# Incentives and Other-regarding Preferences

Four Essays in Experimental Economics

Inaugural-Dissertation
zur Erlangung des Grades
Doctor oeconomiae publicae (Dr. oec. publ.)
an der Ludwig-Maximilians-Universität München

2011

vorgelegt von

Sebastian Strasser

| | |
|---|---|
| Referent: | Prof. Dr. Martin G. Kocher |
| Korreferent: | Prof. Dr. Klaus M. Schmidt |
| Promotionsabschlussberatung: | 8. Februar 2012 |

Datum der mündlichen Prüfung: 30. Januar 2012

Namen der Berichterstatter: Martin G. Kocher, Klaus M. Schmidt, Florian Englmaier *to my parents*

# Acknowledgements

There are numerous people that I am grateful to for sharing with me their knowledge, their thoughts and their support.

Most importantly, I thank Martin Kocher who - both in his role as a supervisor and a co-author - introduced me to the field of experimental economics and guided me with an indescribable patience across all major and minor obstacles that I encountered during the last four years. It is by no means an exaggeration to say that without his support, this dissertation would not have been possible.

I owe special thanks to Klaus Schmidt and Florian Englmaier for agreeing to serve as supervisors of my thesis and for the numerous very helpful discussions, ideas and inspiring comments all along my doctoral studies. Furthermore, my thanks go to my co-authors Julius Pahlke, Ferdinand Vieider and Joachim Winter for the inspiring and fruitful collaboration on joint projects.

I received valuable comments from friends, colleagues and seminar participants at the MELESSA brown bag seminar, the Micro Workshop in Munich, the ESA conferences in Innsbruck 2009, Copenhague 2010 and Chicago 2011, the Royal Economic Society in Surrey in 2010, the World Meeting 2010 of the Econometric Society in Shanghai, the THEEM Meeting in Kreuzlingen 2011, the Spring Meeting of Young Economists in Istanbul 2009 and the IMEBE Workshop in Granada 2009. I also thank the German Science Foundation for its financial support and the entire staff involved in conducting my experiments at MELESSA.

Lastly, I am grateful to Constanze, my parents and my sister for their patience and their inexhaustible support in everything I do.

Sebastian Strasser

# Contents

# List of Tables

# List of Figures

# Preface

The interaction of economic agents is at the heart of any economic analysis. Achievements in game theory and contract theory over the last decades have provided powerful tools and methods that enabled a rigourous analysis of the optimality of behavior, incentives and contracts. Many of these analyses rested upon the assumption of a so-called ”homo oeconomicus“, i.e., an economic agent that is exclusively motivated by his very own fate, but neglects anything and everybody around him. While models based on this assumption have yielded enormously strong predictions, evidence emerged where observed behavior was clearly at odds with theoretically optimal behavior and hinted at a ”non-selfish“ decision maker who cares also for his social environment. In retrospection it may come as a small surprise - particularly to non-economists - that the economic science has for a long time been losing sight of the fact that the well-being or the utility of an economic agent not only depends on what accrues to him or her, but also on what accrues to others in a (possibly common) interaction. The importance of this human characteristic gave birth to a field of economic research which incorporates this dissertation, namely the study of other-regarding concerns, i.e., the formalization and testing of models as well as the provision of empirical evidence that encompass a ”social“ element in the preference structure of an economic agent.

Up to date, there are two distinct ways other-regarding concerns have found their way into economic theory. Pioneered by the works of Fehr and Schmidt (1999) and Bolton and Ockenfels (2000), economic agents are assumed to have some sort of preferences for a payoff distribution by comparing their outcome to the outcomes of others with a general notion of disliking unequal or ”unfair“ allocations. These models remain to a large extent agnostic about how outcomes are actually achieved but focus on the final distribution of outcomes. A different way to incorporate notions of concerns for others into the decision making process has been formalized by Rabin (1993) and Dufwenberg and Kirchsteiger (2004) who concentrate on the intentions behind actions and rely on notions of kindness and reciprocity to explain an economic agent's non-selfish behavior.

The presence of other-regarding concerns has brought up new challenges particularly for the study of incentives, since the optimality of an incentive scheme under the self-interest model may vanish when an economic agent also takes into account a counterpart's intentions or the consequences for the other's outcome.[1] As an example, non-selfish preferences have been shown to act as a remedy against contractual incompleteness under moral hazard which has important consequences for labor markets. In the gift-exchange framework, a principal's optimal contract offer under selfish preferences may lead to an inefficient outcome when it is perceived as unfair by an agent with non-selfish concerns and reciprocated with rejection or shirking. Optimal incentives in the sense that they yield efficient or desired outcomes may therefore be very different for agents with non-selfish preferences than for selfish ones.

The parsimonious representation of models of other-regarding concerns has contributed greatly to the possibility to test them with empirical data, however, they leave considerable degrees of freedom towards the environment and the circumstances of a decision. E.g. the reference group in the model of Fehr and Schmidt (1999) can not be unambiguously be determined in every situation, or to what extent an action is considered as "kind" in the model of Rabin (1993) varies with the circumstances of the decision. The environment of a decision determines to a great degree the predictive power of models of other-regarding preferences and the desired effect of incentives. From the study of these models, it is not clear where they precisely apply and if different institutional environments amplify or reduce concerns for others. The presence of competition as an example may under certain conditions completely eliminate the effects of social preferences.

The main contribution of this dissertation is therefore to provide evidence how other-regarding concerns translate into behavior when the environment changes and the circumstances of decisions are altered. A further contribution relates to the well-documented heterogeneity of other-regarding concerns in the population and to what extent this has to be taken into account for the design of appropriate incentives. The strength of models in

[1]For an assessment of the impact of non-selfish preferences on general equilibrium outcomes, see e.g. Dufwenberg et al. (2011).

this field for predictions and policy implications crucially depends on a profound knowledge of under what conditions other-regarding concerns play which precise role. Next to fostering an understanding of real world phenomena, it is the purpose of all chapters in this dissertation to provide evidence about how to set the behaviorally optimal incentives when facing economic agents that are not solely driven by selfish motives.

Theoretical achievements in the field of social preferences have been paralleled by the development of experimental methods in economics. Laboratory studies appear particularly useful given the high degree of control they provide in contrast to standard empirical analyses. For the study of the interaction of incentives and other-regarding concerns, laboratory experiments act complementary to the development of empirical methods (amongst them field experiments) and theoretical models by providing a way to come up with clean empirical evidence and the possibility to test competing theories against each other or refine theoretical analyses e.g. in the presence of multiple equilibria. As well-documented in the history of natural sciences, experiments can also be explorative in nature and can thus be sources of unexpected discovery that subsequently inspire new ways of thinking and foster the development of new theories. All four chapters in this dissertation use the experimental method to provide evidence about instances where other-regarding concerns shape outcomes and interact with the incentives under which experimental subjects take their decisions.

The first two chapters are embedded in the field of experimental labor markets and use the gift-exchange game as a workhorse of the analysis. The third chapter is closely linked to the prior chapter through accounting for individual heterogeneity and its consequences, but does this in a public goods setting. In chapter four, concerns for others are looked at in the field of choices under risk, a field where social preferences have only started to be accounted for recently. All four chapters contribute to assess experimentally how other-regarding preferences translate into changing and competitive environments and focus on the interaction of incentives and the underlying (possibly heterogeneous) preference structure of individuals.

The first chapter, which is joint work with Martin Kocher, focuses on the existence of the so-called "fair-employment-hypothesis“ on labor markets which stipulates that firms employ workers even if it is not profitable for them in the short run and refuse to use unemployment as a disciplining device when they suffer from a negative productivity shock.

Many studies have confirmed the well-known fair-wage hypothesis (Akerlof, 1982), i.e., a robust positive relationship between the wage and non-enforceable effort on labor markets characterized by moral hazard. However, none of these studies has accounted for the fluctuations from demand or productivity shocks and their effects on the robustness of gift-exchange outcomes. In the presence of a negative productivity shock a firm may not be able to pay the same levels of wages, such that a static interpretation between wages and efforts at one point in time has severe limitations. The idea behind this chapter is therefore to test the dynamic counterpart of the fair-wage hypothesis where reciprocity stretches over to behavior across periods. To do so, we pay particular importance to the concept of relational contracts, i.e., the possibility for firms and workers to engage endogenously in long-term relationships to see how robust these relationships are in the presence of labor market fluctuations that affect firms.

In the design of the experiment, we extend the framework of Brown et al. (2004) by the existence of commonly known productivity shocks to firms where in a one-sided auction firms submit public or private wage offers and workers - identifiable through an identification number - can accept posted contract offers. There is excess supply of workers and firms are restricted to hire a maximum of one worker.

Our results indicate that reputational mechanisms between firms and workers help to prevent unemployment or even a market breakdown in recessions and confirm gift-exchange to be a robust phenomenon also under unstable productivity levels. When the economy enters a recession, firms cut wages significantly but workers reduce effort levels only marginally such that a new wage-effort relation emerges in periods of low productivity. There is evidence for intertemporal reciprocity between firms and workers in the sense that relationships with high levels of effort in the past are characterized by a stronger decrease in wages when the economy enters a recession compared to newly formed firm-worker pairs. We interpret this finding through an increased flexibility of the labor market inside the firm through relational contracts that is absent when hiring a new worker from outside the firm. In contrast to stable conditions of productivity, however, we identify a series of frictions in the contracting behavior on a labor market with varying productivity where efficiency is also lowest.

In two control treatments, reputation building is made impossible through allocating worker IDs randomly every round and wages are exogenously fixed to isolate the impact of reputation mechanisms and intentions behind the wage level in presence of productivity shocks. We find that reputation mechanisms are crucial to sustain high employment levels in recessions and that fixing the wage level has a surprisingly positive effect on market efficiency especially in recessions. We explain the latter finding through reciprocity between being offered a job itself (rather than the wage level) and effort levels as well as through subjects caring strongly about the surplus split.

We furthermore extend our analysis to a setting where firms are allowed to hire a second worker at a lower marginal productivity. Firms can now decide whether they reduce wages for their entire workforce or lay a low-performing worker off when hit by a recession, see also Bewley (1999). Wages are either flexible or exogenously fixed according to the treatment. We find that when firms are not able to reduce wages in a recession, this produces a strongly cyclical employment pattern with close to full employment in booms, but a significantly reduced workforce in recessions. Under flexible wages we do not observe any cyclicality in employment, since firms buffer shocks through lower wages, which is accepted by workers.

Our study is subjecting the results from the prior experimental literature on labor markets to a rigorous test as to what behavioral patterns survive and emerge when introducing fluctuations or - more generally - when adding more realistic features. The results underline the importance of dynamic considerations for individuals who have some sort of other-regarding concerns and foster our understanding of the incentives on markets in an incomplete contract environment.

In a joint project with Florian Englmaier and Joachim Winter, which constitutes chapter two, we analyze the impact of available information about worker characteristics again in a gift-exchange setting. We address the question how - and if at all - worker characteristics are able to predict behavior in a subsequent moral hazard situation and whether information about worker traits is used by firms when writing contracts. In contrast to much of the experimental literature on labor markets that focuses on the enforcement of incomplete contracts, our study is among the first to shed light on the adverse selection element in the process of entering the labor market. Given the well-documented heterogeneity of individuals with respect to their skill set and their preferences, we examine how this heterogeneity is taken into account by labor market participants and how it shapes final outcomes in interactions characterized by non-enforceable effort.

We design an experiment where we elicit in a first part two dimensions of a worker's characteristics that we consider to be important for the worker's effort decision in a subsequent gift-exchange game. We focus on the worker's productivity in a real-effort task and the behavior in a binary trust game as a proxy for social preferences and make this information accessible to employers. In the second part of the experiment, firms can condition their wage offers on the information they have about workers before firms and workers interact in a standard gift-exchange framework. In a one-shot setting, we explicitly exclude the possibility to build up a reputation and hence concentrate on the pure effect of information from sources that are exogenous to the relationship in contrast to studies where information about worker types arises endogenously within a firm-worker relation.

Our results indicate that firms pay wage premia to workers with a high productivity and high trustworthiness measure. Gift-exchange is robust in our real-effort task across all types of workers, but turns out to be strongest for trustworthy workers. Firms use the information about workers to tailor incentive schemes and to make use of complementarities between wage levels and worker types. From the effort choices, we find that optimal wages are highest for workers that combine both traits (productivity and trustworthiness) which lead to maximal profits for firms. Only in the interaction with trustworthy workers an increase in the wage produces significantly higher profits for firms. In a control treatment, we show that subtle differences in the presentation of information about productivity (switching from a binary measure to a continuous measure) have non-negligible effects for outcomes and produce different endogenously generated wage distributions.

We explain our findings by the importance of different levels of complementarities between incentives and worker types. A wage increase has a different effect on workers with high or low degrees of productivity or other-regarding concerns yielding different wage-effort relations for worker types. The results suggest that worker heterogeneity particularly with respect to trustworthiness has to be taken into account when setting the right incentives in gift-exchange situations. Information about the contracting partner shows up as a vital ingredient in the contract writing phase. Our analysis underlines the importance of available worker information since it allows firms to control the interaction between worker heterogeneity and the solution to the moral hazard problem.

The third contribution of this dissertation also highlights the importance of heterogeneity of other-regarding preferences, but this time with respect to tournament incentives in a social dilemma. Tournament incentives are widely used - explicitly or implicitly - in organizations to elicit additional effort from agents or when individual effort is prohibitively costly to monitor and cannot be easily observed in contrast to final output. Social dilemmas constitute an important class of situations where contributions to a common pool have positive externalities on other group members. Hence, individually rational and socially efficient behavior diverges. The effect of competition between teams in social dilemma situations has been well-documented. However, many tournament incentives in organizations work at the individual level such that an individual tournament in a social dilemma creates a trade-off: Contributions to the public good have a greater positive externality on others as they improve others' chances of a tournament prize, but decrease the own individual likelihood of obtaining the prize. Prominent examples within organizations include e.g. promotions at the individual level while working in a team. The question that I address in this chapter is how different types of subjects react differently to the introduction of tournament competition in a social dilemma framework. Pertaining to the design of the experiment, all subjects are classified according to the conditional contribution exercise by Fischbacher et al. (2001). In a public goods framework, I subsequently expose subjects to tournament incentives at the individual level such that there is an exogenously fixed prize for a subject if the proceeds from cooperation within a group exceed those of a competitor under identical incentives in a different group. In a first experiment, subjects interact repeatedly in a partner setting to allow for reputation and strategic behavior over time and prizes are awarded on the basis of cumulative earnings. In a second experiment, the prize is awarded on the basis of behavior in a one-shot game using the strategy method to see how competition affects the classification of subjects into types.

The results suggest that in the repeated setting, competition destroys the incentives for freeriders to build up a strategic reputation of being a cooperative "type". Subjects classified as conditional cooperators are unaffected by the presence of competition. In the absence of reputational concerns in a one-shot setting, there is evidence that in the presence of competition some subjects cease to act as conditional cooperators altogether such that the measured proportion of freeriders is significantly higher under competition compared to the baseline treatment. This confirms the argument that other-regarding concerns stop to play a role when individuals lose control over the final allocation of payoffs which is caused by competition.

The study is among the first to provide evidence for a differential impact of an exogenous treatment variation on different types of subjects which warrants a careful interpretation of treatment effects. The results furthermore have important consequences for the impact of a change of the institutional environment on observed individual behavior and naturally for setting the right incentives when one is willing to induce a desired behavior. The same incentive may have very different effects on different types identified, which underlines the importance of availability of information about types for the contract designer. Since randomization of subjects into treatments takes place, the observed effects constitute a strict lower bound of the differential effect from the competitive incentive on different types of agents. The presence of selection or sorting of different types into different incentive schemes will amplify the variance of outcomes from the implementation of the same incentive scheme for different types.

In the last chapter, which is joint work with Julius Pahlke and Ferdinand Vieider, we focus on the interaction of other-regarding preferences and individual risk attitudes to see how the latter are affected when the decision maker decides not only for herself but for another person as well. To date, risk preferences have been studied predominantly in a context of individual decision making. We argue that the consequences of a substantial part of decisions taken under risk in real world situations affect not only the decision maker himself but also other persons and are hence taken under responsibility for others. Examples include parents taking decisions for their kids or managers for an entire company and thus their shareholders. Any differences found between individual risk preferences and risk preferences under responsibility will undermine the predictability of decisions from the knowledge acquired about risky decision taken individually. Given that decisions under responsibility constitute an important class of decision situations - and indeed one that in its economic importance may even surpass individual decisions - we present evidence on differences between the two and hence provide important insights for descriptive as well as prescriptive policy purposes.

In a between subjects design we compare lottery decisions observed in an individual treatment, i.e., a decision maker's choice only affects her own income, with those made in a responsibility treatment, i.e., a decision maker's choice affects her own income as well as the income of an anonymous recipient in exactly the same way. With perfect income matching for the decision maker and her recipient, we explicitly exclude preferences over outcome distributions and inequality concerns to have an impact on the decision under risk in the responsibility treatment. If the decision maker accommodates any presumed preferences of the recipient or when following some social norm, the own payoffs are affected in the same way and she will incur an actual cost compared to a decision taken only on her own account or only for somebody else. Therefore, our findings are a conservative measure and constitute only the lower bound on the effects of responsibility on risk preferences that we aim to investigate.

Given the systematic differences of risk preferences over probability and outcome spaces, we expose in a first experiment subjects to simple lottery choices that are either in the gain domain, the loss domain and the mixed domain including gains and losses holding the probabilities constant at 50%. In a second experiment, we extend the analysis to different probability levels.

We confirm the intuition that being responsible for somebody else's payoffs increases risk aversion for gains, while in the loss domain we find increased risk seeking in our first experiment. In the second experiment, we replicate the finding of increased risk aversion

for large probabilities of a gain, while for small probabilities we find an increase of risk seeking under conditions of responsibility. These findings discredit the hypothesis of a unilateral cautious shift through taking over responsibility for somebody else, but indicate that the fourfold pattern of risk attitudes predicted by prospect theory is accentuated in our responsibility treatment.
Our results suggest that risky decisions under responsibility differ significantly from individual choices under risk, which appear to be reinforced by the presence of responsibility for others. These findings have important policy implications for the design of optimal contracts as the identified patterns under responsibility may be seen as suboptimal from a risk neutral principal's point of view.

All four chapters of this dissertation are self-contained and include their own introductions and appendices such that they can be read independently. The respective appendices contain the instructions of the experimental protocols.

# Chapter 1

# The Fair-Employment Hypothesis: Reciprocity in unstable environments[0]

*Good times, bad times, you know I had my share.* (Led Zeppelin)

## 1.1 Introduction

In their seminal papers, Akerlof (1982) and Akerlof and Yellen (1988, 1990) postulate a positive relationship between non-contractible work effort and wage. If such a relationship exists, it may be optimal for employers to pay wages that are above the market-clearing level. Numerous laboratory experiments, using the gift-exchange game to assess the existence of a fair wage-effort relationship, established that with incomplete contracts average wages indeed exceed the marginal product of labor (Fehr et al., 1993, 1998a,b; Fehr and Falk, 1999). While gift-exchange is not robust under all conditions in the field (Gneezy and List, 2006), it still seems to be a widespread phenomenon on real labor markets (Bellemare and Shearer, 2009; Kube et al., 2010).[1]

This chapter investigates experimentally whether the static fair wage-effort hypothesis

[0] This chapter is based on joint work with Martin Kocher.

[1] Excellent overviews of the literature are provided by Fehr and Gächter (2000), Cooper and Kagel (2009), and Charness and Kuhn (2011).

has a dynamic sister, the 'fair employment hypothesis'. Extending the original fair wage-effort hypothesis, the fair employment hypothesis stipulates that, in an incomplete contracts setup, firms keep workers employed after negative productivity shocks (in times of economic distress for the firm) in order to induce them to behave (more) reciprocally. Reciprocity or gift-exchange in a setting with alternating high potential profits (simply denoted 'good times' in the following) and low potential profits or even losses (henceforth 'bad times') can be established through additional channels than the ones that are available in a static setting. In the static fair wage-effort relationship kindness is signaled by higher than minimal wages and higher than minimal effort levels. With alternating good and bad times, kindness can also be indicated by keeping a worker employed in bad times or by specific inter-temporal patterns of wage offers, respectively wages, as well as effort levels. Whether such inter-temporal forms of gift-exchange exist, to what extent they can be put to work, and under which prerequisites they are beneficial for firms and workers, is at the heart of this chapter.

We model an incomplete contracts environment (Brown et al., 2004) with firms (principals) and workers (agents). Firms offer contracts consisting of a wage and a non-binding desired effort level to workers on a labor market. Workers can accept contract offers and, if they accept, have to choose an effort level, with a cost-of-effort function that is convex in effort. The higher the effort level, the more beneficial it is for the firm. In order to induce alternating periods of high potential profit and low potential profits, we add market-wide and symmetric negative productivity shocks to this standard setup. The productivity shocks can be viewed as an analogy for the cyclical nature of the economy. More precisely, we analyze gift-exchange in phases with high and low levels of effort and productivity. Hitherto, almost all gift-exchange experiments have implemented stable economic conditions. Our design provides both comparative static results for different productivity levels as control treatments and real dynamics. In addition to the productivity shocks, our experiment exogenously varies in a systematic way (i) the maximal size of the firm (i.e., how many workers a firm can employ in a given period), (ii) the level of wage rigidity, and (iii) the possibility for both interacting parties, firms and workers, to form reputation over time, because the impact of the three dimensions could interact with the variation in productivity levels.[2]

Our extension of the basic gift-exchange setup allows us to answer several research questions that have not been addressed so far in the literature. *First*, we are able to analyze how negative productivity shocks affect reciprocity among firms and workers as well as unemployment rates on the labor market. In most treatments, our experimental design implies full employment in good times and full unemployment in bad times, when using standard assumptions. Our results, on the contrary, indicate that the existence of (inter-temporal) gift-exchange can overcome the dire implications of negative productivity shocks, and unemployment in excess of the natural rate (due to excess supply of labor on the market by construction of our setup) is very small in bad times. Wages and effort levels are significantly higher than predicted by standard theory, but they respond to the productivity shocks. In bad times, wages drop sharply without inducing a strong negative reaction of worker's effort levels. For given wage levels, employees voluntarily work much harder in bad times than in good times.

In a *second* step, we assess the driving forces behind our main results. Unemployment in bad times could remain low because of the wage flexibility firms have in our setup. However, even if the introduction of wage rigidity[3] makes gift-exchange between firms and workers more difficult, unemployment is non-existent such that wage flexibility is not the main driver of our results. Gift-exchange is surprisingly robust over time with rigid wages. In a similar vein, the elimination of the option to form long-term relationships, i.e., to invest in reputation, reduces employment levels, but to a much lesser extent than expected. Reciprocity is still strong enough to sustain a considerable number of completed contracts between firms and workers in bad times. Moreover, rigid wages lead to a higher overall market efficiency than flexible wages. With flexible wages firms tend to lower wages to an extent that reduces overall market efficiency in bad times. From the comparison of our results from the markets with and without reputation mechanisms, we can conclude that it is not the disciplining effect from the threat of becoming unemployed in bad times that

[2]In the following we will for the sake of succinctness always refer to the labor market analogy of the gift-exchange game. Our setting also captures the (potential long-term) relationship between a buyer and a seller that interact under incomplete contract conditions, when the seller can choose the quality of the provided good after the payment was made. For instance, in business-to-business relationships productivity shocks in the way we model them might influence the relationship between the buyer and the seller.

[3]Our setup is similar to Charness (2004), who also eliminates all wage-related reciprocity.

induces workers to work harder in bad times, because without reputation in a repeated interaction there is no such threat. It is the worker's kind reaction on the kind action by firms to offer them a contract in times in which the profit prospect for firms is bad that drives our main results.

A *third* contribution of this chapter is the study of the micro-mechanisms that help firms and workers overcome bad times through inter-temporal gift-exchange. We find evidence for workers exerting higher effort levels in bad times, when they experienced a higher wage-effort relation in the past. Furthermore, we identify relational contracts as a means of making the labor market *inside* the firm more flexible, with the existing workforce accepting wage cuts in bad times, facilitating the emergence of long-term relationships. In contrast, the labor market *outside* the firm for hiring new workers is characterized by a higher degree of wage rigidity during bad times.

Our experimental design is also inspired by Bewley (1999), whose starting point was the observation of rigid wages, see also Blinder and Choi (1990) and Campbell III and Kamlani (1997) in recessions. The persistent occurrence of rigid wages are in contrast to standard economic theory. Complementing his empirical results from unstructured interviews with a large number of, primarily, human resource managers in the Northeast of the United States, our experiment allows to analyze to what extent and, potentially, under what circumstances wage rigidity is bad and how the availability of reputation mechanisms (i.e., the development of endogenous relational contracts) influences (labor) markets in times of economic distress.[4] As already mentioned, our results indicate that with incomplete contracts, wages are not rigid and that exogenously implemented wage rigidity is not causing considerable unemployment. However, one of the main arguments in Bewley (1999)'s book is that employers prefer to lay off workers in times of economic distress in order to get 'the problem out of the door'. Such a motive cannot be captured in our one-firm-one-worker setting. The *fourth* contribution of this chapter is therefore to experimentally implement a multiple worker setting that is able to provide an empirical test of the motive. More specifically, we implement additional experimental treatments in which firms can employ more than one worker, and the second worker exhibits a lower

[4] Bewley (1999) distinguishes between 'primary-' and a 'secondary-sector' jobs that differ with respect to the possibility of developing a reputation as a firm or worker. It is exactly this distinction that is also important in our experiment.

productivity than the first one. These treatments allow us to assess the trade-off between wage and employment policies of the firm. Even though we observe more unemployment in this multiple worker setting, the level of unemployment is still much lower than expected theoretically. Depending on the degree of exogenous wage rigidity, firms use both the employment and wage levels to counter a productivity shock.

The remainder of this chapter is structured as follows: After a short overview of the related literature in section 1.2, we outline the experimental design, describe the laboratory protocol, and discuss the theoretical predictions in section 1.3. Section 1.4 provides the experimental results for our main treatments. In section 1.5, we extend our main treatments to a multiple worker setting. Finally, section 1.6 discusses our results and concludes the chapter.

