or not you participate in the second stage. Only the participants with the two highest results in your group will participate in the second stage. If two group members have the same result, the participant in the second stage will be drawn by lot. The two group members with the lowest results in the first stage will not participate in the second stage. They receive 5.73 Euro minus the cost of their decision number.

Stage 2

The process in stage 2 resembles the process in stage 1. The two group members who participate in the second stage again get to **choose a decision number**, which is associated with the costs found in the cost table. Additionally, the computer chooses a **new random number** for each of the two participants **individually**. As in stage 1, your result is determined according to the following formula:

Your result (stage 2) = your decision number + your random number

Your result in the first stage thus has **no influence** on your result in the second stage. At the end of the second stage, you will see your decision number and your random number, and you will know if you got the lower or the higher result in this stage. The participant with the lower result in stage 2 receives 12.11 Euro minus the cost of his decision numbers (in both stages). The participant with the higher result receives 20.00 Euro minus the cost of his decision numbers (in both stages).

At the end of the second stage, each participant will be informed about their income determined by their decisions and the random numbers. After this notification the experiment is over.

How is your income calculated?

Your income in this experiment depends on your result and the results of the other group members in stage 1. If you have one of the two highest results in your group, your income additionally depends on your result and the result of the other participant in stage 2.

Three cases can be distinguished:

Case 1: You do not have one of the two highest results in stage 1. In this case you do not participate in stage 2 and your income is calculated in the following way:

Income = 5.73 Euro - cost of your decision number in stage 1

Case 2: You achieved one of the two highest results in stage 1 and the lower result in stage 2. In this case your income is calculated in the following way:

Income = 12.11 Euro - cost of your decision number in stage 1 - cost of your decision number in stage 1

Case 3: You achieved one of the two highest results in stage 1 and the higher result in stage 2. In this case your income is calculated in the following way:

Income = 20.00 Euro - cost of your decision number in stage 1- cost of your decision number in stage 2

In sum, your income is higher, the higher the designated amount of money you are entitled to and the lower your costs. Please note that in cases 2 and 3 your costs incurred in **stages 1 and 2** are relevant for the calculation of your income.

If you have any further questions, please raise your hand and we will come over to answer them.

D.2 Schedule of Effort Costs

Cost Table

Decision Number	Cost in Euro
0	0.000
1	0.000
2	0.002
3	0.00
4	0.01
5	0.01
6	0.02
7	0.02
8	0.03
9	0.04
10	0.04
11	0.05
12	0.06
13	0.08
14	0.09
15	0.10
16	0.11
17	0.13
18	0.14
19	0.16
20	0.18
21	0.20
22	0.22
23	0.24
24	0.26
25	0.28
26	0.30
27	0.32
28	0.35
29	0.37
30	0.40
31	0.43
32	0.46
33	0.48
34	0.51
35	0.54
36	0.58
37	0.61
38	0.64
39	0.68
40	0.71
41	0.75
42	0.78
43	0.82
44	0.86
45	0.90
46	0.94
47	0.98
48	1.02
49	1.07
50	1.11
51	1.16
52	1.20
53	1.25
54	1.30
55	1.34
56	1.39
57	1.44
58	1.50
59	1.55
60	1.60
61	1.65
62	1.71

Decision Number	Cost in Euro
63	1.76
64	1.82
65	1.88
66	1.94
67	2.00
68	2.06
69	2.12
70	2.18
71	2.24
72	2.30
73	2.37
74	2.43
75	2.50
76	2.57
77	2.64
78	2.70
79	2.77
80	2.84
81	2.92
82	2.99
83	3.06
84	3.14
85	3.21
86	3.29
87	3.36
88	3.44
89	3.52
90	3.60
91	3.68
92	3.76
93	3.84
94	3.93
95	4.01
96	4.10
97	4.18
98	4.27
99	4.36
100	4.44
101	4.53
102	4.62
103	4.72
104	4.81
105	4.90
106	4.99
107	5.09
108	5.18
109	5.28
110	5.38
111	5.48
112	5.58
113	5.68
114	5.78
115	5.88
116	5.98
117	6.08
118	6.19
119	6.29
120	6.40
121	6.51
122	6.62
123	6.72
124	6.83
125	6.94

D.3 Elicitation of Risk Attitudes

In the following, we describe the lottery procedure that was used to elicit participants' risk attitudes for the analysis in Chapter 4. The same procedure (with slightly different payoffs) was also used in the experiment of Chapter 1.

Instructions

Before the start of today's main experiment, this first part of the experiment deals with individual decision making. This means that your income in this part only depends on your personal decisions and the outcomes of lottery drawings.

This part of the experiment deals with choices between a **lottery** and a **safe payment**. In the following, you are presented 15 situations. In each situation the lottery is the same, but the safe payment varies. In the lottery you receive 4 Euro with a probability of 50% and 0 Euro with a probability of 50%.

On the following screen you can see the 15 situations. Please decide for each situation if you want to choose the lottery or the safe payment.

At the end of the experiment one of the 15 situations is randomly chosen by the computer. According to your decision in this situation you then either take part in the lottery or you receive the respective safe payment. If you have chosen the lottery in this situation, the outcome of the lottery will also be randomly determined by the computer at the end of the experiment. Your income will be paid to you in cash at the end of the experiment.

Please press CONTINUE to make your decisions.

Lottery Choices

Please decide for every situation whether you choose the lottery or the safe payment!

Situation	Lottery	Safe Payment	Your Decision	
			Lottery	Safe Payment
1	50% Chance 4 Euro and 50% Chance 0 Euro	0.25 Euro		
2	50% Chance 4 Euro and 50% Chance 0 Euro	0.50 Euro		
3	50% Chance 4 Euro and 50% Chance 0 Euro	0.75 Euro		
4	50% Chance 4 Euro and 50% Chance 0 Euro	1.00 Euro		
5	50% Chance 4 Euro and 50% Chance 0 Euro	1.25 Euro		
6	50% Chance 4 Euro and 50% Chance 0 Euro	1.50 Euro		
7	50% Chance 4 Euro and 50% Chance 0 Euro	1.75 Euro		
8	50% Chance 4 Euro and 50% Chance 0 Euro	2.00 Euro		
9	50% Chance 4 Euro and 50% Chance 0 Euro	2.25 Euro		
10	50% Chance 4 Euro and 50% Chance 0 Euro	2.50 Euro		
11	50% Chance 4 Euro and 50% Chance 0 Euro	2.75 Euro		
12	50% Chance 4 Euro and 50% Chance 0 Euro	3.00 Euro		
13	50% Chance 4 Euro and 50% Chance 0 Euro	3.25 Euro		
14	50% Chance 4 Euro and 50% Chance 0 Euro	3.50 Euro		
15	50% Chance 4 Euro and 50% Chance 0 Euro	3.75 Euro		