## 1.2 Related Literature

Contractual incompleteness is an omnipresent characteristic on many markets, and it has been studied extensively in economics. Often, obligations of market participants are only specifiable or deliberately specified imprecisely, and relations are influenced by informal rules or unwritten norms. It is hard or even impossible for third parties to enforce such relational or implicit contracts because, typically, outsiders are unable to verify whether contractual obligations have been met. In the context of labor markets it is obvious that, even if many agreements are very explicit about the compensation that accrues to a worker, they are imprecise when it comes to specify the tasks and obligations a worker has to accomplish. As a consequence, potential incentive problems arise. The theoretical literature offers several ways to overcome the moral hazard created by contractual incompleteness. In the context of labor markets, there are theories that focus on the disciplining version of the efficiency wage hypothesis (Gintis, 1976; Shapiro and Stiglitz, 1984; Bowles, 1985) as well as the above-mentioned fairness versions of the efficiency wage hypothesis (Akerlof, 1982; Akerlof and Yellen, 1990).[5] In addition, theoretical models that formalize the role of reputation and implicit contracts in repeated interactions on labor

[5]Fairness models such as Rabin (1993), Fehr and Schmidt (1999), Bolton and Ockenfels (2000), Charness and Rabin (2002), and Dufwenberg and Kirchsteiger (2004) provide similar predictions as the fairness versions of the efficiency wage hypothesis.

markets have been developed to capture the incentive effects arising from self-enforcing non-written firm-worker agreements (Bull, 1987; MacLeod and Malcomson, 1989, 1993, 1998; Baker et al., 2002; Levin, 2003; MacLeod, 2003; Fuchs, 2007).

Empirical assessments are especially desirable in the context of contractual incompleteness because many models in the theoretical literature exhibit multiple equilibria.[6] Based on their earlier work, Brown et al. (2004) provide evidence that long-term relationships between trading parties emerge endogenously in the absence of third party enforcement. Low effort is penalized by the termination of the relationship, which is a powerful contract enforcement device. If third-party enforcement is not available, markets are split up into bilateral long-term interactions that are sustained through reputation mechanisms.

Adding cyclical ups and downs in the form of productivity shocks to the setting is not only interesting in itself, because they can be viewed as an analogy for real business cycles, but it also helps to assess the anatomy of reciprocity and the determinants of the effects that allow for a bilateralization of interactions.[7] For instance, one of our main results – the low levels of unemployment even in bad times, alongside relatively high earnings of workers – together with the comparison of the results from exogenously implemented rigid wages and flexible wages indicate strongly that fairness theories are able to explain the data to a better degree than disciplining theories.

To the best of our knowledge, we are the first to address the research questions arising from cyclical market instability on reciprocity in the laboratory. Most closely related to this chapter is a study by Linardi and Camerer (2010), who subject relational contracts to stochastic interruptions where firms cannot hire workers for three periods after being hit by an idiosyncratic shock. Despite the random shocks they find firm-worker relationships to be robust. Gerhards and Heinz (2011) run a two-period gift-exchange game with a positive probability of an "economic crisis“ to realize in the second period. Even though the setup is quite different from ours, the reduction in wages and the stability of effort levels in bad times that we find is replicated in their study. Other features of the labor market that are relevant in the context of our study and their effect on experimental gift-

[6] In the following, we are discussing only a very small fraction of the whole empirical/experimental literature. The surveys mentioned in footnote 1 provide an excellent overview.

[7] Albeit of different length and character, cyclical patterns are a persistent phenomenon on many markets such as, for instance, labor and capital markets (Lucas, 1977; Kydland and Prescott, 1982; Long and Plosser, 1983; Hansen, 1985).

exchange outcomes have been studied more extensively, e.g. minimum wage laws by Falk et al. (2006). Their main result is that minimum wages have a persistent impact on workers' reservation wages and fairness perceptions by creating entitlement effects. Brandts and Charness (2004) do not find significant differences between markets with excess supply of labor or excess supply of firms, but they are also able to document that minimum wages have a negative effect on effort provision and, thus, overall market efficiency, even though the effect is small. Falk et al. (2008) find strong negative effects of dismissal barriers implemented as the obligation to hire workers for several consecutive periods at initially agreed wages; an effect that is offset by introducing non-enforceable bonus payments that firms can pay after observing actual effort levels. The role of unemployment as a disciplinary device is analyzed by Brown et al. (2011), who find that unemployment is indeed not a necessary device to motivate workers to provide above-minimum effort levels on markets with excess demand for labor.

There is also a nascent literature on the effects of social comparison on the wage-effort relation in experimental gift-exchange games with multiple workers that is relevant in the context of our treatments in which firms can employ more than one worker (see, e.g., Maximiano et al. (2007), Mittone and Ploner (2009), Abeler et al. (2010), Gächter and Thöni (2010), Kocher et al. (2010), Angelova et al. (2011), Gächter et al. (2011)). We will discuss the relevant literature directly in section 1.5.

## 1.3 Model, Experimental Design, and Theoretical Predictions

### 1.3.1 The Basic Model and Productivity Shocks

We use a simple gift-exchange environment (similar to Brown et al. (2004)), consisting of two stages per period. In the first stage, firms can make binding contract offers $\{w, \tilde{e}\}$ to workers, with $w$ denoting the wage and $\tilde{e}$ the desired effort level.[8] Workers that accept a contract have to exert effort $e$ in the second stage. Effort determines the employer's benefit

[8]We make interchangeable use of the terms 'firm' and 'employer' henceforth. Along the lines of Cabrales and Charness (2003), we will refer to the firm as being female and the worker as being male.

$\beta B(e)$ at some cost to the worker $\zeta C(e)$. We make the following standard assumptions regarding the functional forms: (i) $B'(e) > 0, B''(e) \leq 0$, and (ii) $C(0) = 0, C'(e) > 0, C''(e) > 0$. In the experiment, the first stage is a trading period of 180 seconds in which employers can make contract offers either publicly on a posted offer market or privately to a specific worker, a feature introduced into the literature by Kirchsteiger et al. (2001). Contracts were restricted to the range $\{w \in \{0, 1, ..., 100\}, \tilde{e} \in \{1, 2, ..., 10\}\}$. Workers can accept any standing contract offer during the trading period, and if one does, he exerts (costly) effort $e \in \{1, 2, ..., 10\}$, where $\tilde{e}$ is non-binding (i.e., contracts are incomplete). The cost-of-effort function that we use is shown in Table 1.1.

Table 1.1: Cost of Effort

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| $e$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| $c(e)$ | 0 | 1 | 2 | 4 | 6 | 8 | 10 | 12 | 15 | 18 |

Each firm can hire at most one worker, and each worker can be hired by only one firm. The material payoff of a firm is given by

$$\pi_f = \begin{cases} p_t \cdot e - w & \text{if a contract was concluded} \\ 0 & \text{otherwise} \end{cases}$$

and the payoff of a worker is given by

$$\pi_w = \begin{cases} w - c(e) & \text{if a contract was concluded} \\ b_{unemp} & \text{otherwise} \end{cases}$$

Employers can make as many offers as they want during a trading period, as long as none of them is accepted. After a worker has accepted a contract of a specific firm, all standing contract offers of this firm are automatically deleted. Public offers are public to all other firms and all workers on a market. At any time in the trading phase all market participants know which firms and which workers have already concluded a contract.[9]

[9]See Figures A1.12 and A1.13 in appendix A1.6 for screenshots of the computer screens used in the experiment.

After the worker's effort choice, both one's own profit as well as the decisions and profit of one's trading partner in the respective period are displayed, and a new period begins. At the beginning of each period, all workers are unemployed. The experiment lasts for 15 periods, and all participants keep the same role throughout an experimental session.
The productivity parameter of the firm, $p_t$, is used to implement market-wide productivity shocks.[10] We decided to apply the most basic form of productivity instability one can think of, with two levels of $p_t$ and common knowledge of their occurrence for all subjects right from the start of the experiment, according to Table 1.2. The two levels of $p_t$, $p_t = 10$ and $p_t = 5$, are denoted good times (GT) and bad times (BT), respectively. Remember that we are interested in the incentive and potential disciplining effects of economic fluctuations. In order to be able to focus purely on them, we deliberately eliminate any kind of uncertainty and any noise from the order, timing and size of the productivity shocks. The simplification buys us a high level of control over the experimental setup. Productivity shocks apply market-wide to all employers symmetrically.

Table 1.2: Productivity Parameter of the Firm in GTBT

| **period** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $p_t$ | 10 | 10 | 10 | 5 | 5 | 5 | 10 | 10 | 10 | 5 | 5 | 5 | 10 | 10 | 10 |

Our laboratory markets consist of twelve subjects each, with five firms and seven workers. We thus follow the standard in the literature of implementing excess supply of labor. Unemployed workers receive a fixed unemployment benefit of 6 experimental currency units. Even in BT, maximal effort is efficient, since the marginal benefit of effort ($p_t$=5) still exceeds the highest marginal cost ($\Delta c_{max} = 3$). Furthermore, the incentives for workers are unaffected by the condition of the economy as only firms' profits depend on $p_t$. The contracted wage serves exclusively as a distributional device, splitting up the surplus between the firm and the worker, but it has no direct impact on efficiency.

[10] It can be interpreted in terms of demand fluctuations, variations in prices over time, or other exogenous shocks to profits.

### 1.3.2 Experimental Treatments and Laboratory Protocol

We implement an unbalanced 3x2x2 factorial design with the following factors: (i) productivity levels, (ii) wage flexibility, and (iii) reputation. Productivity levels can either be varying over time according to Table 1.2 (Treatment GTBT), or be fixed over the 15 periods on the high level $p_t = 10$ (Treatment FGT for 'fixed good times'), and fixed over the 15 periods on the low level $p_t = 5$ (Treatment FBT for 'fixed bad times'). Treatments FGT and FBT can be interpreted as control treatments for GTBT.[11] Treatments GTBT, FGT and FBT are implemented with flexible wages and with fixed IDs over the course of the experiment, i.e., we allow for reputation formation.[12] Treatment GTBT is also implemented with the wage level fixed at $\overline{w} = 30$ in order to capture the possible effects of wage rigidity (Treatment GTBT_FIX) and with changing IDs over time in order to make reputation formation impossible (Treatment GTBT_RI for 'random ID').[13]

The experiment was conducted computerized using the software package zTree (Fischbacher, 2007) at the MELESSA laboratory of the University of Munich in 2009. Participants were randomly recruited from the undergraduate population of the University with the help of ORSEE (Greiner, 2004). No subject participated in more than one session. Upon arrival, students were seated at computer screens divided by blinds. The instructions were distributed and read aloud by the experimenter. They were framed in neutral terms (e.g., employers were called 'type A', the wage was denoted 'transfer', and so on).[14] Before the first period, subjects were assigned their roles. They kept their roles throughout the entire experiment. Yet, subjects were completely anonymous in all treatments. After 15 periods the experiment ended, and participants were paid the sum of their earnings in private. A typical session took less than two hours, including instruc-

[11] FGT is close to a replication of one of the treatments in Brown et al. (2004), with a different size of the market and a couple of other smaller differences. Our results for this treatment are indeed very similar to those in Brown et al. (2004).

[12] All subjects in all treatments were given a documentation sheet in which they were asked to take note of outcomes (wage, effort, profits, ID of transaction partner) in every period in order to help them track the history of interactions. We did not require subjects to fill out these documentation sheets, but told them it could be useful for their decisions. Almost all subjects took great care in following what happened during the experiment and filled out the sheets throughout the entire experiment.

[13] The level of the fixed wage, $\overline{w}$, in GTBT_FIX was chosen after GTBT was conducted. See details regarding the choice of $\overline{w}$ in the results section.

[14] The experimental instructions to the FGT, GTBT, and FBT treatments can be found in appendix A1.7. Others are available on request.

tions and payment.

All 372 participants were endowed with an initial endowment of 400 points. Experimental currency units in the experiment (called 'points' in the instructions) were converted to Euro at the pre-announced exchange rate of 0.10 €/point in all treatments. Table 1.3 shows the number of markets (with twelve subjects each) for each treatment as well as the average earnings in each treatment.

Table 1.3: Treatment Overview

| | Flexible Wages | | Fixed Wages | |
|---|---|---|---|---|
| | Reputation | No Reputation (GTBT_RI) | Reputation (GTBT_FIX) | No Reputation |
| FGT | 5 markets (35.7 €) | - | - | - |
| GTBT | 6 markets (24.6 €) | 8 markets (18.5 €) | 6 markets (28.3 €) | - |
| FBT | 6 markets (19.6 €) | - | - | - |

Treatment overview with number of experimental markets and average earnings per subject in parentheses.

### 1.3.3 Predictions and Hypotheses

**Standard Predictions: Selfish Players**

Assuming payoff maximization, risk neutrality, and common knowledge, there is only one subgame perfect equilibrium of the finitely repeated game: Workers in each period exert minimal effort levels and, anticipating this, firms offer contracts $\{w = 6, \tilde{e}\}$, with wages that are equal to their outside option (i.e., the unemployment benefit of 6 points in the experiment).[15] In FGT each firm will employ a worker and reap all the rent from the interaction (i.e., 4 points). Two workers on the market remain unemployed by construction of our design. In FBT, it is easy to see that the outside option is more attractive for workers than the wage that firms are able to offer ($w = 5$). Thus, no firm will engage in trade, and all workers will be unemployed. The same holds for GTBT_FIX, because the fixed wage $\overline{w} = 30$ is much too high to allow market exchange in equilibrium.

[15] The desired effort level is cheap talk. We, moreover, assume that workers that are indifferent accept the contract.

In GTBT and GTBT_RI firms will employ workers in periods of GT for the contract $\{w = 6, \tilde{e}\}$ (i.e., in periods $1-3$, $7-9$, and $13-15$) and not employ any workers in periods of BT (i.e., in periods $4-6$, and $10-12$).

**Hypothesis 1.1 (Employment)** *There will be full employment[16] in treatment FGT and in good times in treatments GTBT and GTBT_RI. There will be full unemployment in treatments FBT and GTBT_FIX as well as in bad times of treatments GTBT and GTBT_RI.*

**Hypothesis 1.2 (Wages, effort, efficiency)** *All workers – if employed – will be employed at $\{w = 6, \tilde{e}\}$ and will exert minimal effort levels. Efficiency is minimal.*

### Alternative Predictions: Non-Selfish Players

If decision makers are fair-minded (not entirely selfish) or if the fraction of fair-minded subjects on the market is sufficiently large, incomplete contracts in gift-exchange relations can be enforced through reciprocal behavior. One way to analyze the effects of the presence of fair-minded subjects is to assume that workers compare their profits with those of their employers. A tractable model that incorporates this idea is the model of inequity aversion by Fehr and Schmidt (1999). However, there is a plethora of (Bayesian) Nash equilibria when one applies the model to our setup. Rather than characterizing the entire set of equilibria, we just aim to illustrate in a stylized way how contractual incompleteness and our treatment variables interact in determining fairness contracts and unemployment, in the presence of fair-minded subjects.

More specifically, assume that a fraction $0 \leq \gamma \leq 1$ of workers has egalitarian preferences such that they are assumed to fulfill the contract offered to them, as long as they are offered at least the equal split of the surplus being generated by the transaction (see also Fehr et al. (2007)).[17] If they are offered less than the equal split, workers accept

[16] This is a slight abuse of the term. We disregard the two workers that are unemployed in our setting by design.

[17] We do not rule out that firms also exhibit fairness preferences, but since workers are always reacting to contracts offers by firms with their effort decisions, the workers' preferences over outcomes are crucial for the transaction. Adding fairness preferences of firms would complicate the analysis and provide little additional insight. Restricting our analysis to only two types of players might be a simplification that does not always seems to be warranted by the data (see Dittrich and Ziegelmayer (2010)), but it is useful to derive clear-cut predictions.

the contract but even fair-minded workers shirk by providing the minimum level of effort as their equity fairness norm had been violated. The remaining fraction $(1-\gamma)$ of workers is purely selfish and would shirk in a one-shot interaction. The utility function of a fair-minded worker (suppressing individual and time indices) can be expressed by

$$u(w,e,\tilde{e}) = \begin{cases} w - c(e) & \text{if } w - c(\tilde{e}) < \frac{1}{2}[p_{GT,1w}\tilde{e} - c(\tilde{e})] \\ w - c(e) - k\max[\tilde{e} - e; 0] & \text{if } w - c(\tilde{e}) \geq \frac{1}{2}[p_{GT,1w}\tilde{e} - c(\tilde{e})] \end{cases} \tag{1.1}$$

where the cost-of-effort function is $c(\cdot)$, $p_{GT,1w}$ indicates the productivity parameter of the firm in good times and $k$ is a fairness parameter which ensures that a fair-minded worker chooses the desired effort level $\tilde{e}$ as long as he obtains (at least) half of the net surplus from the contract. In order to guarantee this, we furthermore assume that $k > \max c'(e)$ such that $e = \tilde{e}$ is indeed optimal, since the fairness costs associated with choosing a lower level of effort than desired in the contract outweigh the material benefit from a lower effort level. Utility functions for bad times are constructed analogously.[18]

We assume further that firms are risk neutral and maximize their expected monetary payoffs. We solve the game by backward induction. In the last period of the experiment $T$, a firm can either offer a 'fair' ('trust') contract that equalizes payoffs between the firm and the worker characterized above or offer workers their outside option in the form of a contract according to the standard predictions (henceforth, the 'standard' contract). This contract has to account for the level of the unemployment benefit of 6, i.e., firms have to offer at least a wage of 6 in order to induce workers to accept.

Table 1.4 gives an overview of the predictions derived in appendix A1.1 for the last period of each of our five treatments. It contains the predicted contracts $[w,\tilde{e}]$ and the minimal proportion of fair-minded workers $\gamma^*$ to make sharing the surplus an optimal strategy for firms and workers in the different treatments and/or the GT- and BT-phases within the GTBT treatments. Since there is no last period with BT in the game-theoretic sense in GTBT and in GTBT_FIX (both end with GT; only GTBT_RI has final BT-periods in the game-theoretic sense because of the random ID draws every period), we put BT-

[18] More details can be found in appendix A1.1.

predictions for GTBT and for GTBT_FIX in curly brackets. They are irrelevant for deriving theoretical predictions for our setup.

Table 1.4: Theoretical Predictions

| | Flexible Wages | | Fixed Wages | |
|---|---|---|---|---|
| | Reputation | No Reputation (GTBT_RI) | Reputation (GTBT_FIX) | No Reputation |
| FGT | $[w, \tilde{e}] = [59, 10]$ | - | - | - |
| | if $\gamma^* \geq 0.59$ | - | - | - |
| GTBT | In GT: $[w, \tilde{e}] = [59, 10]$ | In GT: $[w, \tilde{e}] = [59, 10]$ | In GT: $[w, \tilde{e}] = [30, 5]$ | - |
| | if $\gamma^* \geq 0.59$ | if $\gamma^* \geq 0.59$ | if $\gamma^* \geq 0.50$ | - |
| | {In BT: $[w, \tilde{e}] = [34, 10]$ | In BT: $[w, \tilde{e}] = [34, 10]$ | {In BT: $[w, \tilde{e}] = [30, 9]$ | - |
| | if $\gamma^* \geq 0.64$} | if $\gamma^* \geq 0.64$ | if $\gamma^* \geq 0.63$} | - |
| FBT | $[w, \tilde{e}] = [34, 10]$ | - | - | - |
| | if $\gamma^* \geq 0.64$ | - | - | - |

Theoretical predictions for the final period with fair-minded workers

Generally speaking, if $\gamma \leq \gamma^*$, a firm is better off offering the 'standard' contract in the last period. Otherwise, in all but the final period, the firm will employ a policy of contingent contract renewal in treatments FGT, FBT, GTBT and GTBT_FIX, i.e., re-employ the worker if he exerted the desired effort level and dismiss him if he shirked. A fair worker will accept any offer that shares the surplus equally, and selfish workers will cooperate if future rents from cooperating exceed the gains from shirking. This is clearly the case, such that selfish workers prefer to cooperate in the pre-final period and reveal their type only in the last period when they shirk. The mechanism works analogously in good times and in bad times, but in bad times wages are lower.[19] Since the game finishes with three periods of high productivity, the incentives for the selfish types are unaffected by the presence of periods of low productivity. They fulfill the contract and shirk only in the last period. Optimal behavior in GTBT with flexible wages therefore can be characterized by the following two conditions depending on the value of $\gamma$:

- **Condition 1:** $\gamma > \gamma^*_{GT,1w}$: Firms offer the 'fair' contract $[w, \tilde{e}] = [59, 10]$ in GT and $[w, \tilde{e}] = [34, 10]$ in BT. The contract is accepted by all workers who exert the desired effort level in all but the final period. Workers earn $\pi^w_{GT} = 59 - 18 = 41$ in

[19] It is interesting to note that the relatively small difference between the cut-off levels of $\gamma$ between GT and BT is a consequence of the different outside options of firms. In GT employers can always offer $w = 6$, whereas in BT the outside option is employing nobody and earning nothing.

GT and $\pi^w_{BT} = 34 - 18 = 16$ in BT until the last period T in which fair-minded workers earn $\pi^w_T = 59 - 18 = 41$, but selfish workers shirk to obtain a payoff of $\pi^w_T = 59 - 0 = 59$. In all non-final periods, firms earn $\pi^f_{GT} = 10 \cdot 10 - 59 = 41$ in GT and $\pi^f_{BT} = 5 \cdot 10 - 34 = 16$ in BT. In the last period, firms earn $\pi^f_T = \gamma(41) + (1-\gamma)(-49) \geq \pi^f_T(6,1)\ \forall \gamma > \gamma^*_{GT,1w}$.

- **Condition 2:** $\gamma^*_{GT,1w} \geq \gamma$: Firms offer the 'standard' contract $[w, \tilde{e}] = [6, 1]$ in GT and do not make an offer in BT. The contract is accepted by all workers who exert the desired effort level of 1 in all periods. Workers earn $\pi^w_{GT} = 6 - 0 = 6$ in GT and the unemployment benefit $b = 6$ in BT. Firms earn $\pi^f_{GT} = 10 \cdot 1 - 6 = 4$ in GT and nothing in BT.

In the case of fixed wages (treatment GTBT_FIX) the mechanism works identically. Optimal contracts and cut-off levels are different, as can be seen in Table 1.4, but the qualitative predictions when non-selfish workers are present are analogous. Most importantly under the fixed wage, effort levels from fair-minded types are not maximal at $e = 10$ but only $e = 5$ in good times. BT induce an increase of the effort level up to $e = 9$.
In the absence of a reputation mechanism (treatment GTBT_RI) every period can be treated as the final period. Thus the contract offers are identical to those in GTBT, assuming the same level of $\gamma$ across treatments, but the average effort levels should be clearly lower because selfish workers will always shirk and only fair-minded workers provide effort. Moreover, there is a third condition that has to be added. Condition 1 now only holds for $\gamma > \gamma^*_{BT,1w}$, condition 2 remains unchanged. For intermediate levels of $\gamma$, condition 3 is relevant for treatment GTBT_RI.

- **Condition 3:** $\gamma^*_{BT,1w} > \gamma > \gamma^*_{GT,1w}$: Firms offer the 'fair' contract $[w, \tilde{e}] = [59, 10]$ in GT and do not make an offer in BT. In GT, the contract is accepted by all workers where only the fair-minded workers exert the desired effort level and the selfish workers shirk. Workers earn $\pi^w_{GT} = 59 - 18 = 41$ (fair-minded) or $\pi^w_{GT} = 59 - 0 = 59$ (selfish) in GT and the unemployment benefit $b = 6$ in BT. Firms earn $\pi^f = \gamma(41) + (1-\gamma)(-49) \geq \pi^f(6,1) \forall \gamma > \gamma^*_{GT,1w}$ in GT and nothing in BT.

Note that firms have to rely entirely on beliefs about $\gamma$ throughout the experiment in

GTBT_RI, when they enter the offer phase at the beginning of each period in GTBT_RI. The reputation mechanism in GTBT keeps the incentives for selfish workers not to shirk. As a consequence, the effects of productivity shocks could be more severe without a reputation mechanism.

**Hypothesis 1.3 (Employment)** *Regardless of the fraction of non-selfish workers, there will be full employment in treatment FGT and in good times in treatment GTBT as well as in treatment GTBT_RI. In the presence of fair-minded workers, either full unemployment or full employment can occur in treatments FBT and GTBT_FIX as well as during bad times of treatments GTBT and GTBT_RI, depending on the fraction of non-selfish workers in the population (with different cut-off requirements in different treatments).*

**Hypothesis 1.4 (Wages, Effort, Efficiency)** *In treatments FBT, FGT, GTBT and in GTBT_RI efficiency should either be maximal (with sufficiently many non-selfish workers) or minimal. In the presence of sufficiently many non-selfish workers, wages in treatments GTBT and GTBT_RI fluctuate over GT and BT, whereas effort levels will always be maximal. In contrast, in GTBT_FIX effort levels (and efficiency) will be higher in BT than in GT with enough non-selfish workers being present.*

It is important to add that the cut-off level requirements regarding the presence of non-selfish workers are very demanding. All calibration results from the literature (e.g., in Fehr and Schmidt (1999), but also in more recent contributions) indicate that the proportion of non-selfish decision makers is clearly below the requirements for observing trust contracts. However, we have only modeled an inequity aversion motive and not an efficiency motive (Charness and Rabin, 2002) for non-selfish workers. The latter is important in the context of the gift-exchange environment, but it makes predictions messier because it adds at least another degree of freedom for the analysis. When discussing our results, we will relate them to potential efficiency concerns of workers without resorting to a formal model.

Over and above the effects captured by assuming a fraction of fair-minded subjects in our setup, the switch between GT- and BT-phases might have additional effects on the development of reciprocity. The switches between regimes could hinder the emergence of reciprocal relationships and, thus, have a detrimental effect on the endogenous emergence of relational contracts. Although there is no clear theoretical guidance, it seems intuitive to expect an additional negative effect of the ups and downs in productivity on reciprocity. Whether it is significant is an empirical question that we are able to study by comparing the relevant periods in the GTBT treatments with treatments FGT and FBT.

## 1.4 Results

We begin by presenting results for the three treatments GTBT, FGT and FBT (section 1.4.1). Subsequently, our analysis for the two additional treatments GTBT_FIX and GTBT_RI is introduced in section 1.4.2. In section 1.4.3, we provide evidence on micro-mechanisms behind the stability of gift-exchange in the presence of productivity shocks.

### 1.4.1 Overview of Results of the Three Main Treatments

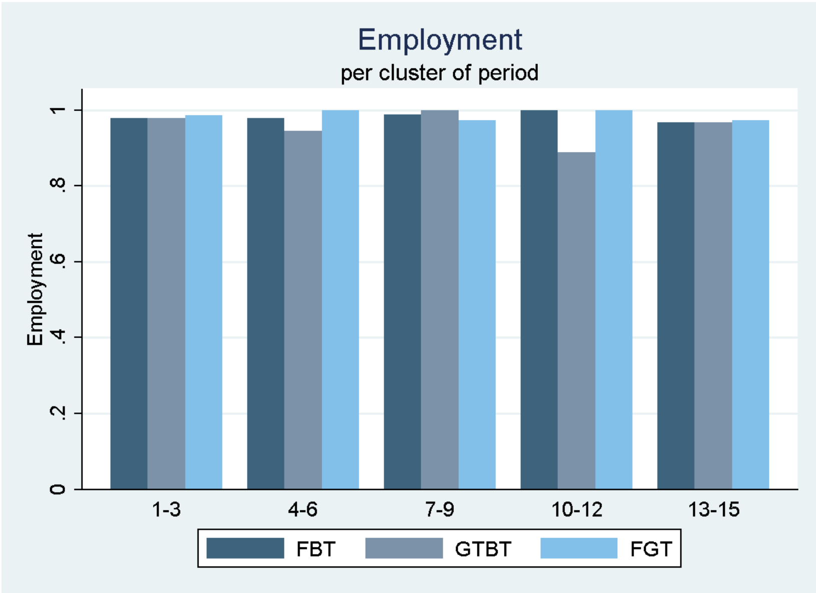
This section presents an overview of results regarding employment, wages, and effort levels and continues with reporting our results regarding the effects of productivity shocks on the contract nature.

#### Employment, Wages, Effort, and Reciprocity in Good and Bad Times

Figure 1.1 shows the percentage of concluded contracts per three periods in order to make treatments FGT and FBT comparable to treatment GTBT. Unemployment in excess of the natural rate (i.e., the excess supply of two workers by construction of our design) is almost non-existent. On average excess unemployment is highest in GTBT, but the rate of 4.4% is far below the 40% predicted by making standard assumptions. Even in the second negative productivity phase (i.e., in periods $10 - 12$) in GTBT, 80 out of 90 possible employment contracts are still realized. Hypothesis 1.1 is clearly refuted.

**Result 1.1** *In accordance with the model taking non-selfish worker behavior into account (see Hypothesis 1.3) there are practically no effects on employment levels from introducing market-wide fluctuation in economic productivity. Non-selfish concerns of workers and/or the available reputation mechanism are able to overcome negative effects of BT.*

Figure 1.1: Employment Levels in FBT, GTBT, FGT



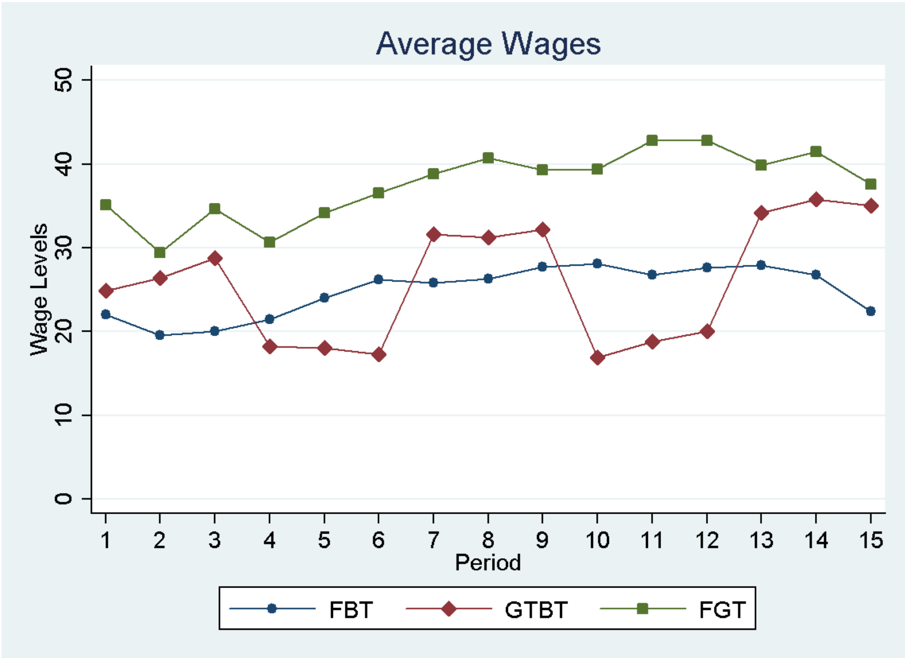

Next we look at wages and efforts. Figure 1.2 displays the average levels of wages for all three treatments. Wages are, not surprisingly, significantly higher in treatment FGT than in treatment FBT (Mann-Whitney-U-test, $p = 0.02$)[20]. We observe market averages between 22 and 27 in FBT and between 27 and 49 in FGT. In both treatments they slightly increase over time if we disregard the endgame effect, and they are somewhat below the alternative predictions based on our stylized inequity aversion model.

Treatment GTBT follows a clear cyclical pattern. In GT of GTBT, wages are significantly higher than in BT (Wilcoxon-signed-ranks-test between average wages from all GT- and all BT-periods, $p = 0.03$).[21] Wage decreases in BT compared to GT are considerable. They amount to $-33\%$ in the first BT-phase and $-42\%$ in the second BT-phase. It is noteworthy that wage levels in the GT- and BT-phases in GTBT are always below their corresponding levels in the treatments FGT and FBT. This is not predicted by any of

[20] In the following, all statistical tests use market averages as the unit of observations unless otherwise indicated. All reported $p$-values are two-sided.

[21] If we take the differences in average wages between single GT- and BT-phases we obtain the same results (Wilcoxon-signed-ranks-test, $p < 0.05$, for all six possible comparisons). Furthermore, we examine the changes in contracted wage levels for the periods adjacent to the switch from GT to BT and from BT to GT. Again the differences are significant (Wilcoxon-signed-ranks-tests on average per-period wages on the market level, $p = 0.03$ for all examined cases (periods 3 to 4, periods 6 to 7, periods 9 to 10, and periods 12 to 13)), i.e., wages decline significantly in all markets when the economy enters BT and increase if it goes to GT.

Figure 1.2: Average Wages in FGT, FBT, and GTBT.

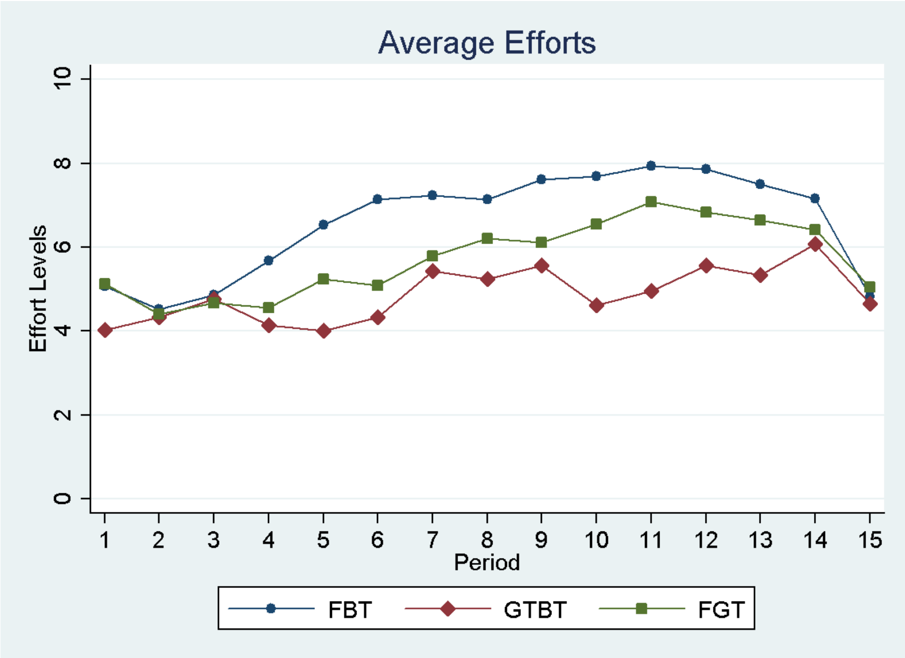



the models, and it hints at an additional effect of the instability on the labor market that cannot be captured by existing models.

**Result 1.2** *While there is virtually no effect of productivity shocks on employment levels, wages react strongly to the shocks. Wage levels drop significantly in BT, but recover in GT. Average wage levels are much higher than predicted by the standard model, but somewhat lower than our alternative predictions.*

Our next look is at effort levels. Workers' effort choice is neither contractible nor enforceable in any of our treatments, but remains observable by firms who can threaten to end the relationship by not offering a contract in the subsequent period. Figure 1.3 displays the average actual effort provided over time in the three treatments. The first observation is that effort levels are clearly higher in FBT than in FGT, even though average wage levels were lower in FBT than in FGT. Lacking previous results, it is noteworthy that the wage-effort relationship is obviously different for different productivity levels (Mann-Whitney-U-test, $p < 0.01$). Our main focus, however, is on the GTBT treatment. Although one can observe fluctuations in effort levels over the GT- and BT-phases, they

Figure 1.3: Average Effort Levels in FGT, FBT, and GTBT




are much smaller than for wages. While wages for the first regime switch from GT to BT fall by 33%, effort levels dip by a mere 5%. The same holds true for the second switch (−42% in wages versus −7% in effort levels). In line with the development of wages, there is a slight upward trend in effort levels in all treatments, as long as one disregards the last period. Again, the GTBT-treatment's effort levels are below the corresponding levels in the GT- and the BT-treatments, indicating an effect of the instability of productivity per se.[22] Effort level predictions are clearly above the predictions from Hypothesis 1.2 and clearly below the predictions from Hypothesis 1.4.[23]

**Result 1.3** *Despite the strong drop in wages when negative productivity shocks hit, average effort levels fall only slightly (and non-significantly). Average absolute levels are*

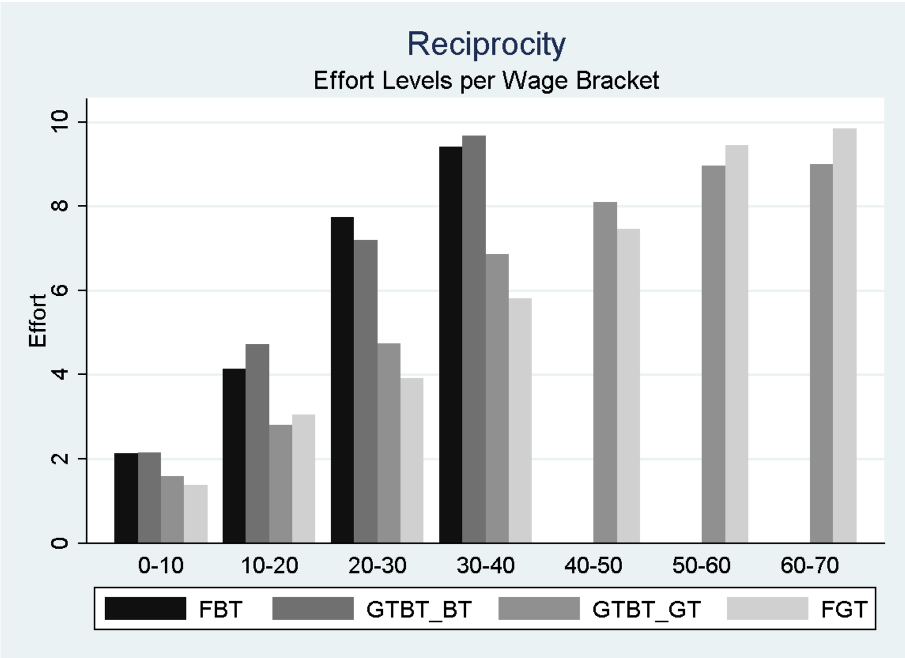
[22]More precisely, efforts in GT-phases in the GTBT-treatment are lower than the corresponding effort levels in the FGT-treatment, but not significantly so. The difference between the BT-phases effort levels in GTBT and the FBT effort levels is, however, highly significant (Mann-Whitney-U-test, $p = 0.01$).

[23]We also ask firms to state an "expected" level of effort before they learn worker's actual decisions, although in a non-incentivized manner. We find that expected levels are very close to the actual levels. In particular, they track the slight decline in effort levels at the regime switch between GT and BT periods. We take this as evidence that firms, to some extent, expect workers not to withdraw effort entirely when they pay them significantly lower wages in BT-periods. Desired effort levels from the contract offers are flat at about 2 points higher than actual levels with a decline when the markets enters the second BT phase in period 10; see Figure A1.4 in appendix A1.6.

*clearly above standard predictions but below the predictions from the inequity aversion model.*

Bringing wage levels and effort levels together allows us studying the level of reciprocity. We follow the convention by defining reciprocity as the slope of the wage-effort relationship.[24] Reciprocity levels in the three treatments and separately for GT and BT in GTBT for different wage classes are displayed in Figure 1.4. We distinguish between BT and GT in the GTBT treatment to see if there are differences in reciprocity between GT and BT within GTBT.[25]

Figure 1.4: Reciprocity in FGT, FBT, and GTBT




We find a positive relationship between the accepted wage and exerted effort levels in all three treatments. Moreover, we observe that the level of reciprocity differs strongly between FGT and FBT. Close to maximal effort levels prevail in the wage bracket between 30 and 40 in FBT, whereas in FGT firms can only expect similar effort levels at

[24]We acknowledge that positive reciprocity can also be defined differently; for instance, by the voluntary effort over and above the equal-split effort or the desired effort. Since we are not interested in the level of reciprocity but only in the treatment differences in reciprocity in this chapter, the choice of definition does not really affect our conclusions regarding treatment differences.

[25]We exclude wage levels above 40 (in BT) and above 70 (in GT) due to very small number of observations in these categories.

wages above 60. Also within GTBT, the levels of reciprocity differ between BT and GT. Interestingly, the levels of reciprocity in BT and GT in treatment GTBT are close to their counterparts from the FGT and the FBT treatment, respectively.

Table 1.5: Panel Regressions on Effort

| **Dep. Var.** | FBT and FGT, RE | | All treatments, RE | | All treatments, FE | |
|---|---|---|---|---|---|---|
| effort | (1) | (2) | (3) | (4) | (5) | (6) |
| GTBT-dummy | | | 0.897*** | 0.243 | | |
| | | | (0.285) | (0.246) | | |
| FBT-dummy | 0.338 | -0.375 | -0.337 | -0.393 | | |
| | (0.492) | (0.330) | (0.317) | (0.314) | | |
| Wage | | 0.125*** | 0.124*** | 0.125*** | 0.116*** | 0.116*** |
| | | (0.008) | (0.007) | (0.007) | (0.011) | (0.011) |
| Private | 1.873*** | 0.821*** | 0.807*** | 0.843*** | 0.890*** | 0.924*** |
| | (0.264) | (0.193) | (0.178) | (0.174) | (0.243) | (0.239) |
| Tenure | 0.379*** | 0.137*** | 0.158*** | 0.147*** | 0.125*** | 0.116*** |
| | (0.048) | (0.042) | (0.033) | (0.033) | (0.032) | (0.031) |
| Wage*FBT | | 0.099*** | 0.098*** | 0.099*** | 0.093*** | 0.095*** |
| | | (0.011) | (0.011) | (0.011) | (0.018) | (0.018) |
| Wage*GTBT | | | -0.008 | 0.004 | -0.012 | 0.002 |
| | | | (0.011) | (0.008) | (0.014) | (0.011) |
| Private*FBT | | 0.168 | 0.147 | 0.180 | 0.108 | 0.135 |
| | | (0.219) | (0.209) | (0.204) | (0.267) | (0.263) |
| Private*GTBT | | | 0.197 | -0.094 | -0.021 | -0.104 |
| | | | (0.234) | (0.223) | (0.341) | (0.313) |
| BT-dummy | | | | 1.015*** | | 0.851*** |
| × GTBT-dummy | | | | (0.228) | | (0.219) |
| Period | 0.417*** | 0.233*** | 0.261*** | 0.196*** | 0.272*** | 0.218*** |
| | (0.153) | (0.081) | (0.067) | (0.062) | (0.071) | (0.067) |
| Period$^2$ | -0.031*** | -0.017*** | -0.019*** | -0.015*** | -0.019*** | -0.016*** |
| | (0.009) | (0.005) | (0.004) | (0.004) | (0.004) | (0.004) |
| Constant | 2.618*** | -0.379 | -0.425 | -0.242 | 0.066 | -0.039 |
| | (0.455) | (0.358) | (0.316) | (0.314) | (0.164) | (0.174) |
| N | 812 | 812 | 1242 | 1242 | 1242 | 1242 |
| (Pseudo)$R^2$ | 0.41 | 0.72 | 0.72 | 0.73 | 0.68 | 0.71 |

Panel random effects (RE) and fixed effects (FE) regressions on the level of effort. Standard errors are given in brackets, clustering on the market level. *** indicates significance at the 1 % level, ** at the 5 %, and * at the 10 % level, respectively.

In order to control for covariates, we also run simple panel random effects (RE) and fixed effects (FE) estimations on effort levels, including treatment dummies, the wage, a dummy for a private offer, the tenure of the worker with that particular firm and a (quadratic) time trend.[26] Table 1.5 gives the results of the estimations. Models (4) and (6) add a

[26]The tenure is defined as the number of consecutive periods that a firm-worker couple has stayed together including the current period. One can argue that the tenure of the relationship is not an

dummy for the BT-periods ($4-6$ and $10-12$) in the GTBT treatment.

Our results indicate a robust reciprocal relationship between firms and workers, i.e., the wage is highly significantly related with effort, regardless of the treatment and control variables. A one-point increase in the wage, leads to an effort increase between 0.1 and 0.2 according to the treatment. Moreover, the tenure at a specific firm has a significantly positive effect on effort exerted, and contracts concluded from a private offer yield an effort level that is significantly higher by roughly one point than from public offers. With the FGT-treatment being our control group, the induced effort level from one additional wage increment increases highly significantly by 0.1 point in the FBT treatment, which leads to a higher effort level of about 2 points controlling for the wage. This result is hard to reconcile with any outcome-based fairness model that takes only efficiency concerns into account, because it is more efficient to exert effort in GT than in BT. The interplay of efficiency concerns and a compassion for the least-well off in the reference group (the firm in our case) – as in the model by Charness and Rabin (2002) – is, however, in principle consistent with our findings. The difference between GTBT and FGT is negligible if we include a BT-dummy for the GTBT treatment. Hence, reciprocity in GT in GTBT and in FGT is not significantly different, controlling for everything else. However, the highly significant coefficient on the BT-dummy captures the new wage-effort relationship in BT. Workers exert more effort in BT ceteris paribus when it is actually less efficient to do so. Reciprocity in BT in GTBT is clearly strongest, and it does not seem to be compromised by the fluctuations in GTBT. If at all, it is rather reinforced. We do not find contracts based on private offers to have a different effect on efforts across treatments. Fixed effect estimations in models (5) and (6) yield quantitatively and qualitatively similar results.[27]

**Result 1.4** *The wage-effort relationship changes significantly over periods of GT and BT. This is in contrast to standard predictions, but also consistent with our predictions based on the assumption that workers are inequity averse. Negative productivity shocks result in steeper wage-effort relationships. The level of reciprocity is reinforced in BT.*

exogenous variable. Therefore, it is important to note that dropping the variable tenure does not change any of our results reported in the following.

[27]Since 429 out of 1242 contracts in our data lead to either minimal ($e = 1$) or maximal ($e = 10$) effort levels, we have also run tobit panel regressions. They corroborate our conclusions, and estimation results can be found in Table A1.4 in appendix A1.5.

Wages and effort levels directly lead to employers' and workers' profits which are presented in Figure A1.6 in appendix A1.6. In FBT and FGT, workers secure themselves an almost constant average income of 15 points (FBT) and 30 points (FGT) per period. In GTBT, profits decrease for both employers and workers in BT periods, however to a lesser extent for workers than for employers. The latter earn, on average, close to zero profits in BT-periods. In appendix A1.2, we provide evidence that efficiency levels are lowest in GTBT and surplus sharing is similar in the three treatments at about 40 % for firms and 60 % for workers. Both outcomes can readily be inferred from wages and effort levels.

**Contract Nature**

The characteristics of the relationships between workers and firms that emerge endogenously illustrate the importance of relational contracting. The possibility of directing offers at particular workers enables firms to build up long-lasting profitable bilateral relationships. Under constant conditions it has been shown by Brown et al. (2004) that employers greatly care about the identification and thus reputation of their worker. Under varying conditions, long-term contracts are also possible, but supposed to be more difficult to establish if firms and workers cannot implicitly agree on a wage-effort relationship over GT- and BT-cycles. Indeed, the average contract length in our GTBT treatment amounts to 1.6 and in the two control treatments to 2.1 (FGT) and 2.3 (FBT), respectively.

The difference between GTBT and either control treatment in contract length is statistically significant (Mann-Whitney-U-test: $p = 0.03$ (GTBT vs. FBT), $p = 0.07$ (GTBT vs. FGT)). Comparing all three treatments, we obtain a similar picture from a Kruskal-Wallis test ($p = 0.05$). These results indicate that variations in economic conditions hamper the formation of long term contracts. A further confirmation for the hypothesis that variations in economic conditions have a negative impact on long-term contracting comes from a closer look at the overall length of a relationship in the three treatments, where Figure A1.2 in appendix A1.3 shows the cumulative frequency of the length of all concluded relationships. In the GTBT treatment 46% of all relationships lasted only one period compared to 34% and 31% in the two control treatments, respectively. Not surprisingly and already documented by Brown et al. (2004), concluded contracts are more likely to origin from a private wage offer in the course of the experiment in all three treatments.

This is in line with a bilateralization of trade on competitive markets in the form of relational contracts which is robust under variations of the economic conditions. Table 1.6 displays the percentage of contracts in clusters of three periods that were concluded on the basis of a private wage offer.[28] Especially the BT-periods in GTBT lead to a more

Table 1.6: Percentage of Contracts from Private Wage Offers

| | **FBT** | **GTBT** | **FGT** |
|---|---|---|---|
| Period 1-3 | 0.43 | 0.24 | 0.28 |
| Period 4-6 | 0.86 | 0.36 | 0.45 |
| Period 7-9 | 0.93 | 0.64 | 0.63 |
| Period 10-12 | 0.91 | 0.63 | 0.77 |
| Period 13-15 | 0.90 | 0.71 | 0.84 |

frequent use of public offers. In periods 4 and 10 (i.e., the first BT-periods in GTBT), there is a significant difference in the proportion of privately concluded contracts between the treatments (Fisher exact test and $\chi^2$-test, $p < 0.013$).[29] It is interesting to note that bilateralization is highest in treatment FBT (the difference between FGT and FBT as well as between GTBT and FBT is highly significant (Mann-Whitney-U-tests, $p = 0.01$ for both comparisons)), i.e., the fact that only small surpluses can be shared seems to reinforce the tendency of firms to make private offers. In appendix A1.3 we show that the renewal probability of contracts is lowest in GTBT compared to the stable FGT and FBT treatments. More precisely in periods 4 and 10 of GTBT, the proportion of renewed contracts is only 21 % and 39 %, whereas in the stable treatments this amounts to 42 % and 70 % (FBT) as well as 40 % and 64 % (FGT).

**Result 1.5** *Average contract length is lowest in GTBT. Firms increasingly use private offers across all treatments, but even if contract renewal is positively correlated to previous efforts in all treatments, this relationship is weakest in GTBT and particularly so in the BT-periods.*

[28] In appendix A1.6 (Figure A1.5), we report the precise development over the 15 periods. We also split up our three treatments into two parts (periods 1-7 and 8-15) and check for the percentage of concluded contracts on the basis of a private offer to find a higher share of private contracts in the second half of the experiment in all but one market. This increase is also confirmed to be statistically significant in all three treatments (Wilcoxon-signed-rank-tests, $p = 0.04$ (FBT), $p = 0.03$ (GTBT) and $p = 0.04$ (FGT)).

[29] In period 4, we have 82 % (FBT), 31 % (GTBT) and 40% (FGT) and in period 10 93% (FBT), 61% (GTBT) and 72% (FGT) of all contracts concluded following a private offer.

### 1.4.2 Two Control Treatments: No Reputation and Rigid Wages

Productivity shocks have a small negative effect on labor markets with incomplete contracts, but compared to theoretical predictions, reciprocity is very stable, and hence employment levels are almost maximal even in BT. Two straightforward arguments can be brought forward for the functioning of the markets on these surprisingly high levels: (i) unemployment is almost absent because of wage flexibility; and (ii) unemployment is almost absent because of the possibility to build reputation. Both features of our labor market so far – reputation formation and wage flexibility – have important implications for labor market outcomes that are supposed to interact with economic instability/fluctuations as captured by our treatment GTBT. Therefore, we implement two further control treatments: a treatment in which reputation formation is impossible because IDs are assigned each period randomly, and this is common knowledge among workers and firms (denoted GTBT_RI) and a treatment in which wages are completely rigid at $\overline{w} = 30$ (denoted GTBT_FIX).

Additional details regarding the design and theoretical predictions for GTBT_RI and GTBT_FIX have been discussed in section 1.3. The chosen level of the fixed wage at $\overline{w} = 30$ in GTBT_FIX is even higher than the overall average wage in GTBT, which amounts to 26.1, with 31.1 in GT and 18.2 in BT-periods. Everything else was kept exactly identical to treatment GTBT (including the outside option of workers and the possibility to set desired effort levels in GTBT_FIX). In the following, we will only compare GTBT with GTBT_RI and GTBT_FIX and disregard treatments FGT and FBT in oder to allow for a more succinct comparison.[30]

In our main condition GTBT, engaging in a long-term relationship was a possibility for both parties to retain the relational contract even in BT by relying on the past performance of the trading partner. More precisely, if a firm contracted repeatedly in GT with a worker who provided high levels of effort, it could reasonably infer that the worker will also abstain from shirking in BT through combining efficiency wages with the threat of firing. In GTBT_RI, firms can still pay efficiency wages but are deprived of the firing

[30] We conducted sessions with six markets of GTBT_FIX and with eight markets of GTBT_RI. The size of these markets were identical to the size of the markets in GTBT; the experimental instructions were only marginally adapted from the original GTBT-instructions in order to account for the different institutional setting. They are available from the authors upon request.

threat since they cannot identify workers. In a similar manner, workers cannot signal their willingness to provide high levels of effort in the future. So in the absence of reputation mechanisms and thus the ability to form relational contracts, we expect outcomes to be closer to the standard theoretical predictions from section 1.3.3. In GTBT_FIX, firms can use the firing threat to discipline workers, but are deprived of the signaling content of their wage offer. However, the fact that a worker receives an offer can already be interpreted as a kind signal. Hence, the comparison between GTBT_FIX and GTBT allows us to disentangle the value of a kindness signal through wage offers and through an offer in itself.

Figure 1.5: Employment Levels in GTBT, GTBT_RI, and GTBT_FIX

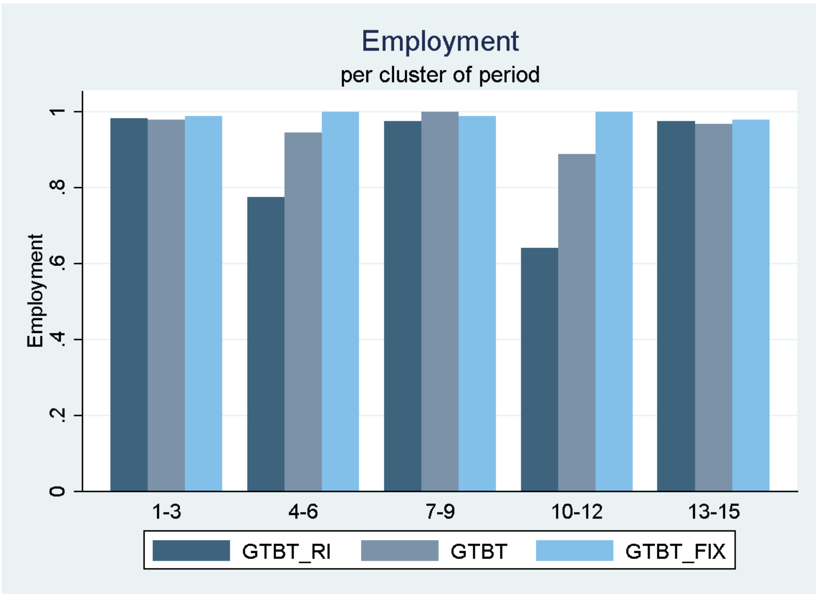



Figure 1.5 displays the employment level measured as the percentage of firms with a contract in clusters of periods. Fixing the wage (GTBT_FIX condition) at a level of $\overline{w} = 30$ has surprisingly no detrimental effect on employment since we observe close to full employment across BT and GT periods. Indeed there are 446 out of 450 possible contracts that have been concluded across all six markets. In GTBT_RI, however, we document a significant drop in employment in BT periods. Employment declines from close to full employment in GT to levels of 78% and 64% in BT periods.[31] Without repu-

[31] For GTBT_RI, Wilcoxon-signed-rank-tests with the number of contracts in clusters of periods as employment criterion give p-values of 0.01, 0.02, 0.01 and 0.01 for the transitions between BT and GT. The same test for treatment GTBT yields p-values clearly above 0.05.

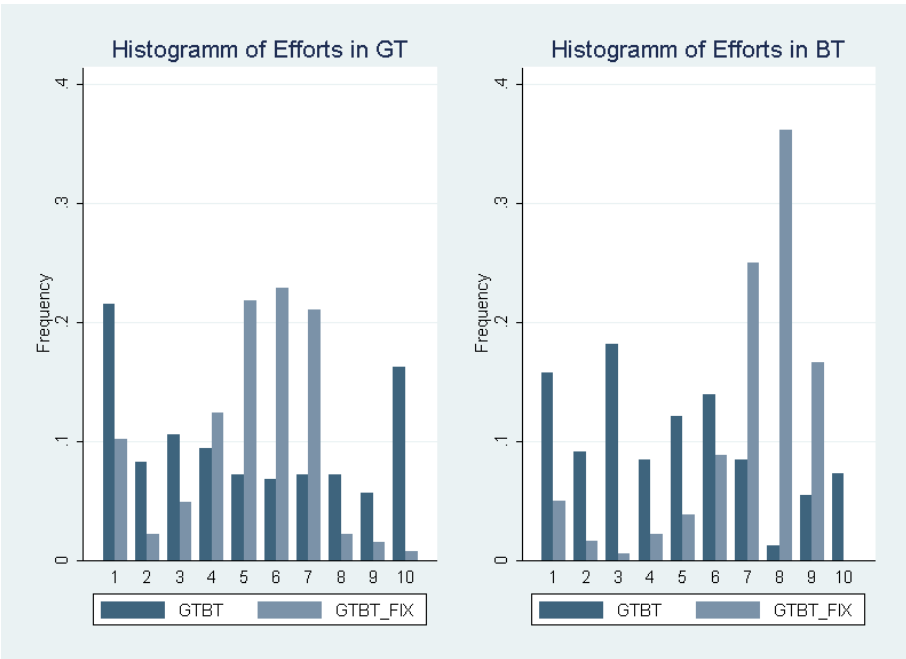
tation in GTBT_RI, productivity shocks have the expected effects on wages and efforts. We observe lower wages in GTBT_RI than in our main condition GTBT, as can be seen from the left panel of Figure 1.6. Firms are still offering higher wages in GT-periods than

Figure 1.6: Average Wages and Efforts in GTBT, GTBT_RI and GTBT_FIX


in BT-periods. Since they have no information about the inclination towards reciprocity of their worker in GTBT_RI anymore, firms especially reduce the number of risky high wage offers in GT-periods. Overall wage levels are consistently below the levels in GTBT; particularly in GT-periods the wage gap between GTBT and GTBT_RI amounts to more than 10 points (Mann-Whitney-U-test; $p = 0.05$). Effort levels in GTBT_RI overall are significantly lower than in GTBT (Mann-Whitney-U-test; $p = 0.02$,) and display no unraveling of types at the end of the experiment through shirking workers. More interestingly, average effort levels are virtually flat over BT- and GT-periods. As already mentioned, in the absence of reputation mechanisms workers have no incentives to invest in the relationship by providing high levels of effort, and firms are less willing to pay efficiency wages to motivate workers. Hence, the only rational reason why a worker would provide more than minimal efforts are other-regarding preferences like in Fehr and Schmidt (1999), indicated by effort levels that remain clearly above the minimum level throughout the whole experiment.

In treatment GTBT_FIX, effort levels match the predictions from our stylized model of inequity aversion with highly significant increases in the BT periods $4-6$ and $10-12$ (Wilcoxon-signed-ranks-test, $p = 0.03$). Figure 1.7 shows the density of efforts in all BT- and GT-periods separately, i.e., periods $1-3$, $7-9$ and $13-15$, for GTBT and GTBT_FIX. Comparing mean level of efforts in GT, we find no significant difference between the two treatments (5.0 against 5.1 overall effort average, Mann-Whitney-U-test, $p = 0.75$). In contrast, looking at the density functions we clearly reject the hypothesis of identical distributions (Kolmogoroff-Smirnov test, $p < 0.01$). Under fixed wages close

Figure 1.7: Frequency of Effort in all GT and BT Periods in GTBT and GTBT_FIX




to 80% of the effort levels are between 4 and 7, and we observe only a couple of outliers. Workers do not choose very high effort levels, as the fixed compensation is too low to induce maximum levels of effort. Moreover, the future value of employment in GTBT_FIX can be higher than in GTBT, as receiving a wage of 30 in BT is very attractive. Therefore, on average only 10% of workers shirk in GT in GTBT_FIX, compared to more than 20% in GTBT. In BT under flexible wages, a close to uniform distribution emerges where workers provide 4.6 in average efforts which is only slightly below their GT effort levels. When we fix wages at 30, however, effort levels increase to an average of 7.1 such that the difference between the two treatments is largely significant. Close to 80% of the workers now provide effort levels between 7 and 9. We again reject the hypothesis that the distributions are the same (Kolmogoroff-Smirnov test; $p < 0.01$).

The general shift to high levels of effort in BT in treatment GTBT_FIX is a consequence of the generous wage level of 30, compared to an average wage of slightly above 18 in BT in treatment GTBT. Fixed wages at a level which is clearly above the wage that emerges naturally in a comparable situation (in GTBT) are able to stabilize relationships and to overcome some of the problems associated with fluctuations. Hence, wage rigidity does not necessarily have to have negative consequences. When we look at profits in BT periods, we find that firms' profits are not significantly different between GTBT and GTBT_FIX in BT periods (Mann-Whitney-U-test on BT market averages, $p = 0.87$), but workers succeed in earning substantially more under fixed wages (Mann-Whitney-U-test, $p = 0.004$). In other words, in our experiment fixing the wage is efficiency-enhancing.

Our descriptive overview of the results is confirmed in a series of panel regressions on the effort level in Table 1.7 that take only the treatments with productivity shocks into account. Wage, contracts based on private offers, and the tenure of a relationship increase effort levels in all treatments, as expected, with one exception: the important effect of the reputation mechanism is outlined by the negative coefficient on a private offer in GTBT_RI across all specifications. Both coefficients jointly are not significantly different from zero ($p = 0.29$ and $p = 0.39$ in regressions (7) and (8)), indicating that the positive effect from private offers is eradicated by the impossibility to build up a relational contract. The increase in provided effort levels in the BT-periods, everything else equal, is significantly higher in treatment GTBT_FIX than in treatment GTBT. Note that with the inclusion of the interaction terms, the treatment dummies become insignificant, i.e., the positive (negative) effect on efforts in the three treatments is entirely captured by the BT-dummy and the interaction with the respective treatments. Comparing GTBT and GTBT_RI in BT-periods shows that the impact of the wage on effort in GTBT_RI is not lower than in GTBT when controlling for the reputation effect. So even in the absence of reputation mechanisms, subjects show concern for the situation of their employer by exerting higher efforts for the same wage in BT-periods. The fixed effect regressions (9) and (10) confirm the above findings.

Table 1.7: Panel Regressions on Effort in GTBT, GTBT_FIX, and GTBT_RI

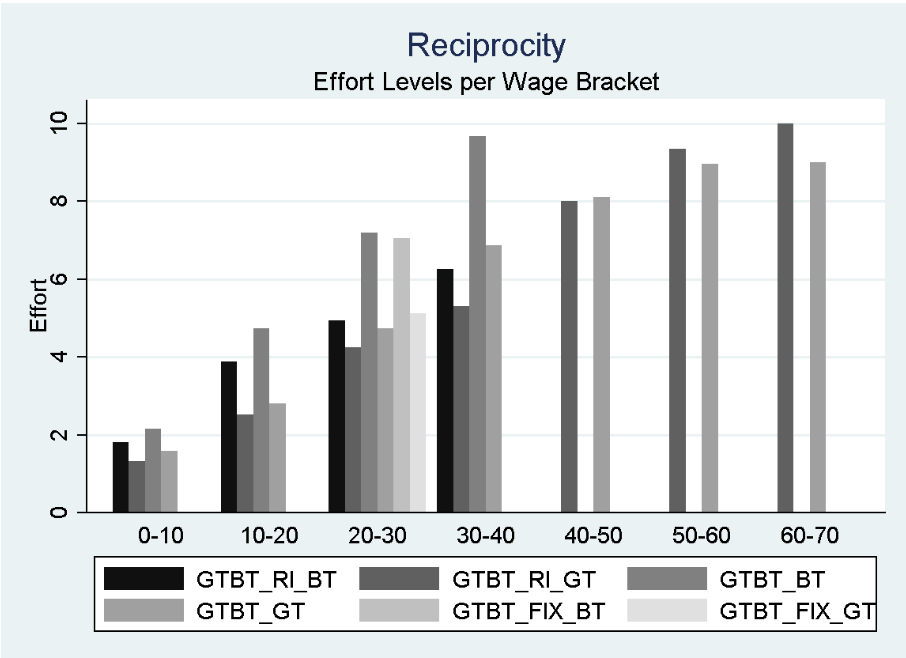
| **Dep. Var.** | Random Effects | | Fixed Effects | |
|---|---|---|---|---|
| effort | (7) | (8) | (9) | (10) |
| GTBT_FIX-dummy | 0.522* | 0.192 | | |
| | (0.296) | (0.285) | | |
| GTBT_RI-dummy | -0.668** | -0.209 | | |
| | (0.275) | (0.162) | | |
| Wage | 0.121*** | 0.130*** | 0.104*** | 0.116*** |
| | (0.008) | (0.005) | (0.006) | (0.007) |
| Private | 0.878*** | 0.856*** | 0.693*** | 0.717*** |
| | (0.138) | (0.147) | (0.223) | (0.212) |
| Tenure | 0.163*** | 0.152*** | 0.159*** | 0.151*** |
| | (0.037) | (0.032) | (0.034) | (0.033) |
| Wage*GTBT_RI | 0.021* | 0.012 | 0.029*** | 0.018* |
| | (0.012) | (0.011) | (0.009) | (0.009) |
| Private*GTBT_FIX | -0.245 | -0.230 | -0.250 | -0.280 |
| | (0.240) | (0.203) | (0.296) | (0.283) |
| Private*GTBT_RI | -0.713*** | -0.702*** | -0.583** | -0.607** |
| | (0.202) | (0.228) | (0.276) | (0.265) |
| BT-dummy | | 1.055*** | | 0.854*** |
| | | (0.221) | | (0.187) |
| BT-dummy | | 0.666** | | 0.873*** |
| × GTBT_FIX-dummy | | (0.298) | | (0.238) |
| BT-dummy | | -0.590 | | -0.474** |
| × GTBT_RI-dummy | | (0.396) | | (0.241) |
| Period | 0.402*** | 0.201*** | 0.395*** | 0.213*** |
| | (0.078) | (0.055) | (0.046) | (0.047) |
| Period$^2$ | -0.026*** | -0.014*** | -0.025*** | -0.014*** |
| | (0.005) | (0.003) | (0.003) | (0.003) |
| Constant | -0.134 | -0.190 | 0.213 | 0.143 |
| | (0.230) | (0.162) | (0.164) | (0.196) |
| N | 1398 | 1398 | 1398 | 1398 |
| (Pseudo)$R^2$ | 0.64 | 0.67 | 0.59 | 0.66 |

Panel random effects (RE) and fixed effects (FE) regressions on the level of effort in GTBT, GTBT_FIX, and GTBT_RI. Standard errors are given in brackets, clustering on the market level. *** indicates significance at the 1 % level, ** at the 5 % and * at the 10 % level, respectively.

**Result 1.6** *Rigid wages do not have a detrimental effect on employment. They significantly increase effort levels in BT-periods. In GTBT_RI, employment levels and effort levels decrease significantly compared to GTBT. The employment effect is exclusively driven by BT-periods.*

We next look at the level of reciprocity, again defined as the slope of the wage-effort relationship. Comparing GTBT and GTBT_FIX in the wage bracket 20-30, we find identical levels of "reciprocity" despite the fact that the level of the wage itself does not convey any intention to the worker (see Figure 1.8). There are three possible explanations for this finding: (i) the threat of unemployment after exerting low levels of effort drives effort levels up; (ii) social preferences are rather outcome-based than intention-based; if we, however, assume that at least part of other-regarding concerns are driven by intentions, then explanation (iii) stipulates that the mere fact of receiving an offer is interpreted as a kind act and reciprocated by workers. The threat of unemployment must be weak, given that workers choose effort levels which secure themselves the bigger share of the surplus. Nevertheless they still leave, on average, part of the surplus on the table for firms. Our results tentatively suggest that the behavior of the participants in the experiments can be rationalized rather by outcome-based preferences like Fehr and Schmidt (1999) over the split of the surplus than by traditional intention-based models acting purely between wage and effort levels within a period.

Figure 1.8: Reciprocity in GTBT, GTBT_RI and GTBT_FIX




From figure 1.8, reciprocity is found to be higher in GTBT than in GTBT_RI, especially in BT where the reputation mechanism ensures that workers exert higher amounts of effort for the same wage bracket. In GT-periods, GTBT and GTBT_RI do not display significant differences in the level of reciprocity. This yields further evidence for the particular importance reputation mechanisms play in BT-periods where they succeed in generating higher levels of reciprocity compared to GTBT_RI.

**Result 1.7** *Levels of reciprocity are identical in GTBT and GTBT_FIX. Especially in BT periods, the absence of the reputation mechanism reduces reciprocity in GTBT_RI compared to GTBT.*

Wages and effort levels directly lead to employers' and workers' profits which are presented in Table A1.3 in appendix A1.4. Not surprisingly, GTBT_FIX realizes the highest possible surplus. As already mentioned, it seems efficiency-maximizing in our environment with incomplete contracts and productivity instability to fix wages, at least within a certain range of wage levels. The absence of a reputation mechanism in GTBT_RI harms efficiency, particularly in BT-periods which we show in detail in appendix A1.4. The split of the surplus is almost identical between GTBT, GTBT_FIX and GTBT_RI over time.

Concerning the contract nature, we find that bilateralization of contracts is more pronounced in GTBT_FIX than in GTBT. 73% of all contracts have been concluded on the basis of a private offer in contrast to only 52% in GTBT, which is indicative of stronger relational contracts in GTBT_FIX also explaining high effort levels especially in BT.

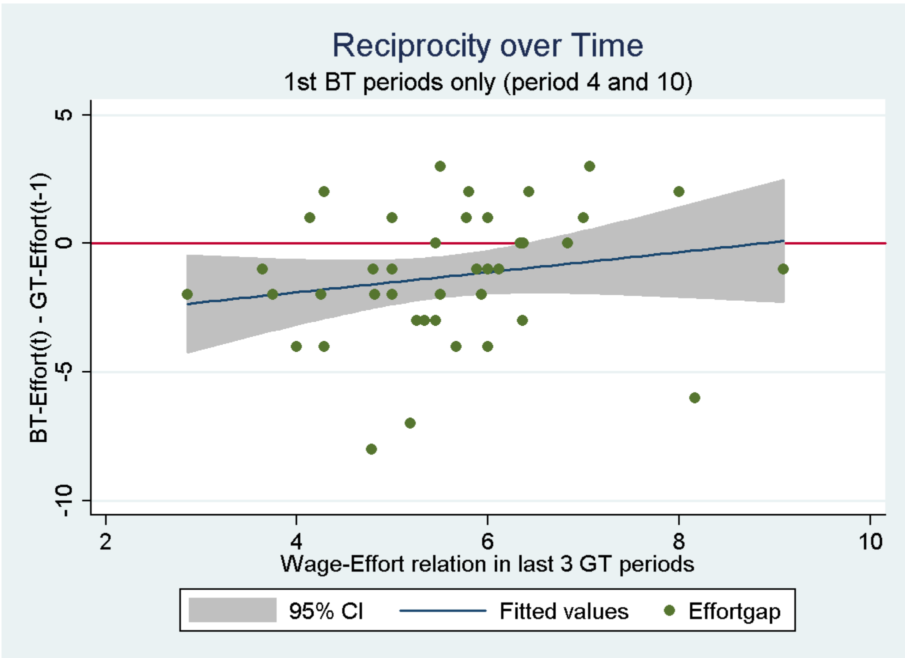
### 1.4.3 Driving Forces behind the Stability of Reciprocity with Productivity Shocks

In this section, we study the micro-mechanisms that explain the stability of employment that we observe on our experimental labor markets. At the heart of the fair employment hypothesis is the idea of dynamic reciprocity between firms and workers that goes beyond the static relationship between wages and effort within the same period. We first analyze outcomes focusing of worker behavior, before adapting the firm's point of view. From our analysis above, we know that workers provide significantly more effort, controlling for the wage in BT than in GT, on average. We further qualify this finding in this section by taking a closer look at worker and firm behavior in BT-periods and at the transition between GT- and BT-periods. The analysis is restricted to the GTBT treatment in a first step, before we add findings from GTBT_RI and GTBT_FIX.

### Micro-Mechanisms in Treatment GTBT

In GTBT, the focus of our analysis is on the changes in behavior between periods 3 and 4 as well as 9 and 10, i.e., between the last period of the GT-phase and the first period of three consecutive BT-periods. If the fair employment hypothesis holds, the extent to which a worker adapts his effort choice in the first BT-period compared to the last period of GT will depend on the contracting outcomes in the previous periods in GT. We therefore are interested in relating the effort gap of a worker between periods 3 and 4 as well as 9 and 10 to outcomes in periods 3 (respectively, 1-3) and 9 (respectively, 7-9). As a measure of outcome, we refer to the wage-effort relation (i.e., (average) wage

Figure 1.9: Effort Gap (BT-GT) and Past Wage-Effort Relation in GT for GTBT




over (average) effort) that prevailed in the previous period (on average, in the previous three periods) for the worker. In all six markets, we have 46 instances where a worker is employed in both periods adjacent to a shift from GT to BT. Figure 1.9 displays a scatterplot between these two variables which shows a positive relationship where the effort gap, displayed on the ordinate, is defined as the difference in effort between period

4 (10) and 3 (9) accordingly.[32]

The higher the wage a worker earned in the last three GT-periods for a given level of effort, the lower is the extent to which he cuts down his effort in the first BT-period if at all. Reciprocity hence acts in a dynamic manner between GT and BT, and it is able to sustain high effort levels in BT. We confirm this finding in a series of regressions (see Table 1.8) that explain the effort gap, where we also include the wage gap between BT- and GT-periods that naturally explains a substantial part of the effort gap. In all three specifications, we find the same significant relation between the past wage-effort relation and the effort gap when entering the BT-periods. As one would expect, tenure at the firm

Table 1.8: OLS Regression on Effort Gap between GT- and BT-periods in GTBT

| **Dep. Var.** $\Delta$ Effort: $e_t - e_{t-1}$ | (11) | (12) | (13) |
|---|---|---|---|
| $\Delta$ Wage: $w_t - w_{t-1}$ | 0.105*** | 0.110*** | 0.118*** |
| | (0.008) | (0.010) | (0.013) |
| Tenure | 0.232** | 0.294** | 0.251** |
| | (0.069) | (0.075) | (0.079) |
| Wage-effort relation in last three GT periods | | 0.144** | 0.136** |
| | | (0.049) | (0.049) |
| Proposed surplus to the worker | | | 0.304 |
| | | | (0.404) |
| Proposed surplus to the worker $_{t-1}$ | | | 1.066 |
| | | | (1.374) |
| Constant | -0.071 | -1.108** | -2.381** |
| | (0.346) | (0.421) | (0.442) |
| N | 46 | 46 | 46 |
| $R^2$ | 0.56 | 0.66 | 0.67 |

OLS regressions on the effort gap between GT- and BT-periods in GTBT within the same firm, i.e., periods 4 and 10. Standard errors clustered on the market level are given in brackets. *** indicates significance at the 1 % level, ** at the 5 % and * at the 10 % level, respectively.

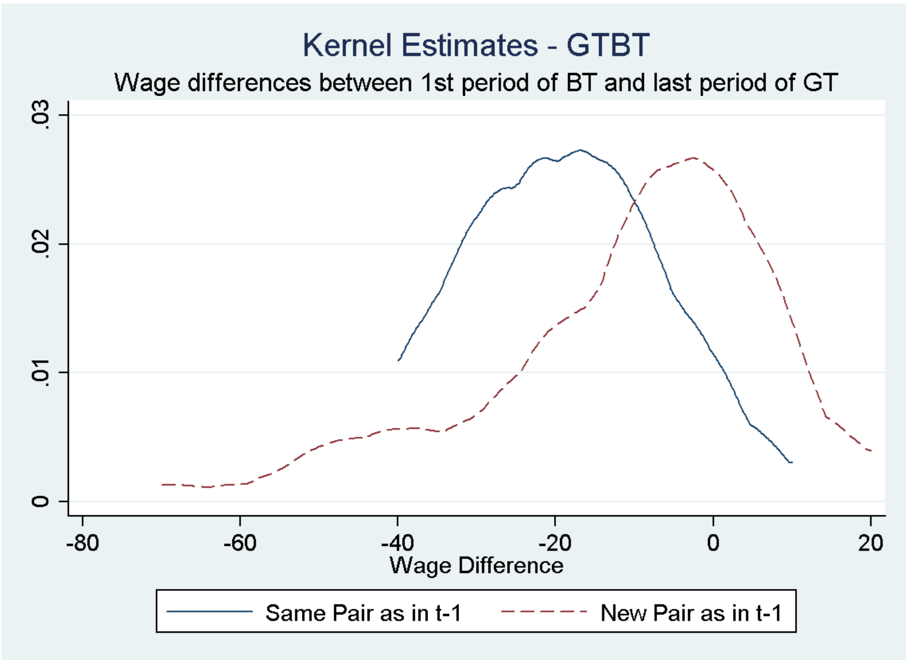
also has a positive impact, underlining the importance of relational contracts in coping with the negative productivity shock. Another possible driving force of effort choices could be the proposed surplus to the worker through the desired effort level. We compute the proposed share of the total surplus that accrues to the worker in the first BT- and the last GT-period on the basis of the desired effort level from the firm's contract offer that was accepted. It turns out, however, that the variable has no effect on the effort gap

[32]We drop four outliers ((wage-effort relation/effort gap): (40/1), (15/1), (15/2), (12/-2)) from the graph for reasons of succinct presentation, but naturally include them in the regressions. The results are unaffected by the inclusion or the exclusion.

between BT and GT.[33]

Pertaining to firm behavior, employers cut wages considerably when the market enters a BT-period. We take a closer look at the 57 instances in which a firm participated in the market in periods 3 and 4 as well as 9 and 10. We distinguish between those in which the firm-worker pairing is identical in the last GT-period and the first BT period (17 cases) and those in which firms employ a different worker in the periods adjacent to a regime-switch from GT to BT (40 cases). The estimated distribution of the wage gap between periods 3 and 4 as well as 9 and 10 separately for these two categories is given in Figure 1.10. We observe a clear difference between the distributions, where firms reduce the wage

Figure 1.10: Kernel Estimates for Distribution of the Wage Gap between BT- and GT-periods




by a higher amount in firm-worker pairs that stay together than otherwise. The difference between the two distributions is highly significant (Kolmogorov-Smirnov-test, $p = 0.02$). There is evidence that when firms interact with the same worker within the same relational contract, they can afford to pay lower wages than when they want or have to hire a new worker (from the market) for the first BT-period. Workers in an existing relationship

[33] We conduct the same analysis by comparing the wage gaps and effort gaps when taking the average values (wages, efforts, wage-effort relation and suggested surplus sharing) per worker across the entire three GT- and BT-periods, respectively. We obtain qualitatively and quantitatively very similar results.

hence accept lower wages than new workers who have to be motivated through stronger wage incentives, because for new matches of firms and workers, relational capital still has to be accumulated. Since the effort distributions do not differ significantly (Kolmogorov-Smirnov-test, $p = 0.19$) between relationships that persist over a change from GT to BT and those who do not, we observe a lower wage-effort relation in long-term firm-worker relationships in BT-periods. Firms in BT pay lower wages for the same levels of effort in relationships with the same interaction partner than with a different interaction partner. We interpret these results as evidence for an increased degree of wage flexibility *inside* the firm through relational contracts in the presence of productivity shocks. The labor market *outside* the firm for new hires is comparatively less affected when the economy goes from GT- to BT-periods.

In order to explain the driving forces behind the level of the wage-effort relationship that we observe in Figure 1.9, we construct a dummy variable that takes on the value one when the contract partners are identical in periods $t$ and $t + 1$ and the contract was concluded on the basis of a private wage offer.[34] When we consider all 15 periods in

Table 1.9: Panel Regressions on the Wage-Effort Relation in GTBT

| **Dep. Var.** Wage-effort relation | (14) | (15) |
|---|---|---|
| Same partner as $(t-1)$ & private offer | -1.076*** | -0.399 |
| | (0.197) | (0.365) |
| BT-periods | | -2.318** |
| | | (0.924) |
| Same partner as $(t-1)$ & private offer × BT-periods | | -1.540*** |
| | | (0.599) |
| Constant | 6.804*** | 7.676*** |
| | (0.334) | (0.625) |
| N | 430 | 430 |
| $R^2$ | 0.01 | 0.07 |

Random effects panel regressions on the wage-effort relation in every period in GTBT. Standard errors clustered on the market level are given in brackets. *** indicates significance at the 1 % level, ** at the 5 % and * at the 10 % level, respectively.

GTBT, we find a significantly negative impact by roughly one point on the wage-effort relation in renewed contracts. That is to say that in these interactions the "price" of effort that firms need to pay workers for a given level of effort is on average one point lower

[34] This is to exclude contracts where interaction partners met a second consecutive time over the market.

than in newly established firm-worker relationships. Adding a dummy for the BT-periods and an interaction term, we find that the effect is driven by the BT-periods in general and, even more so, strengthened in the contracts that are renewed with a private offer in BT. Relational contracts hence allow firms to ”buy“ effort at a lower price from the existing workforce than on the market through new hires in BT-periods. This lends further support to our general hypothesis that dynamic forms of reciprocity shape labor market outcomes in the presence of productivity shocks by acting as a crucial element to sustain comparatively high levels of efforts in BT-periods, particularly in existing firm-worker relationships.

**Result 1.8** *The stability of the labor market in BT-periods crucially depends on the presence of inter-temporal mechanisms of reciprocity in GTBT. Relationships that persist over the transition from GT- to BT-periods are characterized by a higher degree of flexibility than those concluded on the market.*

#### Effects in Treatments GTBT_RI and GTBT_FIX

As a next step, we want to analyze the effectiveness of the mechanisms outlined above when reputation mechanisms are absent, in GTBT_RI. Unlike in GTBT, we find no relation between the effort gap between GT- and BT-periods and the wage-effort relation in the last three GT-periods. The slope of the fitted line as in Figure 1.9 is not significantly different from zero. Furthermore, when we conduct the same regressions as in GTBT for the data from treatment GTBT_RI – these regressions can now be viewed as placebo regressions – we unsurprisingly find insignificant coefficients for the wage-effort relation in the past three GT-periods on the effort gap between GT and BT. So if reputation mechanisms are absent, the inter-temporal effect of reciprocity cannot show up. Obviously, in GTBT_RI firms cannot apply different wage schemes to existing workers and to new hires.

When wages are fixed in GTBT_FIX, we again cannot document any inter-temporal effects of the wage-effort relation in the three GT-periods on the effort gap of workers in BT-periods. Similar regressions as above for GTBT yield insignificant coefficients for the past wage-effort relation. In the absence of incentives through wage offers by firms, the predominant mechanism that guarantees high effort levels in BT-periods in GTBT_FIX almost exclusively arises from the motivations to share the surplus between firms and workers. The surplus sharing norm also governs behavior in GTBT, but is enriched by the dynamic elements of reciprocity outlined above. These mechanisms are absent in GTBT_FIX where firms only can act on employment and not on wages. The fact that there are other mechanisms at work in treatment GTBT_FIX is an indication that inter-temporal reciprocity is induced by firms through their contract offers and not by workers through effort signals in the last GT-period before the first BT-period. However, workers understand the restrictions for firms in treatment GTBT_FIX and increase effort levels nevertheless, which leads to the reported high levels of efficiency under fixed wages.

**Result 1.9** *The absence of reputation mechanisms in GTBT_RI makes inter-temporal relationships and incentives impossible. Under fixed wages (GTBT_FIX) the norm of surplus sharing overcomes the lack of signaling through the wage offer, and it promotes high levels of efficiency.*

### 1.4.4 Discussion

Our experiments revealed several remarkable features of markets with incomplete contracts that experience productivity shocks. First, markets are extremely stable in terms of employment despite the productivity shocks and despite the prediction of full unemployment based on standard assumptions. Wages of workers go down considerably, but effort levels fall only slightly, on average, when productivity plummets. There are small effects of productivity shocks on the functioning of the markets such as shorter average firm-worker relationships and slightly lower wages and effort levels than in the control conditions, but the effects are much less severe than predicted by the standard model.

Two main features of our design – entirely flexible wages and the possibility to build up reputation over time – could be the driving force behind the stability of the markets. Implementing two control treatments that systematically drop these features in the design allows us to conclude that it is definitely not the wage flexibility that is responsible for the functioning of the markets. Fixing the wage in the contract offer on a comparably high level does neither lead to higher levels of unemployment, nor does it create any other frictions. On the contrary, it leads to even higher levels of efficiency on the market through higher levels of effort provided than in our main treatment. Wage rigidity should, therefore, not be viewed as negative per se; on markets with incomplete contracts and instability of productivity or profits, it could even be beneficial for certain ranges of wage levels.

In contrast to wage rigidity, taking away the possibility to form relational contracts over time makes it harder for firms and workers to overcome the negative effects of productivity shocks on the markets. Even though unemployment is still by far lower than predicted by the standard model, it is significantly higher than in all other conditions that we have analyzed. The negative effect becomes particularly apparent in phases with low productivity. Furthermore, efficiency is lowest in the treatment without reputation formation.

One interesting finding from our analysis is that workers exert more effort in phases of low productivity than in phases of high productivity per unit of wage. Remarkably, effort levels are highest in our control treatment with constant low productivity. Such a behavior is completely in line with models of other-regarding preferences, but it is obviously not efficiency-maximizing. From an efficiency perspective it would be optimal to shift effort provision to high productivity phases. Apparently, relational contracts are not strong or stable enough to allow for such an intertemporal rationale.

Nevertheless, there are important intertemporal aspects in our data that are able to explain the well-functioning of the market with productivity shocks. First, relational capital that is accumulated through reciprocal behavior between firms and workers over time allows for a smoother transition from GT to BT. Indeed, firms are able to pay lower wages to workers in the first period of the BT-phase that have already been employed in GT. Hence, relational contracts seem to allow for a higher degree of wage flexibility within the firm compared to the market outside the firm. Second, while there is no rigorous test available because of the endogeneity inherent to the interaction, our result indicates that (i) the mere fact of receiving a (private) offer is already considered to be a kind signal that is reciprocated by many workers, and (ii) that it is not the workers who are initiating the sustaining of the relationship in the transition from GT to BT through signaling with high levels of effort, but the firms who keep on making attractive offers when entering BT. Third, we find supporting evidence for intertemporal reciprocity over regime switches from GT- to BT-periods by looking at the relationship between the wage-effort ratio and the effort response in the transition. Highly reciprocal relationships during GT display smaller decreases or even increases in effort levels from the last GT-period to the first BT-period. Hence, it is safe to conclude that workers take intertemporal considerations into account, but they usually react on contract offers and do not signal intentions in the last GT-period.

The experimental evidence that we obtained so far appears not fully consistent with the results by Bewley (1999) based on interviews of executives about their wage and employment policy during the 1991-recession. Bewley (1999) stresses the importance of ”morale“ to keep workers' propensity to exert costly effort high, especially during recessions. For this reason, he concludes that executives consider a wage cut to be a severe hazard to a worker's morale and thus refrain from cutting wages during a recession, but rather lay the least productive workers off ”to get the problem out of the door all at once“ when productivity of effort is low.

In our setup so far, we restricted firms to hire a maximum of one worker such that it is not possible to disentangle these two motives of firms from each other. If firms are able to hire more than one worker and when workers are differently productive, they obviously have two ways of reacting to a productivity shock: reducing wages of the existing workforce or cutting employment through lay-offs. In order to investigate Bewley's main hypothesis in our setup and to assess the wage policy of firms vis-á-vis its employment policy when hit by productivity shocks, we extend our setup to firms that can hire up to two workers.

## 1.5 Good Times and Bad Times in a Multi-Worker Firm

### 1.5.1 Experimental Design and Theoretical Predictions

The experimental design in this section strongly parallels the design from section 1.3 in many aspects. In our main treatment with multiple workers (denoted GTBT2 henceforth), firms and workers again interact over 15 periods. In each of the 15 trading phases firms can post binding public or private wage offers that can be accepted by workers who choose non-enforceable efforts after accepting an offer. As the major difference, firms are now allowed to hire up to two workers per period, but workers could still only accept at most one contract.[35] Each market consisted of 11 subjects, with 3 firms and 8 workers such that the competitive pressure with two workers in excess supply was identical to the one-worker setting introduced above. In all treatments, firms were exposed to the same cyclical pattern in productivity as in our GTBT treatment from above. The effort costs of workers were again given by Table 1.1 such that incentives for workers were unchanged.[36] As in Altmann et al. (2009), a firm's production technology is now characterized by decreasing returns to scale while being subject to the same productivity shocks in BT-periods that we introduced earlier. The material payoff of a firm during GT (i.e., periods 1-3, 7-9 and 13-15) is given by:

$$\pi_f = \begin{cases} 8 \cdot (e_1 + e_2) - w_1 - w_2 & \text{if two contracts were concluded} \\ 10 \cdot e_1 - w_1 & \text{if one contract was concluded} \\ 0 & \text{otherwise} \end{cases}$$

In BT-periods (i.e., periods 4-6 and 10-12), a firm's profit is given by:

$$\pi_f = \begin{cases} 4 \cdot (e_1 + e_2) - w_1 - w_2 & \text{if two contracts were concluded} \\ 5 \cdot e_1 - w_1 & \text{if one contract was concluded} \\ 0 & \text{otherwise} \end{cases}$$

Note that if a firm hires only one worker, her material payoff is identical to the situation in the one-worker setting. In addition to that, the firm is now able to hire a second worker at a lower marginal productivity than her first worker yields.[37] Full effort of $e = 10$ is

[35]Since firms could not withdraw wage offers, all outstanding offers from one firm were automatically deleted upon acceptance of one contract by a worker (see Altmann et al. (2009)). If a firm wanted to hire a second worker, she could again post new offers that could be then accepted by all remaining workers that were without contract at this time. We implemented this feature to guarantee that if employers wanted to employ only one worker but entered multiple offers for the first hire, they did not end up with two workers unintentionally.

[36]Remember that workers obtain an unemployment benefit of 6 if they are not employed in a given period. This feature of the market is also retained.

[37]It is important to underline that in our specification of firm profits only the effort sum is relevant. It

always efficient with this specification, regardless of the condition of the economy since the second worker's marginal productivity per unit of effort lies at 3 in BT and thus never exceeds his marginal cost of effort that lies between 1 and 3. All parameters were made common knowledge in the instructions for our experimental participants. After the effort choice, participants were only shown the payoffs that resulted from their trade with their contracting partner, i.e., workers only learned the profit the firm was making with them and whether the firm had hired a second worker in the given period, but not any payoffs to the firm or another worker resulting from a possible second contract in the same period. This was done to prevent confounding effects of social comparison between workers (Abeler et al., 2010; Gächter and Thöni, 2010; Gächter et al., 2010) and, hence, to be able to focus on the incentive effects from the productivity shocks.

In addition to treatment GTBT2 where wages were set by employers, we conducted a second multi-worker treatment in which, in a similar manner as in the single-worker setting, we exogenously fixed the wage offers at a level of 30, thereby eliminating all signaling content in the wage offer (denoted GTBT2_FIX henceforth). Our setup allows us to analyze the differences in behavior between single-worker and multi-worker firms with productivity shocks under exogenous wage rigidities. We conducted the experiments in the MELESSA laboratory of the University of Munich in 2009. We implemented six markets of each treatment, such that in total 132 subjects took part in the experiment where they earned 27.8 €, on average, for sessions that lasted less than two hours. No subject that had already taken part in the one-worker experiments took part in the two-worker experiments.

Theoretical predictions for the two treatments are straightforward and very similar to the ones derived for our main treatments. Under "standard" assumptions, it is optimal for profit-maximizing firms to offer wages of 6 when productivity is high since selfish and rational workers cannot be forced to exert higher than minimal effort levels. The firm is indifferent to hire a second worker at minimal efforts as this leaves her profit level unchanged.[38] If the productivity level is low in BT, a firm cannot make any offer that

does not make a difference which worker provided what level of effort when calculating final firm profits. Distinguishing between the workers or assigning worker productivity types would have complicated the experiment considerably without adding much additional insight.

[38]If a firm hires one worker at a wage of 6, she makes $\pi_f = 10 \cdot 1 - 6 = 4$; with two workers exerting minimal effort she earns $\pi_f = 8 \cdot (1 + 1) - 6 - 6 = 4$.

guarantees at least zero profits, neither when she hires one worker ($\pi_f = 5 \cdot 1 - 6 = -1$), nor when she employs two workers ($\pi_f = 4 \cdot (1+1) - 6 - 6 = -4$) such that the theoretical prediction is full unemployment in BT. In treatment GTBT2_FIX with a fixed wage of 30, firms cannot make any offers that leave them with non-negative payoffs when workers exert minimal effort levels. Regardless of the productivity level, full unemployment prevails with money-maximizing workers with a wage of 30.

Similar to the results from the one-worker setting, if there are enough fair-minded workers in the population, firms pay high wages and workers reciprocate by exerting high levels of effort, where the threat of ending the relationship makes it profitable for both parties to adhere to these strategies. We determine optimal behavior in the presence of non-selfish workers in appendix A1.1 in the same way as we did for the single worker case to find critical values $\gamma^*$ of the fraction of non-selfish workers to sustain an efficient equilibrium with either one or two workers. One complicating issue in the context of two-worker firms is the definition of the reference group: equilibrium cut-off levels of $\gamma$ depend on whether the worker compares his profit only with the firm's profit or also with the other worker's profit. Since we give no information about payoffs from a potential co-worker to workers, we consider the first scenario a more reasonable benchmark for our setup. The cut-off levels for sustaining fairness equilibria become slightly more demanding in the two-worker case than in the one-worker case. Details can be found in appendix A1.1. We expect that the negative effect of BT on employment is stronger under rigid wages than under flexible wages.

### 1.5.2 Results

#### Employment Levels

We begin by comparing the one-worker setting and the two-worker setting with respect to employment.[39] Table 1.10 reveals that employment levels are lower in treatment GTBT2 than in treatment GTBT (Mann-Whitney-U-test, $p = 0.01$). The difference is almost identical in GT and in BT. It seems that firms are deliberately rationing jobs, even

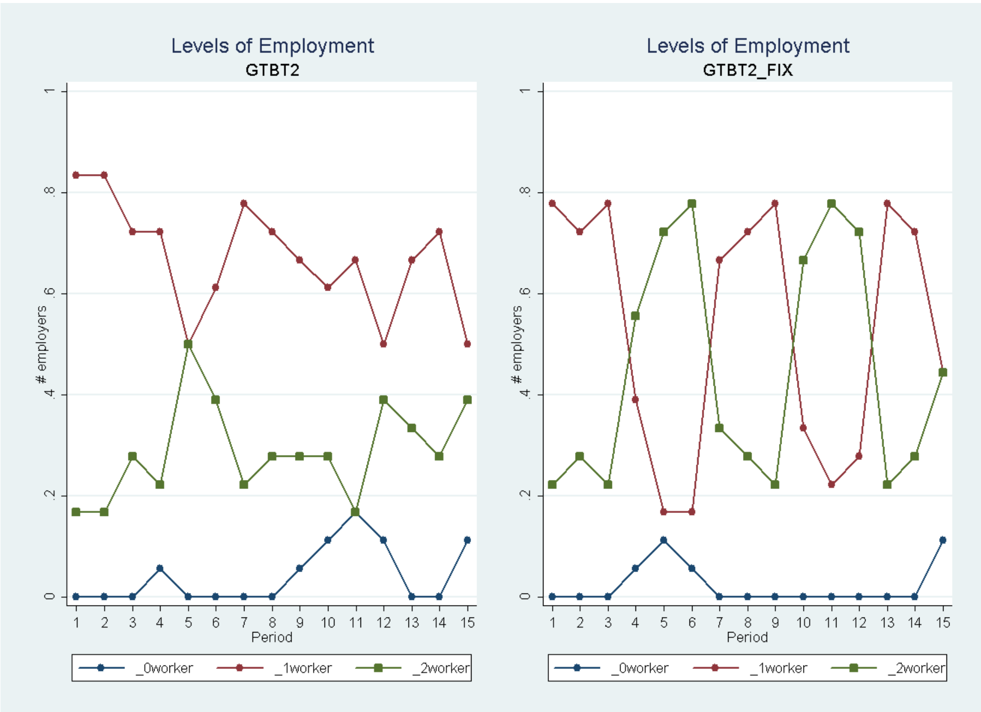
[39] We again refer to a situation as having "full employment" when all firms have hired the maximum number of workers allowed, i.e., 5 workers per market in the single-worker setting and 6 in the two-worker setting.

Table 1.10: Employment Levels for GTBT, GTBT_FIX, GTBT2 and GTBT2_FIX.

| treatment | Overall Employment | GT Employment | BT Employment |
|---|---|---|---|
| GTBT | 95.5 % | 98.2 % | 91.7 % |
| GTBT_FIX | 99.1 % | 98.5 % | 100 % |
| GTBT2 | 81.5 % | 84.9 % | 76.4 % |
| GTBT2_FIX | 75.4 % | 84.9 % | 61.1 % |

though average levels of unemployment in treatment GTBT2 are again significantly below the predicted levels based on standard assumptions. The negative employment effect is amplified in treatment GTBT2_FIX, which is mainly caused by BT-periods where most firms ration employment to one worker. An employment level of more than 61% is of course still far above the theoretical prediction of full unemployment. Figure 1.11 depicts the pattern of employment policy of firms over time.

Figure 1.11: Employment Strategies over Time




Percentage of employers that hire no worker, one worker or two workers in GTBT2 and GTBT2_FIX

**Result 1.10** *When the production technology is characterized by decreasing returns to scale and firms can employ more than one worker, BT-periods show lower levels of employment. This effect is amplified by the presence of wage rigidities, but the level of employment is still much higher than the theoretical prediction under standard assumptions.*

At first sight, the results from the flexible wage treatment lend partial support to the hypothesis of Bewley (1999) that firms prefer to get the least productive worker out of the door in BT rather than suffering from the negative incentive effects when cutting their wages. In order to get a comprehensive picture of the actual policy that multi-worker firms implement in BT, we next turn to the analysis of wages and efforts.

### Effort Levels and Wage Setting Behavior

Under flexible wages, we confirm our findings from GTBT and observe an almost identical wage pattern as in Figure 1.2 based on the single-worker setting (Mann-Whitney-U-test, $p = 0.52$), see Figure A1.10 in appendix A1.6.[40] Firms cut wages by 31% and 40%, respectively, when the productivity shock occurs such that we can refute the hypothesis that firms only act on the employment in BT, leaving wages unaffected. Apparently, firms use a mixture policy by reducing employment and cutting wages after a productivity shock, with a stronger tendency to cut wages.

OLS regressions, given in Table 1.11, find robust positive relationships between the level of effort and the wage levels, private contracts and a quadratic time trend.[41] Comparing the two treatments indicates that also in the two-worker setting, BT elicit additional effort provision. Controlling for the wage, productivity shocks increase effort levels by about one point. However, there is no additional effort from the interaction term of a productivity shock and the fixed-wage treatment, indicating that workers do not provide significantly higher effort in BT for a given wage level. Interestingly, workers exert slightly lower levels

[40] If not otherwise indicated, we consider every contract separately rather than aggregating over firms, i.e., if a firms hires two workers in one period, it concludes two contracts. See Figure A1.9 in appendix A1.6 for average wages and efforts in the multiple-worker treatments and Figure A1.10 also for an effort comparison between the single- and the multiple-worker setting.

[41] We also run tobit regressions which yield very similar results.

Table 1.11: OLS Regression on Effort for GTBT2 and GTBT2_FIX

| | GTBT2 and GTBT2_FIX | |
|---|---|---|
| | (16) | (17) |
| Wage | 0.146*** | 0.163*** |
| | (0.007) | (0.006) |
| Private | 1.004*** | 0.451*** |
| | (0.094) | (0.081) |
| Period | 0.771*** | 0.574*** |
| | (0.085) | (0.082) |
| Period$^2$ | -0.051*** | -0.039*** |
| | (0.006) | (0.006) |
| GTBT2_FIX | 0.213 | 0.138*** |
| | (0.237) | (0.178) |
| BT-periods | | 1.084*** |
| | | (0.268) |
| BT-periods | | 0.286 |
| × GTBT2_FIX | | (0.251) |
| Two-workers | 0.126 | 0.457 |
| | (0.271) | (0.256) |
| Second contract | -0.372*** | -0.362*** |
| | (0.088) | (0.087) |
| Constant | -0.653 | -1.138*** |
| | (0.402) | (0.332) |
| Obs | 847 | 847 |
| $R^2$ | 0.46 | 0.49 |
| | Prob= 0.00 | Prob=0.00 |

OLS regression on effort for GTBT2 and GTBT2_FIX. Standard errors are given in brackets, clustering on the market level. *** indicates significance at the 1 % level, ** at the 5 % and * at the 10 % level, respectively.

of effort if their contract is the second contract to an employer in a given period, although they can only infer this from a later time of acceptance. This could reflect the response of a worker when he is discouraged by the fact that he correctly believes not to be the number-one worker of a given firm.

Workers were only informed about their own payoff but not about the payoff of a potential co-worker at the same firm. We therefore expect no difference in average effort levels compared to the single-worker setting, like we did not obtain any differences in wages either. This is confirmed for the treatments with flexible wages (Mann-Whitney-U-test, $p = 0.26$), but not for the treatment with fixed wages (Mann-Whitney-U-test, $p = 0.02$). A possible explanation for the difference with fixed wages could be latent worker competition given the attractiveness of an employment with a guaranteed wage of 30, with effort levels being slightly higher in GTBT2_FIX than in GTBT_FIX.

Table 1.12: Panel Regression on Firm Profits

| | GTBT2 and GTBT2_FIX, RE | |
|---|---|---|
| | (18) | (19) |
| Average wage | -0.174 | -1.541*** |
| | (0.111) | (0.077) |
| Average effort | 6.884*** | 11.638*** |
| | (0.442) | (0.421) |
| Private | -3.811 | -0.894 |
| | (2.737) | (1.142) |
| GTBT2_FIX | -1.879 | 1.488 |
| | (1.354) | (1.285) |
| BT-periods | | -32.641*** |
| | | (4.200) |
| BT-periods | | -11.236*** |
| × GTBT2_FIX | | (3.562) |
| Two-workers | 12.014** | 4.256*** |
| | (2.954) | (0.867) |
| Two-workers | | -9.350*** |
| × BT | | (2.468) |
| Constant | -19.114*** | 7.319*** |
| | (5.279) | (1.522) |
| Obs | 523 | 523 |
| | Wald-$\chi^2(5)$=368.6 | Wald-$\chi^2(8)$=2958.9 |
| | Prob= 0.00 | Prob=0.00 |

Random effects panel estimation on firm profits. Standard errors are given in brackets, clustering on the market level. *** indicates significance at the 1 % level, ** at the 5 % and * at the 10 % level respectively.

In order to assess the profitability of the hiring strategy of firms, we take a closer look at the determinants of their profits. We therefore aggregate firm behavior in one period into one observation by calculating average efforts and wages in case a firm hired two workers. Table 1.12 with panel regressions confirms our hypothesis that it has indeed a negative impact on total firm profits per period when firms hire two workers in BT. Not surprisingly, firms suffer considerably from the productivity shock in terms of profits, especially when firms are forced to pay a wage of 30. In contrast, profits are significantly higher for two-worker firms in GT.

**Result 1.11** *Under flexible wages and multiple-workers, firms combine both wage cuts and a reduction in workforce to overcome productivity shocks, but workers do not withdraw effort provision accordingly. In the absence of wage flexibility, we observe firms to substitute wage cuts with a more restrictive hiring policy.*

### 1.5.3 Discussion

In our experimental labor markets, we are not able to unambiguously confirm the argument Bewley (1999) has brought forward to explain wage rigidities. Workers accept wage cuts without reducing efforts accordingly such that it is the best strategy for firms to reduce wages when they experience a productivity shock. The fairness norms of workers in our experimental labor markets are not governed by the nominal levels of their wages, but rather by the share of the surplus between them and their employer. Figure A1.11 in appendix A1.6 illustrates that although this share varies from 40 % to 70 % between GT and BT, the share from the treatments GTBT2 and GTBT2_FIX track each other surprisingly well and are also not significantly different from the surplus split in the one-worker setting in Figure A1.1. Pertaining to overall market efficiency, the positive effect of exogenous wage rigidity on effort levels in the single-worker setting (GTBT_FIX) is counterbalanced by the negative employment effect in the two-worker setting (GTBT2_FIX) in BT. Taking both effects together, markets in GTBT_FIX realize 78% of the possible surplus in bad times, whereas GTBT2_FIX markets attain efficiency levels of only 66%. The negative employment effect more than offsets the positive effect from higher efforts in BT in terms of efficiency. The corresponding levels for GTBT and GTBT2 are 49 % and 57 % which indicates that fixing the wage in our setup generally increases efficiency. The increase is more pronounced for the single-worker treatments, where firms cannot adapt employment levels other than not hiring a worker at all.

At this point, we can only speculate about the reasons why we are not able to reproduce the hiring policies put forward by Bewley (1999) under flexible wages in the presence of a "recession". One aspect is possibly related to the difficulties to reproduce the notion of "morale" from real world labor markets in the laboratory without real effort tasks.[42] The results from the two-worker setting show, however, that firms readily act on both - wage and employment levels - when being hit by a productivity shock depending on the level of wage rigidity prevalent in the market. In general, a static interpretation of only wages and efforts in the spirit of the fair-wage hypothesis falls short of capturing important aspects of reciprocity between workers and firms when labor markets are prone to fluctuations.

[42]We did not frame our experiments as labor markets either, but rather as neutral goods markets thereby also removing framing effects from a labor market setting.

## 1.6 Conclusion

In this chapter, we extend the fair wage-effort hypothesis to a setup where firms are facing productivity shocks on labor markets with non-enforceable but observable effort. Workers build up a reputation such that firms and workers engage in long-term relationships in which firms cut wages considerably in BT and workers only reduce efforts marginally. As a consequence, under flexible wages new wage-effort relationships arise endogenously in BT as workers do not withdraw effort accordingly when their wage falls. The disciplining threat of not employing a worker is virtually never carried out by firms as we observe close to full employment in all our treatments. When we restrict offered wages to be constant over time, thereby eliminating wage cuts in BT, effort levels increase significantly compared to GT, without affecting the employment level at all but increasing efficiency. Removing reputational mechanisms, however, has harmful effects for both the employment level and the achieved market efficiency particularly in BT. Our results lend support to what we call the "fair employment hypothesis“, i.e., the fact that reciprocity prevails not only between wage levels and exerted effort in a static way, but even more so in a dynamic manner between firm's actual employment decision and efforts over GT and BT. We provide evidence for this hypothesis in the form of intertemporal elements of reciprocity at the micro level between firms and workers when markets enter times of low productivity. Relational contracts help to increase flexibility for the labor market inside the firm, compared to the outside market. We furthermore show in a multiple-worker treatment that firms substitute changes in the wage level by an adaptation of their employment level to adjust a productivity shock if there is no wage flexibility.

Relating our findings to real-world indicators in the spirit of our stylized model with non-selfish workers, it is somewhat unclear to what precisely workers compare their wage to, as firm's overall profits may yield a very noisy indicator. This notwithstanding and as a response to low firm productivity, increasing individual effort under downward rigid wages on real labor markets in bad economic conditions is consistent with our experimental evidence, which keeps the unemployment rate low as the recent economic crisis demonstrated. After all, it is not too difficult for workers to obtain good signals on the current situation of their firm. Furthermore, our results raise the question if wages are really as downward rigid as the economic literature has been underlining or if real wages are subject to (minor) downward adaptations dependent on the situation of the economy. Our findings show that intertemporal considerations between firms and workers play an important role on labor markets. It is not clear from our results how these dynamic considerations interact with institutions and features of labor market that equally affect behavior of firms and agents, e.g. the level of competition, collective agreements, the level of sick pay, etc. We hope that our study will encourage further analyses of the interplay between factors that are exogenous to the contracting environment and endogenously arising incentives in a relational contract. More specifically, it would be interesting to investigate the effects of asymmetric information in our setting where only firms have knowledge about the state of the economy. What would happen if shocks were unpredictable and idiosyncratic to specific firms? What if firms could go bankrupt? How would markets behave if there was wage rigidity within firms, but not on the market? Answers to all of these questions are important ingredients for the understanding of phenomena that we observe on real-world labor markets. We think that - next to empirical analyses of specific labor market features such as job protection, probation time, etc. - particularly the incentives arising from labor market fluctuations on individual behavior of both firms and workers are largely unexplored and deserve future research.

## 1.7 Appendix A1

### Appendix A1.1: Theoretical Predictions

In this appendix, we want to illustrate theoretically the consequences of market-wide productivity shocks on the profitability of different contracting strategies by firms when there are two types of workers in the market, fair-minded and selfish ones. We thereby do not aim at characterizing the full set of possible Perfect Bayesian Nash Equilibria, but rather show how market outcomes can change in the presence of fairness preferences among workers. More precisely, we analyse how the fraction of fair-minded workers in the population governs a different contracting behavior by firms between GT and BT.

### One worker firms

We assume that workers are of two different types where types are private information. A fraction $0 \leq \gamma \leq 1$ of workers has egalitarian preferences such that they are assumed to fulfill the contract offered to them as long as they are offered at least the equal split of the surplus being generated by the transaction.[43] If they are offered less than the equal split, workers accept the contract but shirk by providing the minimum level of effort as their fairness norm of equal division has been violated. The remaining fraction $(1-\gamma)$ of workers is purely selfish and would shirk in a one-shot interaction. The utility function of a fair-minded worker in GT can be expressed by

$$u(w, e, \tilde{e}) = \begin{cases} w - c(e) & \text{if } w - c(\tilde{e}) < \frac{1}{2}[p_{GT,1w}\tilde{e} - c(\tilde{e})] \\ w - c(e) - k \max[\tilde{e} - e; 0] & \text{if } w - c(\tilde{e}) \geq \frac{1}{2}[p_{GT,1w}\tilde{e} - c(\tilde{e})] \end{cases} \tag{A1.1}$$

where $w$ is the offered wage, $\tilde{e}$ and $e$ desired and actual effort levels with the associated cost function $c(\cdot)$, $p_{GT,1w}$ indicates the productivity parameter of the firm in GT and $k$ is a fairness parameter that implies that a fair-minded worker chooses the desired effort

[43]We do not rule out that firms also have some sort of fairness preferences, but since workers are always reacting to the contract offer by firms with their effort decision, we consider their preferences over outcomes at the heart of the stylized transaction. For the purpose of this appendix, assuming fairness preferences of firms other than in their contracting policy would complicate the analysis but add little insight.

level $\tilde{e}$ as long as he obtains (at least) half of the net surplus created from the contract. In order to guarantee this, we furthermore assume that

$$k > \max c'(e) \tag{A1.2}$$

such that $e = \tilde{e}$ is indeed optimal since the fairness costs associated with choosing a lower level of effort than desired in the contract outweigh the material benefits from a lower effort level.

We assume that firms are risk neutral and maximize their expected monetary payoff. In the last period of the game $T$, a firm can either offer a "trust" contract that equalizes payoffs between the firm and the worker characterized by above minimal efforts and wages or offer workers their outside option in the form of a "standard" contract. This contract has to account for the level of the unemployment benefit of 6, i.e., firms have to offer at least a wage of 6 to induce workers to accept.[44] With the parameters of our experiment this outside option gives firms a last period profit of

$$\pi_T^f(w, \tilde{e}) = \pi_T^f(6, 1) = 10 \cdot 1 - 6 = 4 \tag{A1.3}$$

In the last period when productivity is high, a firm can do better by offering the "trust" contract $[w, \tilde{e}] = [59, 10]$ if the fraction $\gamma$ of fair-minded workers in the population is large enough. This contract yields a payoff of $\pi_w = \pi_f = 41$ for both the firm and the worker, such that fair-minded workers would accept and exert the desired effort level of $\tilde{e} = 10$.[45] For such a "trust" offer to be profitable the following condition has to hold:

$$\gamma(p_{GT,1w} \cdot \tilde{e} - w) + (1 - \gamma)(p_{GT,1w} \cdot 1 - w) \geq \pi_T^f(6, 1) \tag{A1.4}$$

which boils down to $\gamma \geq \gamma^*_{GT,1w}$ with $\gamma^*_{GT,1w} \approx 0.59$. So if $\gamma \leq \gamma^*_{GT,1w}$, a firm is better off offering the "standard" contract $[w, \tilde{e}] = [6, 1]$ in the last period. In all pre-final periods, the firm will employ a policy of contingent contract renewal, i.e., re-employ the worker if

[44] In what follows we will assume that if workers are indifferent between accepting the unemployment benefit and accepting an offer with a wage of 6 and shirking, they choose to do the latter.

[45] Note that if workers are characterized by egalitarian preferences as described above, it is never optimal for a firm to make a payoff equalizing contract offer in which it desires an effort level that is lower than $\tilde{e} = 10$.

he exerted the desired effort level and dismiss him if he shirked. A fair worker will accept any offer that shares the surplus equally and selfish workers will cooperate if future rents from cooperating exceed the gains from shirking. Hence, in the pre-final period a selfish worker has to choose between cooperating and shirking in the last period, which gives him a payoff of $\pi^w_{T-1} + \pi^w_T = w - c(10) + w - c(1) = 100$, and shirking right away and remaining unemployed in the last period, i.e., $\pi^w_{T-1} + \pi^w_T = w - c(1) + b = 65$, such that selfish workers clearly prefer to cooperate in the pre-final period and reveal their type only in the last period when they shirk.

By backward induction this holds also for all prior periods regardless of the productivity level, since the continuation value of a relationship induces selfish types to work and their incentives are unaffected by the presence of periods of low productivity. The fair contract in BT, however, has a decreased wage and amounts to $[w, \tilde{e}] = [34, 10]$ which yields identical payoffs to both parties of $\pi_w = \pi_f = 16$.

Optimal behavior in GTBT therefore depends on the value of $\gamma$:

- $\gamma > \gamma^*_{GT,1w}$: Firms offer the "trust" contract $[w, \tilde{e}] = [59, 10]$ in GT and $[w, \tilde{e}] = [34, 10]$ in BT. The contract is accepted by all workers who exert the desired effort level in all but the last period T. Workers earn $\pi^w_{GT} = 59 - 18 = 41$ in GT and $\pi^w_{BT} = 34 - 18 = 16$ in BT until the last period T in which fair-minded workers earn $\pi^w_T = 59 - 18 = 41$, but selfish workers shirk to obtain a payoff of $\pi^w_T = 59 - 0 = 59$. In all non-final period, firms earn $\pi^f_{GT} = 10 \cdot 10 - 59 = 41$ in GT and $\pi^f_{BT} = 5 \cdot 10 - 34 = 16$ in BT. In the last period, firms earn $\pi^f_T = \gamma(41) + (1 - \gamma)(-49) \geq \pi^f_T(6, 1)$ $\forall \gamma > \gamma^*_{GT,1w}$

- $\gamma^*_{GT,1w} > \gamma$: Firms offer the "standard" contract $[w, \tilde{e}] = [6, 1]$ in GT and do not make an offer in BT. The contract is accepted by all workers who shirk by exerting the desired effort level of 1 all periods. Workers earn $\pi^w_{GT} = 6 - 0 = 6$ in GT and the unemployment benefit $b = 6$ in BT. Firms earn $\pi^f_{GT} = 10 \cdot 1 - 6 = 4$ in GT and nothing in BT.

These strategies imply that a firm can re-hire a worker it had employed in the past through an identification mechanism as present in our main treatments. When wages are fixed at 30 in our GTBT_FIX treatment, the mechanism is identical as outlined above, with the exception that the "trust" contract is $[w, \tilde{e}] = [30, 5]$ in GT $[w, \tilde{e}] = [30, 9]$ in BT. The lower wage of 30 in GT does not destroy the incentives for selfish types in the pre-final period to mimick the non-selfish types.

In the absence of this reputation mechanism in GTBT_RI, every period can be treated like the last period. Firms have to rely entirely on a belief about $\gamma$ throughout the experiment, when they enter the offer phase in every period. Selfish types do not have an incentive to provide effort to stay in the relationship, such that in every period $\gamma$ has to be sufficiently high to induce firms to offer the "trust" contract. $\gamma > \gamma^*_{GT,1w}$ ensures that firms offer a "trust" contract in GT (identical to the last period from above), but in BT there is an additional condition for the critical value of $\gamma$ to sustain an efficient equilibrium. This critical value $\gamma^*_{BT,1w}$ stems from the incentive of firms in BT, i.e., $\pi^f_{BT} = \gamma(16) + (1-\gamma)(-29) \geq 0 \; \forall \gamma > \gamma^*_{BT,1w}$ where $\gamma^*_{BT,1w} \approx 0.64$. In GTBT_RI, the first condition from above only holds for $\gamma > \gamma^*_{BT,1w}$. If $\gamma^*_{BT,1w} > \gamma > \gamma^*_{GT,1w}$, the equilibrium looks as follows:

- $\gamma^*_{BT,1w} > \gamma > \gamma^*_{GT,1w}$: Firms offer the 'fair' contract $[w, \tilde{e}] = [59, 10]$ in GT and do not make an offer in BT. In GT, the contract is accepted by all workers where only the fair-minded workers exert the desired effort level and the selfish workers shirk. Workers earn $\pi^w_{GT} = 59 - 18 = 41$ (fair-minded) or $\pi^w_{GT} = 59 - 0 = 59$ (selfish) in GT and the unemployment benefit $b = 6$ in BT. Firms earn $\pi^f = \gamma(41) + (1-\gamma)(-49) \geq \pi^f(6, 1) \; \forall \gamma > \gamma^*_{GT,1w}$ in GT and nothing in BT.

If $\gamma^*_{GT,1w} > \gamma$, the inefficient equilibrium from above prevails also for GTBT_RI in all periods.

## Two worker firms

When firms are not restricted to hire only one worker, they can choose between different employment strategies. The incentives for firms compared to the one-worker case are unchanged if they restrict themselves to hire indeed only one worker. Decreasing returns to scale of the second worker are set in a way that the marginal benefit of hiring a second worker exceeds marginal costs, making full employment of two workers the efficient outcome. In the last period, their outside option is to offer the "standard" contract $[w, \tilde{e}] = [6, 1]$ to either one or two workers which gives them

$$\pi_T^f(w, \tilde{e}) = \pi_T^f(6, 1) = \begin{cases} p_{GT,1w} \cdot e - w = 10 \cdot 1 - 6 = 4 & \text{if they hire one worker} \\ p_{GT,2w} \cdot e - w = 8 \cdot (1 + 1) - 6 - 6 = 4 & \text{if they hire two workers} \end{cases} \tag{A1.5}$$

which means that they are indifferent between hiring one or two workers in the final period. We again assume egalitarian preferences on behalf of the workers, but these can now take two different forms, depending on their fairness benchmark when a firm hires two workers. It is crucial to note that in our experiment, there was no information about the terms of the contract of a potential co-worker before a worker exerted costly effort. This was implemented to focus on the relationship between firm and worker by avoiding peer effects between multiple workers. So we assume that the worker could condition his effort decision neither on the existence of a second worker (which he learned at the end of every period) nor on the contract terms of a potential colleague (which he was not told at anytime).[46] The way how workers' fairness preferences for multiple worker relationships look like determines the wage firms have to offer in order to induce workers to accept their terms. We distinguish between two cases: relationship-specific ("bilateral") and strictly egalitarian preferences ("trilateral").

**Relationship specific fairness preferences ("bilateral")**

In the two-worker scenario, we therefore assume that a "fair-minded" worker exerts the desired effort level as soon as the contract terms guarantee him at least half of the surplus *from the transaction* with the firm. Furthermore, we make the assumption in this section that a firm hires two workers, since the incentives for the one-worker case are identical to the ones discussed in the previous section.

[46]We thus abstract from the possibility of signaling between firms and workers on the existence of a second through contract terms.

The preferences of a worker in GT under this set of assumptions can be represented by

$$u(w,e,\tilde{e}) = \begin{cases} w - c(e) & \text{if } w - c(\tilde{e}) < \frac{1}{2}[p_{GT,2w}\tilde{e} - c(\tilde{e})] \\ w - c(e) - k\max[\tilde{e} - e; 0] & \text{if } w - c(\tilde{e}) \geq \frac{1}{2}[p_{GT,2w}\tilde{e} - c(\tilde{e})] \end{cases} \tag{A1.6}$$

Thus, the "trust" offer is now given by $[w, \tilde{e}] = [49, 10]$ yielding a payoff from the transaction of $\pi^w_{GT} = 49 - 18 = 31$ for workers and $\pi^f_{GT} = 80 - 49 = 31$ for firms that hire two workers. With a second worker for identical conditions, firms would earn $\pi^f_{GT} = 160 - 98 = 2 \cdot 31 = 62$.[47] We again can compute the fraction of fair-minded workers in the population required to make a "fair" offer $[w, \tilde{e}] = [49, 10]$ profitable for firms in the last period through

$$\gamma^2(p_{GT,2w} \cdot \sum \tilde{e} - 2w) + (1-\gamma)^2(p_{GT,2w} \cdot (1+1) - 2w) + 2\gamma(1-\gamma)(p_{GT,2w} \cdot (\tilde{e}+1) - 2w) \geq \pi^f_T(6,1) \tag{A1.7}$$

The first term captures a firm employing two fair-minded workers which happens with probability $\gamma^2$, the second term if the firm employs two selfish workers with probability $(1-\gamma)^2$ and the third term if she employs one selfish and one fair-minded worker. The critical threshold is given by $\gamma \geq \gamma^*_{GT,2w}$ with $\gamma^*_{GT,2w} \approx 0.60$. All selfish workers have an incentive to mimic fair-minded types in all pre-final periods including BT, such that they shirk only in the last period of the experiment. Similarly to the one-worker case we can describe market outcomes in GTBT2 under a policy of contingent contract renewal of firms which depend on the value of $\gamma$:

- $\gamma > \gamma^*_{GT,2w}$: Firms offer the "trust" contract $[w, \tilde{e}] = [49, 10]$ to two workers in GT and $[w, \tilde{e}] = [29, 10]$ to two workers in BT. The contract is accepted by all workers who exert the desired effort level in all but the last period. Workers earn $\pi^w_{GT} = 49 - 18 = 31$ in GT and $\pi^w_{GT} = 29 - 18 = 11$ in BT until the last period in which fair-minded workers earn $\pi^w_T = 49 - 18 = 31$, but selfish workers shirk to obtain a payoff of $\pi^w_T = 49 - 0 = 49$. In all non-final periods, firms earn $\pi^f_{GT} = 8 \cdot 20 - 49 - 49 = 62$ in GT and $\pi^f_{BT} = 4 \cdot 20 - 29 - 29 = 22$ in BT. In the last period, firms earn $\pi^f_T = \gamma^2(62) + (1-\gamma)^2(-82) + 2\gamma(1-\gamma)(-10) \geq \pi^f_T(6,1)\ \forall \gamma > \gamma^*_{GT,2w}$.

[47] If a firms hires only one worker at these conditions, it would earn $\pi^f_{GT} = 100 - 49 = 51$ and thus less than with a second worker.

- $\gamma^*_{GT,2w} > \gamma > \gamma^*_{GT,1w}$: Firms offer the "trust" contract $[w, \tilde{e}] = [59, 10]$ in GT and $[w, \tilde{e}] = [34, 10]$ in BT to one worker. The contract is accepted by all workers who exert the desired effort level in all but the last period. Workers earn $\pi^w_{GT} = 59 - 18 = 41$ in GT and $\pi^w_{BT} = 34 - 18 = 16$ in BT until the last period in which fair-minded workers earn $\pi^w_T = 59 - 18 = 41$, but selfish workers shirk to obtain a payoff of $\pi^w_T = 59 - 0 = 59$. In all non-final period, firms earn $\pi^f_{GT} = 10 \cdot 10 - 59 = 41$ in GT and $\pi^f_{BT} = 5 \cdot 10 - 34 = 16$ in BT. In the last period, firms earn $\pi^f_T = \gamma(41) + (1-\gamma)(-49) \geq \pi^f_T(6, 1)\ \forall \gamma > \gamma^*_{GT,1w}$.

- $\gamma^*_{GT,1w} > \gamma$: Firms offer the "standard" contract $[w, \tilde{e}] = [6, 1]$ in GT and do not make an offer in BT. The contract is accepted by all workers who shirk by exerting the desired effort level of 1 in all periods. Workers earn $\pi^w_{GT} = 6 - 0 = 6$ in GT and the unemployment benefit $b = 6$ in BT. Firms earn $\pi^f_{GT} = 10 \cdot 1 - 6 = 4$ in GT, are indifferent whether they hire one or two workers in GT and earn nothing in BT.

In comparison to GTBT, an additional condition on $\gamma$ is required to sustain an equilibrium where firms hire two workers instead of one in GTBT2. The mechanism in GTBT2_FIX is identical, only contract offers are different since wages are exogenously fixed at 30. In GT, the "trust" contract is given by $[w, \tilde{e}] = [30, 6]$ such that firms earn $\pi^f_{GT} = 48 - 30 = 18$ and workers $\pi^w_{GT} = 30 - 8 = 22$ if they adhere to the contract terms. In BT, firms offer $[w, \tilde{e}] = [30, 10]$ to earn $\pi^f_{BT} = 40 - 30 = 10$, with $\pi^w_{BT} = 30 - 18 = 12$ left on the table for workers. The critical threshold of $\gamma$ to make it profitable in GTBT_FIX to hire two workers throughout the 15 periods amounts to $\gamma^*_{GT\_FIX,2w} \approx 0.55$.
Hence, depending on the proportion of fair-minded workers, market outcomes take on different forms from a complete market breakdown to efficient contracting despite the moral hazard problem. If $\gamma$ is high enough, it is always better for firms to employ two workers, and only when $\gamma$ decreases we expect negative employment effects.

**Egalitarian fairness preferences ("trilateral")**

In this section, we only briefly discuss the consequences of workers having egalitarian preferences which require them to obtain *a third of the overall generated surplus when* *a firm employs two workers* in order to exert the desired level of effort. Since in our experiment workers were not told the conditions of another contract between their firm and a second worker, we do not consider this theoretical case a real benchmark such that we only briefly comment on it here.

Applying the same reasoning as above, firms now have to make a "trust" offer $[w, \tilde{e}] = [60, 10]$ to both workers in the last period to induce workers to work such that every worker obtains $\pi_T^w = 60 - 18 = 42$ and the firm $\pi_T^f = 8 \cdot 20 - 120 = 40$. The critical threshold of $\gamma$ for this to be profitable naturally increases to $\gamma_{GT}^* \approx 0.75$. In BT, the "trust" offer would have to be $[w, \tilde{e}] = [33, 10]$ leading to profits of $\pi_T^w = 33 - 18 = 15$ for workers and $\pi_T^f = 4 \cdot 20 - 66 = 14$ for firms. Under this stronger case of fairness preferences among workers, efficient market outcomes are even harder to achieve compared to the bilateral case.

## Appendix A1.2: Efficiency and Surplus Sharing in Good and Bad Times

In this appendix, we present the results on efficiency and surplus sharing in the three main treatments GTBT, FGT and FBT that can be inferred from the evidence on wages and efforts in the text.

Remember that exerting maximal effort at $e = 10$ is always efficient, regardless of the treatment since it generates the biggest possible surplus to be shared by the two contracting parties. As a measure of efficiency, we calculate the generated surplus as the sum of the employer's and the worker's profits and subtract the outside option of the worker of 6. We compare this with the maximal achievable surplus under complete contracts of 76 in GT and 26 in BT.[48] Table A1.1 shows efficiency levels according to our measure over time and overall averages that naturally reinforce the conclusions from section 1.4.1. The fluctuations in GTBT lead to a decrease in efficiency, such that in none of the three-period clusters the market reaches the efficiency level of the respective control treatments. The main reason for this are the sharp wage drops that occur in BT of GTBT. Mann-Whitney-U-test tests on aggregate market efficiency confirm significant differences between FBT and GTBT ($p = 0.025$), but not between FGT and GTBT ($p = 0.36$). When looking at BT periods separately, the difference between FGT and GTBT is highly significant ($p = 0.006$). In contrast, there is no difference between GT periods in GTBT and FGT ($p = 0.85$).

**Result 1.12** *Efficiency levels are highest in FBT. The productivity shocks in GTBT result in a reduction in efficiency compared to the relevant periods in FGT and FBT.*

We next turn to the split of the surplus. Taking our stylized model with non-selfish workers seriously, the split of the surplus is a driving force for their behavior. As a measure for the split of surplus, we calculate the share of a worker's profit to the overall profits from trade for both parties. In FBT and FGT, this share converges to a level between 50%

[48] If $p_t = 10$, the surplus from bargaining under complete contracts and maximal efforts is $100 - c(10) - 6 = 100 - 18 - 6 = 76$. Under $p_t = 5$, the maximal surplus is given by $50 - c(10) - 6 = 50 - 18 - 6 = 26$. Our efficiency measures are not well-behaved for minimal efforts since they amount to non-zero values when workers shirk ($0.05$ if $p_t = 10$, $-0.04$ if $p_t = 5$), but this does not distort treatment comparisons.

Table A1.1: Relative Efficiency Levels in FBT, GTBT, and FGT

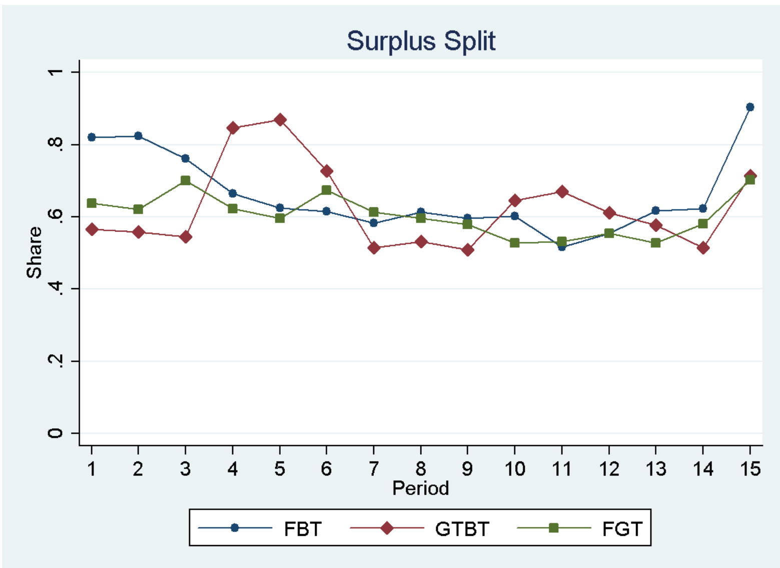
| Treatment | $p_t$ | period $1-3$ | period $4-6$ | period $7-9$ | period $10-12$ | period $13-15$ | Total Efficiency |
|---|---|---|---|---|---|---|---|
| FBT | 5 | 0.46 | 0.65 | 0.73 | 0.78 | 0.62 | 0.65 |
| GTBT | 5 | | 0.38 | | 0.48 | | 0.43 |
| | 10 | 0.43 | | 0.53 | | 0.52 | 0.49 |
| FGT | 10 | 0.46 | 0.48 | 0.60 | 0.68 | 0.59 | 0.56 |

Relative efficiency levels in treatments FBT, GTBT, and FGT measured as the ratio of the realized surplus minus the outside option divided by the theoretically feasible surplus under complete contracts

and 60% after some initial fluctuation. From period 4 onwards, workers generally succeed in obtaining higher profits than firms despite the excess supply of labor on the market. In treatment GTBT, in the first three BT-periods workers reap over 80% of the surplus and a somewhat lower but still slightly higher than 60%-share during the second three BT-periods. In the long run, the split of surplus is very similar in all three treatments.

**Result 1.13** *The split of the surplus appears to be similar in treatments FBT, FBT and GTBT, after an initial phase of different sharing norms. It is around* 40% *for firms and* 60% *for workers.*

Figure A1.1: Surplus Split in FBT, GTBT and FGT




Share of the surplus that accrues to the worker in treatments FBT, GTBT and FGT.

## Appendix A1.3: Contract Nature

In this appendix, we provide further evidence on the effects of productivity shocks on the nature of contracts concluded, focusing on the contract length and the renewal probability of contracts.

Figure A1.2: Cumulative Frequency of Relationship Length in FBT, GTBT and FGT

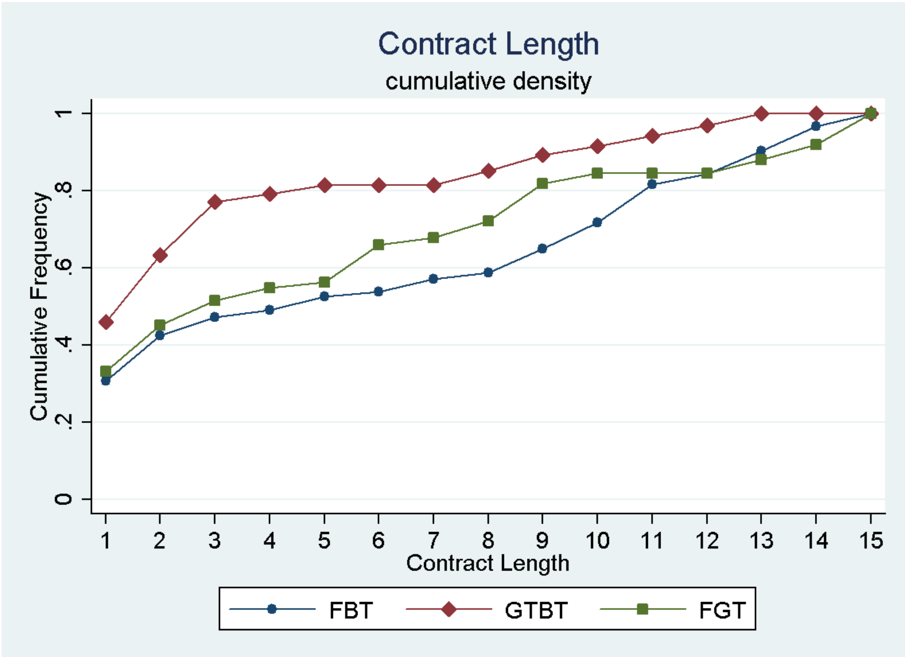



Figure A1.2 shows that the average contract length is lowest in GTBT compared to the two control treatments. We furthermore expect contracting behavior of firms to depend on their experience from past periods. If a worker has provided more than the desired effort, a firm will be willing to continue the relationship by offering him a new contract in the next period. A lower level of effort is thus more likely to induce the end of a relationship. These considerations will crucially depend on what firms and workers think to be a "fair" wage-effort relationship.

We hypothesize that in a stable economic environment a consensus between firms and workers is easier to reach as under varying economic conditions. If the conditions alter, we conjecture that adaptations to the new circumstances lead to higher break-up rates of contracts. Table A1.2 shows that the probability of renewing one's contract with the same employer is increasing in the effort provided in the previous period. This holds for

Table A1.2: Probability of Contract Renewal

| Effort in previous period | **FBT** | **GTBT** | **FGT** |
|---|---|---|---|
| **1** | 0.21 | 0.24 | 0.37 |
| **2** | 0.09 | 0.38 | 0.00 |
| **3** | 0.47 | 0.44 | 0.50 |
| **4** | 0.75 | 0.48 | 0.69 |
| **5** | 0.50 | 0.58 | 0.67 |
| **6** | 0.71 | 0.61 | 0.69 |
| **7** | 0.67 | 0.75 | 0.85 |
| **8** | 0.75 | 0.63 | 0.85 |
| **9** | 0.95 | 0.72 | 0.92 |
| **10** | 0.98 | 0.80 | 1.00 |

all three treatments demonstrating that firms reward a good worker performance with a new contract offer. In the GTBT treatment, this mechanism is qualitatively the same but not quantitatively. A higher effort in the previous period still increases the probability of a renewed contract in the next period, but to a lower extent. To check for this, we compute Spearman's $\rho$ between effort in the previous period and a dummy if a contract is renewed for each treatment to find significantly positive correlations between the two. The correlation is, however, lowest in the GTBT treatment (Spearman's $\rho$: FBT: 0.53, GTBT: 0.37, FGT: 0.53). In a similar manner, a probit regression of a contract renewal dummy on the effort provided in the previous periods yields highly significant coefficients in all treatments, but the lowest in the GTBT treatment. A fortiori, the unconditional probability of a contract being renewed is again lowest in the GTBT treatment (Mann-Whitney-U-test, $p = 0.02$ for FBT and GTBT, $p = 0.07$ for FGT and GTBT; Kruskal-Wallis test, $p = 0.05$).[49] We conclude, that the presence of commonly known productivity shocks for firms weakens the strength of the bilateralization of trade compared to stable conditions as in Brown et al. (2004) and our control treatments FGT and FBT.

[49]The unconditional probability of contract renewal on the basis of all concluded contracts amounts to 56 % in FGT, 53 % in FGT and only 38 % in GTBT.

## Appendix A1.4: Efficiency and Surplus Sharing in the Control Treatments

In this appendix, we report the results on efficiency and surplus sharing in the treatments GTBT, GTBT_RI and GTBT_FIX. We calculate the total generated surplus per contract, subtract the outside option of 6 and divide this measure by the theoretically maximal surplus of 76 in GT and 26 in BT.[50] Note that this measure abstracts from the employment effect, since we only look at concluded contracts.

Table A1.3: Relative Efficiency Levels in GTBT, GTBT_RI and GTBT_FIX

| Treatment | period 1 − 3 | period 4 − 6 | period 7 − 9 | period 10 − 12 | period 13 − 15 | Total Efficiency |
|---|---|---|---|---|---|---|
| GTBT_FIX | 0.48 | 0.68 | 0.58 | 0.77 | 0.47 | 0.60 |
| GTBT | 0.43 | 0.38 | 0.53 | 0.48 | 0.52 | 0.47 |
| GTBT_RI | 0.30 | 0.22 | 0.28 | 0.22 | 0.30 | 0.27 |

Relative Efficiency Levels in Treatments GTBT, GTBT_RI and GTBT_FIX measured as the ratio of the realized surplus minus the outside option divided by the theoretically feasible surplus under complete contracts

In the absence of reputation mechanisms, efficiency is clearly lowest across all control treatments (see Table A1.3). Particularly, BT-periods decrease efficiency sharply, which points again at the importance of reputation in phases of low productivity. In contrast, GTBT_FIX surprisingly displays the highest efficiency values. In fact, they are even higher than in GTBT, with markets performing especially well in BT-periods with realized efficiency levels of more than 70%. Hence, fixed wages do not necessarily have to have a negative impact on efficiency. Mann-Whitney-U-tests on aggregate market efficiency confirm that efficiency in GTBT is significantly different from GTBT_FIX ($p = 0.025$) and from GTBT_RI ($p = 0.02$).

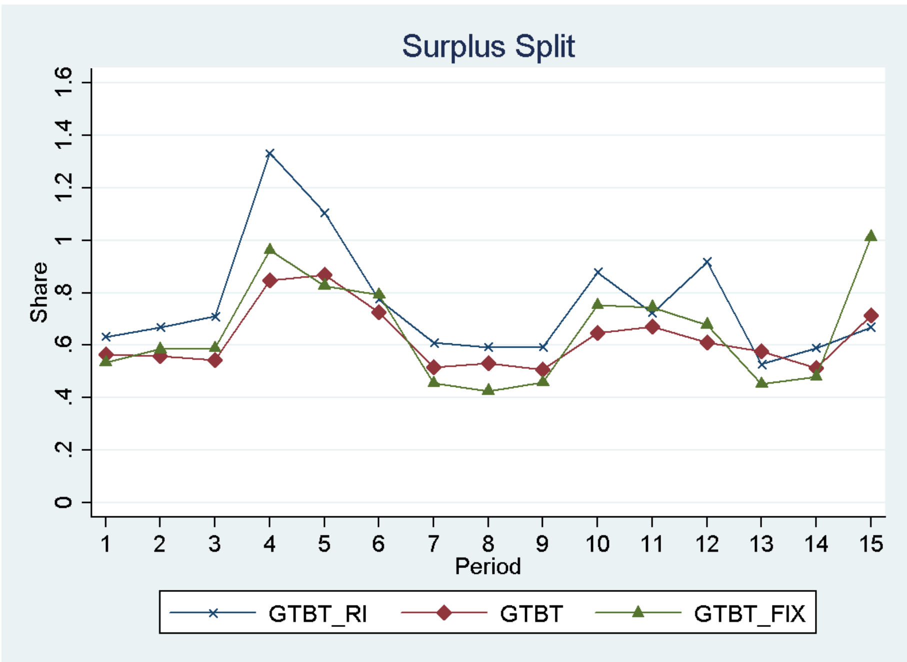
As far as the surplus split is concerned, we find an almost identical pattern across all three treatments. During the first BT-periods, workers reap the full surplus with nothing being left on the table for firms. In GTBT_RI, profits for firms are even negative on average such that workers obtain more than 100% of the surplus. In the second BT-phase, we observe a lower share that goes to the worker. Overall, the surplus split of 60% to the

[50] See appendix A1.2 for details regarding our efficiency measure.

worker and 40 % to the firm is a very persistent pattern across all three control treatments. An overview is provided in Figure A1.3.

**Result 1.14** *The absence of reputation considerably harms efficiency, particularly in BT-periods. Markets in GTBT_FIX realize* 60% *of the possible surplus, which is even higher than that of GTBT markets. Surplus sharing is identical across GTBT, GTBT_RI and GTBT_FIX.*

Figure A1.3: Surplus Split in GTBT, GTBT_RI and GTBT_FIX




Share of the surplus that accrues to the worker in treatments GTBT, GTBT_RI and GTBT_FIX

## Appendix A1.5: Tables

Table A1.4: Panel Tobit Regressions on Effort

| | Main treatments, Panel Tobit, RE | |
|---|---|---|
| | (A1) | (A2) |
| GTBT-dummy | 1.343*** | 0.279 |
| | (0.476) | (0.499) |
| FBT-dummy | -0.256 | -0.343 |
| | (0.573) | (0.558) |
| Wage | 0.162*** | 0.163*** |
| | (0.010) | (0.010) |
| Private | 1.208*** | 1.252*** |
| | (0.314) | (0.309) |
| Tenure | 0.275*** | 0.259*** |
| | (0.036) | (0.035) |
| Wage*FBT | 0.122*** | 0.124*** |
| | (0.018) | (0.018) |
| Wage*GTBT | -0.008 | 0.014 |
| | (0.013) | (0.013) |
| Private*FBT | -0.079 | -0.010 |
| | (0.430) | (0.423) |
| Private*GTBT | -0.004 | -0.150 |
| | (0.402) | (0.397) |
| BT-dummy | | 1.467*** |
| × GTBT-dummy | | (0.261) |
| Period | 0.367*** | 0.269*** |
| | (0.069) | (0.071) |
| Period$^2$ -0.029*** | -0.023*** | |
| | (0.004) | (0.004) |
| Constant | -2.338*** | -2.010** |
| | (0.437) | (0.411) |
| N | 1242 | 1242 |
| | Wald-$\chi^2(11)$=1414.91 | Wald-$\chi^2(12)$=1529.10 |
| | Prob= 0.00 | Prob=0.00 |

Panel Tobit random effects (RE) regressions on the level of effort. Standard errors are given in brackets, clustering on the market level. *** indicates significance at the 1 % level, ** at the 5 % and * at the 10 % level, respectively.

## Appendix A1.6: Figures

Figure A1.4: Actual, Desired and Expected Effort Levels in GTBT

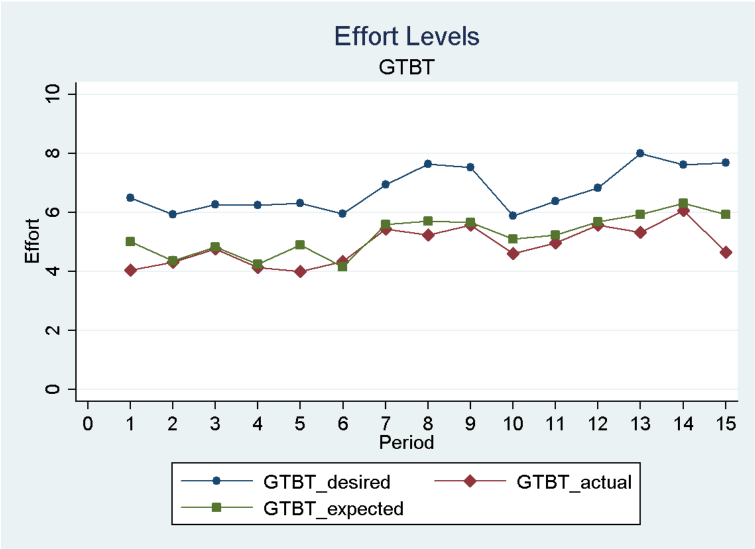



Figure A1.5: Percentage of privately concluded Contracts in FBT, GTBT and FGT

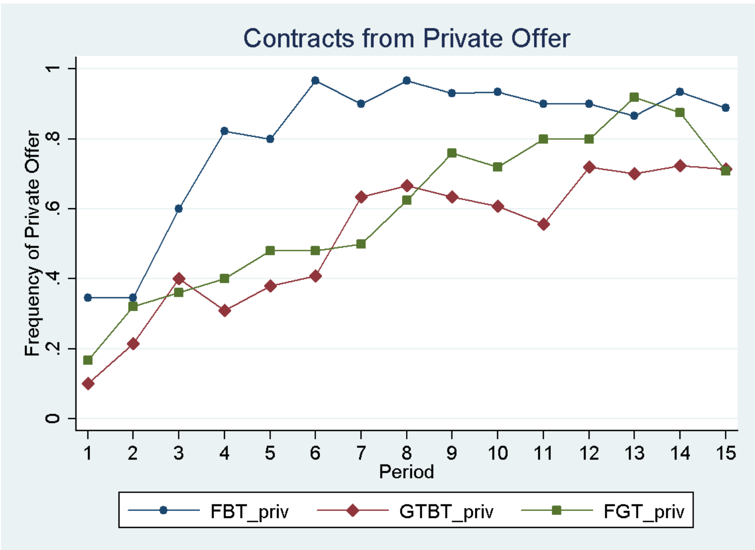


Figure A1.6: Per Period Profits in FBT, GTBT, and FGT

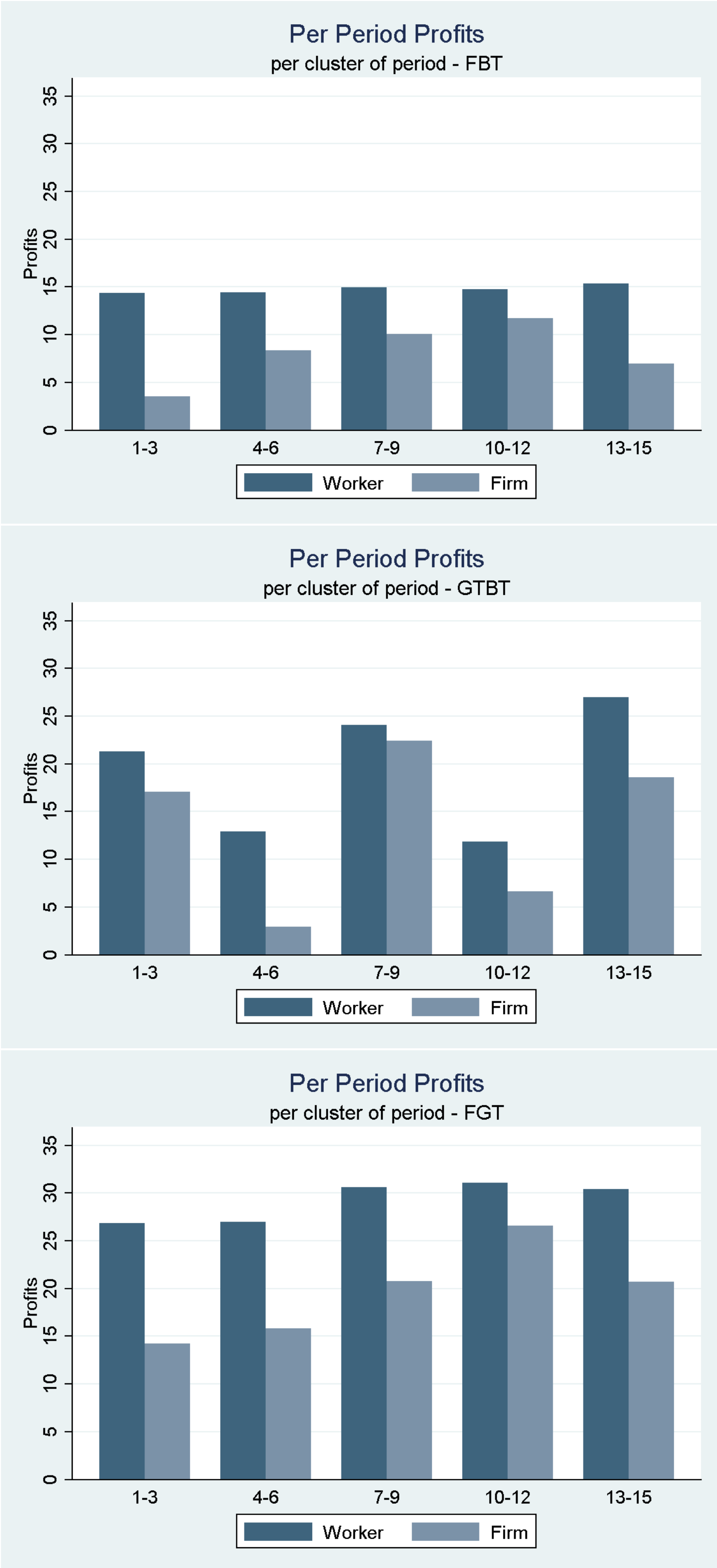


Figure A1.7: Employment Levels in GTBT and GTBT2

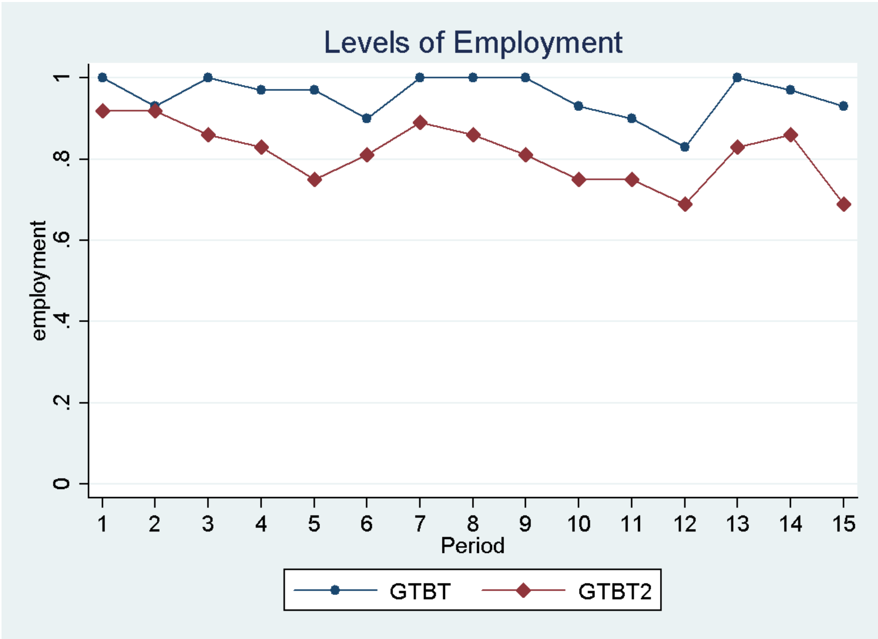



Figure A1.8: Employment Levels in GTBT_FIX and GTBT2_FIX

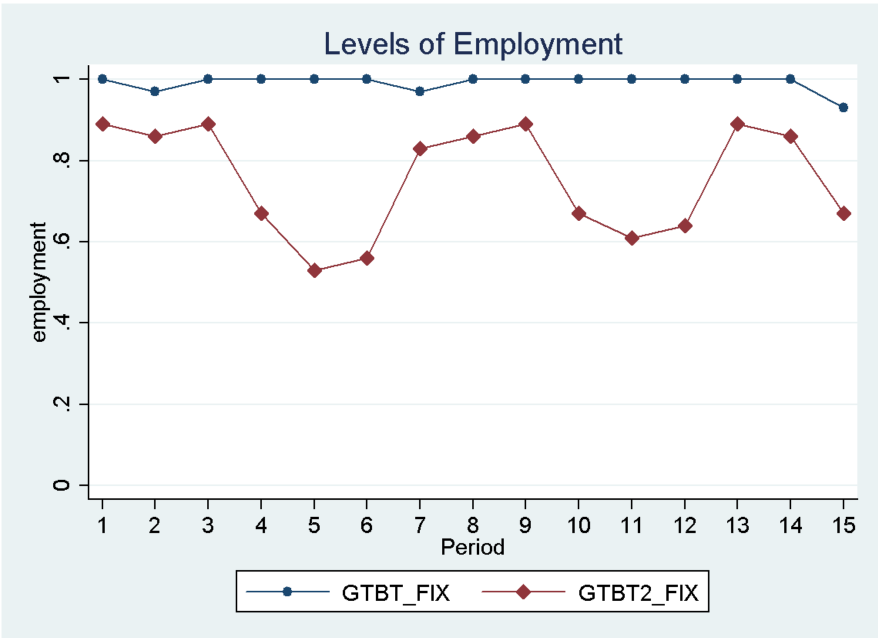


Figure A1.9: Average Wage and Effort Levels in the Two-Worker Setting

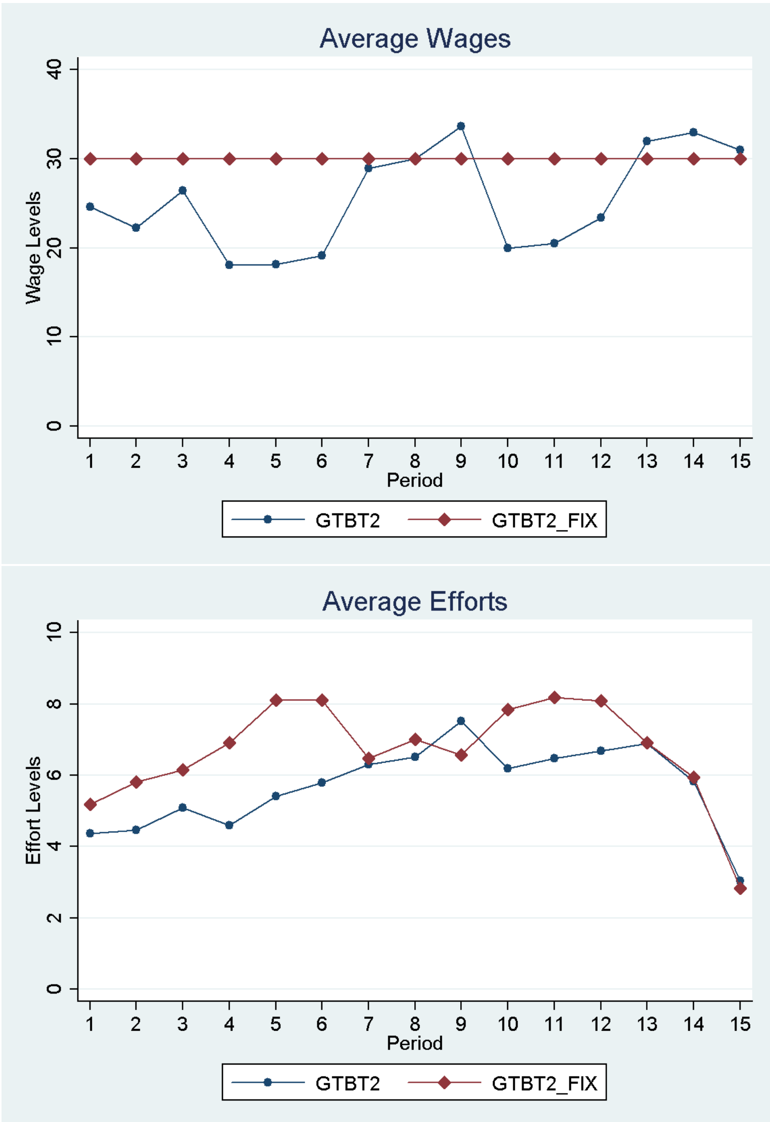


Figure A1.10: Average Wage and Effort Levels in GTBT and GTBT2


Figure A1.11: Share of the Surplus to the Worker in GTBT2 and GTBT2_FIX

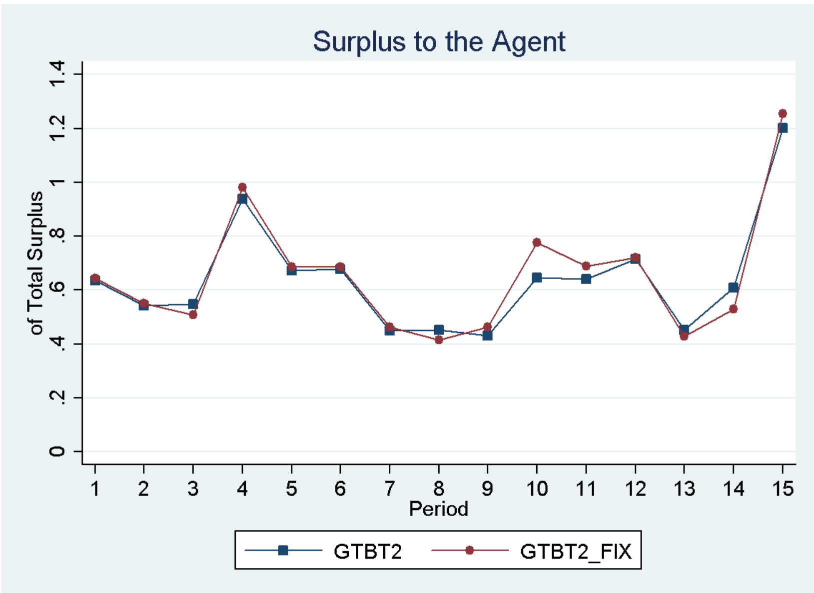
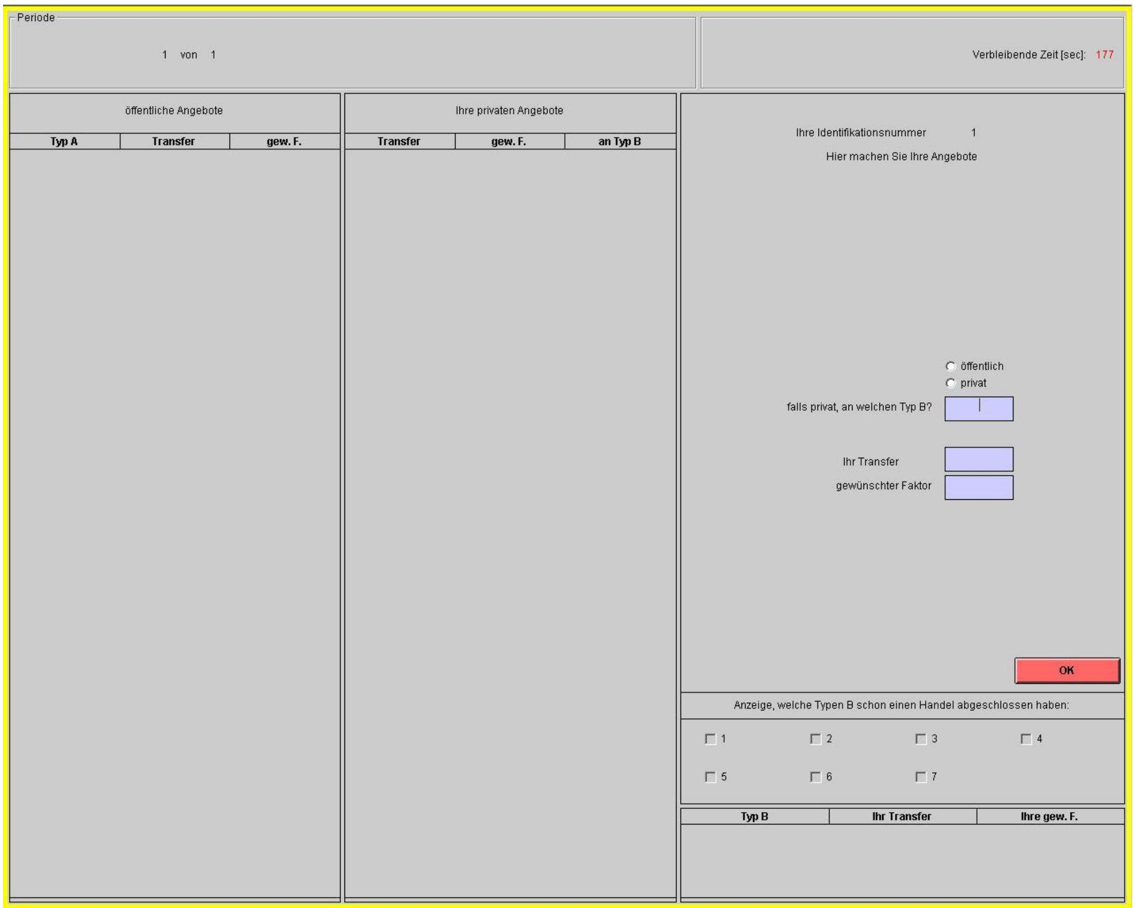


Figure A1.12: Screenshot: Trading Phase (Firm)




Example of the screen of a buyer (firm) during the trading phase with all public offers (left column), private offers (middle column) and the contract offer space (right column)

Figure A1.13: Screenshot: Trading Phase (Worker)


Example of the screen of a seller (worker) during the trading phase with all public offers (left column) and private offers (middle column)

# Appendix A1.7: Instructions (translated from German)

FGT/GTBT/FBT

The objective of this experiment is the analysis of economic decisions. You and all other participants will be taking decisions during the experiment. You will be earning money. Your payoff will depend on both your own decisions as well as the decisions of other participants according to the rules on the following pages.

## Types of participants

There are **two types of participants**: Participant A and Participant B. You will be assigned a role randomly. At the beginning of the experiment, you will be told your role (A or B) on the first screen after the start of the experiment. You will remain in the same role for the whole duration of the experiment. Every participant obtains an identification number. This number also remains unchanged throughout the whole experiment.

## Earnings

At the beginning of the experiment, you obtain an amount of **4 Euro**. In the course of the experiment, you will be earning more money by gaining **points**. All earned points will be converted into Euros at the following conversion rate:

**1 Point = 0.10 Euro (10 Eurocent)**

i.e. **1 Euro = 10 Points**

**At the end of the experiment, you are paid the amount you have earned during the experiment plus the 4 Euro starting stack privately and in cash.**

## Duration

The whole experiment takes about 2 hours. The experiment is divided into 15 periods. In every period you have to take decisions that you enter into the computer in front of you.

## Documentation

You also find a documentation sheet at your desk. Please enter your identification number at the beginning of the experiment into the upper right corner. In every period, you will be entering certain pieces of information into this sheet (see below). Please note that every period has a single line.

## Anonymity

You will not be learning neither during nor after the experiment the identity of other participants. Interaction only takes place via the identification numbers. The other participants do not learn neither during nor after the experiment your role and how much you have earned.

FGT/GTBT/FBT

**Communication is strictly prohibited to during the whole experiment!** We additionally advise you only to make use of those functions at the computer which are relevant for the experiment. Communication or playing with the computer leads to exclusion from the experiment.

# Brief Overview over the Course of events in the Experiment

In every period of the experiment, a participant of type A can conclude a trade with a participant of type B. Type B realizes a gain through the trade, if he obtains a transfer that exceeds his costs. Type A realizes a gain through the trade if earns more through the factor that the transfer costs to him. The level of the costs for type B and the revenues of type A depend on the factor, which is determined by type B.

The whole experiment has 15 identical periods. The course of events in every single period is organized as follows:

1. Every period starts with a **trading phase** which takes **3 minutes**. During this phase type A participants can make offers which can be accepted by type B participants.

   An offer consists of **three things** that have to be specified:

   - the **transfer** type A is offering,
   - the **factor** he is desiring from type B
   - and finally to which type B the offered is directed. Type A participants can make two sorts of offers: private and public offers. Private offers are only directed to one single type B participant and can only be accepted by this particular type B. Public offers are directed towards all type B participants and can be accepted by every type B.

   Type A participants can make as many offers as they want in every period. A standing offer can be accepted at any time. **Every type B can agree to only one single trade per period, i.e. accepting one offer. Type A participants also can only agree upon one trade with one type B participant.** Since there are 5 type A participants and 7 type B participants, there are some type B participants that cannot conclude a trade in every period.

2. Next to the trading phase, all type B participants who have agreed upon a trade must decide upon the factor they are delivering to the type A participant. **Type B is not obliged to deliver the factor type A is desiring from him**. When all type B participants have made their factor choices, the payoffs from this period are determined. After that, the next period starts.

As a reference to reality, you can think of the experiment as a labor market: type A is the employer who is offering work contracts. The transfer constitutes the wage and the factor is the desired performance (hours worked, etc.). Every employer can only hire one employee. Upon his employment (i.e. the acceptance of the contract), the employee decides whether to deliver the desired performance.

# Detailed Procedure of the Experiment

On your market, there are 5 type A and 7 type B participants. Overall, there are two parallel markets in this room with 24 participants. Throughout the whole experiment you remain in the same role and interact on the same market. During the experiment, you enter your decisions into the computer in front of you. Subsequently, it is explained in detail how you can make your decisions in every period.

page 2 of 9

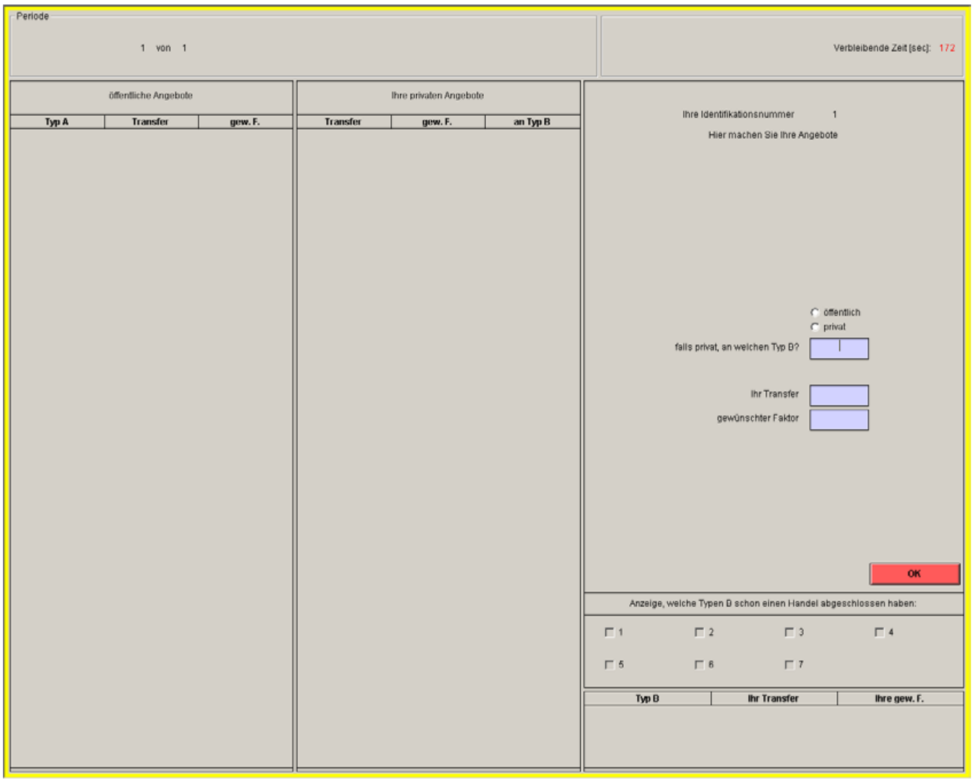
FGT/GTBT/FBT

**1. The trading phase**

***TYPE A***

In the trading phase, type A participants see the following screen:

In the top left corner you can see the number of the period you are currently in. In the top right corner you are shown the remaining time of the trading period in seconds. **The trading phase lasts for 3 minutes, i.e. 180 seconds.** When time is up, the trading phase is over such that you cannot make any offers anymore this period nor accept them.




**Picture 1**

As soon as you see the above screen, a trading phase has started. Type A participants can now make offers. To do so, they have to enter three things at the right of the screen:

a) First, they have to determine whether to make a public or a private offer:

- **Public Offers**
  Public Offers are communicated to all market participants. All type B participants of your market see all public offers of their respective market on their screens. A public offer can hence be accepted by **every type B.** Also all type A participants see all public offers of all other type A participants from their market on their screens.
  To make a public offer, first click on the icon “public”.

page 3 of 9

FGT/GTBT/FBT

- **Private Offers**
  Private offers are only directed towards **one single type B**. Only this particular type B participant can accept the offer. All other participants do not find out anything about this offer.
  To make a private offer, first click on the icon “private”. After that, you have to enter **to which type B** the offer is directed. All 7 type B participants on your market have an identification number (type B1, type B2, …, type B7). All type B participants keep this identification number throughout the whole duration of the experiment. To direct an offer to a particular type B participant, you have to enter the identification number of this type B (e.g. “4” for type B4).

b) Once it has been determined to whom the offer is directed, a **transfer** has to be fixed. This transfer has to be entered into the field “Your Transfer”. The offer must not be smaller than 0 and not exceed 100:

$$\mathbf{0 \leq Offer \leq 100}$$

c) Finally, you have to state which **factor** you desire. This has to be entered into the field “desired factor”. The **desired factor** has to be an integer and must not be smaller than 1 and not exceed 10:

$$\mathbf{1 \leq Desired\ Factor \leq 10}$$

Once an offer has been fully specified, you have to click on the “OK”-button to submit the offer. As long as the “OK”-button has not been clicked, an offer can be modified. After clicking on the “OK”-button, the offer is displayed on screens of the type B participants to whom it has been directed.

On the left hand side of type A’s screen, you can find the headline “Public Offers”. All public offers of the current trading period are displayed here. You will be seeing both your own public offers as well as those submitted by other type A participants. You can identify what participant made the offer, which transfer he is offering and which factor he is desiring. All type A participants also have an identification number that is constant for the whole experiment (type A1, type A2, …, type A5).

In the middle of the screen below the headline “Your private Offers” you are given the private offers that you submitted in the current trading phase. You can see to what type B participants you submitted an offer, what transfer you were offering and what effort you are desiring.

**Every type A can submit as many public and private offers in every period as he wants to.** Every submitted offer can be accepted at any time during the trading phase. A submitted offer cannot be withdrawn.

**Every type A can conclude only one contract per period.** As soon as an offer has been accepted, the respective type A participant gets to know which of his offers has been accepted by which type B participant. In the bottom right corner of the screen appears the identification number of the type B who has accepted one offer, the transfer and the desired factor. Since every type A can only conclude only one contract, all other offers of this type A are automatically deleted upon acceptance of one of the outstanding offers. No further offers can be made this period.

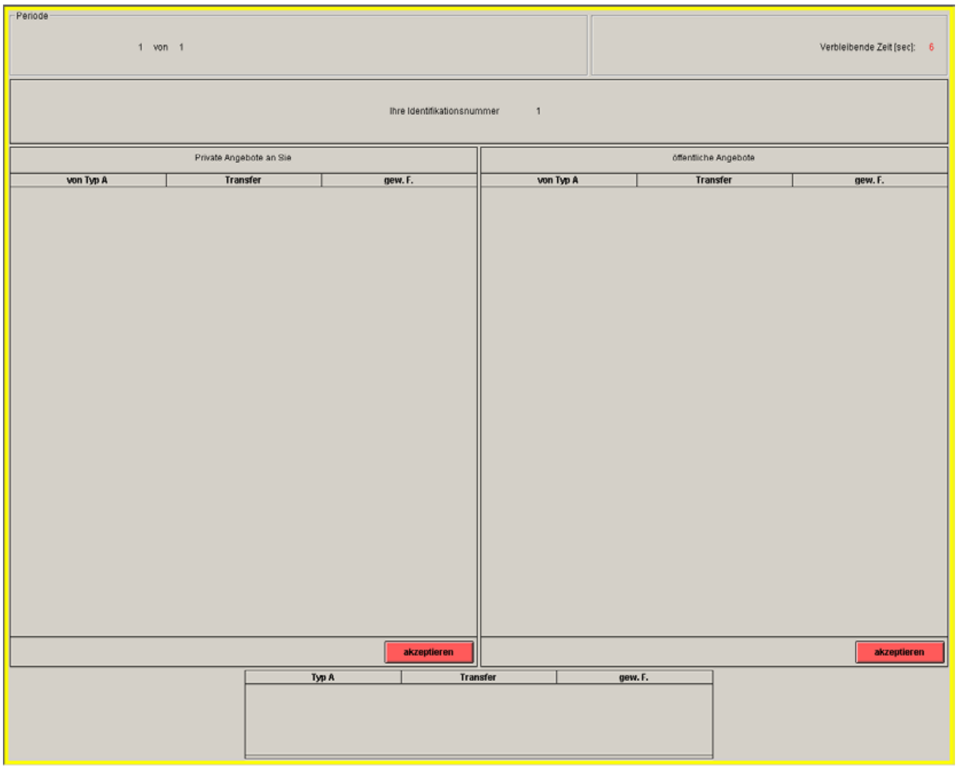
**Every type B can conclude only one contract per period.** The type A participants are continuously informed about which type B participants have **not** yet concluded a contract. Below the headline “Information which type B participants have already concluded a contract” you can find 7 fields. If one type B has accepted an offer, the white box in front of his identification number is ticked. It is not possible to make private offers to type B participants who have already concluded contract.

page 4 of 9

FGT/GTBT/FBT

***TYPE B***

In the trading phase, type B participants see the following screen:




**Picture 2**

In the top left corner you can see the number of the period you are currently in. In the top right corner you are shown the remaining time of the trading period in seconds. **The trading phase lasts for 3 minutes, i.e. 180 seconds.** When time is up, the trading phase is over such that you cannot make any offers anymore this period nor accept them. As soon as you see the above screen, a trading phase has started. The type B participants now can accept offers that have been directed to them by type A participants.

- **Private Offers to You**
  If you obtained private offers, these offers appear on the left hand side of the screen, below the headline "Private Offer to You". The offer contains The following pieces of information: the identification number of the respective type A who submitted the offer, the transfer and the desired factor. In order to accept a private offer, you have to mark the entry with the corresponding offer by clicking on it. To definitively accept the offer, you have to click on the "Accept" button at the bottom of the table. As long as the "Accept" button has not been clicked, you can still modify your choice.

- **Public Offers**
  All public offers appear on the right hand side of the screen below the headline "Public Offers". These offers have to be accepted in the same way as the private offers described above. Please take care to click on the right "Accept" button when accepting an offer.

page 5 of 9

FGT/GTBT/FBT

As soon as an “Accept” button has been clicked, the accepted offer appears in the bottom line of the screen. As soon as 5 type B participants have concluded a trade this period or the three minutes have expired, the trading phase is over. No type A is obliged to submit an offer; no type B is obliged to accept an offer.

**2. Determining the Factor**

After the trading phase, all type B participants who have concluded a trade have to determine their actual factor. **The desired factor stipulated in the offer is not binding**. Type B can choose exactly the desired factor or a higher or a lower factor. Types B have to enter their actual factor on the following screen:


**Picture 3**

To choose the actual factor, you enter the value of the factor into the field “Choose your actual factor” and click the “OK” button. As long as the “OK” button has not been clicked, you can modify your choice. The factor has to be an integer between 1 and 10:

$$\mathbf{1 \leq Actual\ Factor \leq 10}$$

During the time when type B enters the actual factor, type A enters on a separate screen, what actual factor he expects and how certain this assessment is.

page 6 of 9

FGT/GTBT/FBT

# How are incomes calculated?

## Income *Type A*:

- If **no trade has been concluded in a trading phase**, type A gets a payoff of **0 points** this period.
- If **one offer has been accepted**, your income depends on the offered transfer and the actual factor chosen by the type B participant who accepted your offer. Your income is calculated as follows:

**Income Type A = K*(Factor) – (Transfer)**

[**GTBT**: The value of K depends on the period you are currently in. The following table indicates the value of K dependent on the period (the number of the current period is shown in the first line of every screen):]

| **Period** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** | **12** | **13** | **14** | **15** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| K | 10 | 10 | 10 | 5 | 5 | 5 | 10 | 10 | 10 | 5 | 5 | 5 | 10 | 10 | 10 |

[**FGT/FBT**: The value of K is always 10/5 across all periods.] From the formula above it is clear that the income is higher, the higher the actual factor chosen by type B. At the same time, the income is higher, the lower the transfer that was offered.

## Income *Type B*:

- If **no trade has been concluded in a trading phase**, type B gets a payoff of **6 points** this period.
- If **one offer has been accepted**, your income depends on the transfer minus the costs of the factor that you have to bear. Your income is calculated as follows:

**Income Type B = Transfer – Costs of the Factor**

The income of type B is higher, the higher the transfer and the lower the actually chosen factor is. The costs of the actual factor are higher, the higher the actually chosen factor is. The costs of each factor are given by the following table:

| **Factor** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** |
|---|---|---|---|---|---|---|---|---|---|---|
| Costs of Factor | 0 | 1 | 2 | 4 | 6 | 8 | 10 | 12 | 15 | 18 |

**All incomes of type A and type B participants are calculated in the same way. Every type A can calculate the income of the type B participant who concluded a trade with him. Every type B can calculate the income of the type A participant who concluded a trade with him. In addition**, both type A and type B learn the identification of their trading partner in every period.

page 7 of 9

FGT/GTBT/FBT

Please be aware of the fact that both type A and type B participants can make **losses** in every period. You have pay them through your starting stack or from gains accumulated in other periods. A (very unlikely) overall loss can either be paid in cash or through student work at the laboratory ( 5 EUR per half an hour).

You are given your income and the income of your trading partner on an **"Income" screen:**


**Picture 4**

This screen displays the following pieces of information:

- number of identification of your trading partner
- the offered transfer
- the desired factor
- the actual factor
- the income of your trading partner
- your income this period

Please enter all the information on the joint documentation sheet. After the "Income" screen, a period is finished. The next period starts with a new trading phase. Once you are finished with entering the information on the documentation sheet, please click on the "Continue" button.

The experiment does not start until all participants are completely familiar with the course of events and the calculations. To assure this, we kindly ask you to solve the exercises that you find on the next two pages. Additionally, we will conduct **two test trading periods** to make you familiar to the computer program. These test trading phases do not count towards the final results and are not paid out. After the two test periods, the actual experiment starts consisting of 15 periods.

page 8 of 9

FGT/GTBT/FBT

# Exercises

Please solve the following questions by indicating the way you obtained your result. Once you have finished the exercises, pleases raise your hand and the experimenter will control your results. Wrong answers do not have any consequences whatsoever for you.

**Exercise 1**
Type A has not submitted an offer in a trading phase. What is his income this period?

Income *Type A* =

**Exercise 2**
Type B did not accept any offer in a trading phase. What is his income this period?

Income *Type B* =

**Exercise 3:**
In the first period, an offer with a transfer of 30 and a desired factor of 9 is accepted. Type B chooses an actual factor of 9.

Income *Type A* =
Income *Type B* =

**Exercise 4:**
In period 5, an offer with a transfer of 30 and a desired factor of 9 is accepted. Type B chooses an actual factor of 6.

Income *Type A* =
Income *Type B* =

**Exercise 5:**
In period 9, an offer with a transfer of 10 and a desired factor of 2 is accepted. Type B chooses an actual factor of 5.

Income *Type A* =
Income *Type B* =

**Exercise 6:**
In period 12, an offer with a transfer of 20 and a desired factor of 4 is accepted. Type B chooses an actual factor of 5.

Income *Type A* =
Income *Type B* =

**Exercise 7:**
In the last trading phase, type A has made several offers. Neither has been accepted. What is his income in this period?

Income *Type A* =

When you have finished the exercises, we recommend to you to take a closer look at the exercises and the solutions again. Subsequently, please think about what decisions you want to take in the experiment.

page 9 of 9

# Chapter 2

# Worker Characteristics, Contracting, and Wage Differentials: Evidence from a Gift-exchange Experiment[0]

## 2.1 Motivation

Firms spend substantial resources in their hiring procedures to select the "best" candidate for a job. In particular, an increasing fraction of firms uses both ability and personality tests in their hiring processes, see, e.g., Autor and Scarborough (2008). While the rationale for selecting the most "able" candidate is obvious, recent work has highlighted that personality tests can generate additional information, too, so as to improve the chance to hire a suitable candidate for a given position. One important piece of additional information about a potential candidate are her social preferences. In particular in the presence of moral hazard, it is valuable for firms to have access to employees that can be motivated by "social incentives" via gift exchange, see for example Englmaier and Leider (2011).

Despite the importance of information acquisition in real world contracting, the understanding of the impact of the availability of information on the terms of a contract is surprisingly limited and in particular empirical evidence on the issue is scarce. It is

[0]This chapter is based on joint work with Florian Englmaier and Joachim Winter.

precisely this gap that we want to fill: We provide controlled evidence from the laboratory on how specific pieces of information about a contracting partner are used and how they interact with the contracting behavior between a principal and an agent in a gift-exchange situation.

**Research Question** When making employment and contracting decisions, firms naturally desire to minimize the risk of hiring an unsuitable candidate. They try to learn about the qualification of a candidate, his education, his family background, etc. before offering an employment contract. As a necessary simplification of reality for our experiment, we concentrate on two dimensions of information about a candidate that we regard as essential on real world labor markets. First, information about what we call *productivity* henceforth is meant to capture an objective assessment whether the candidate is good at the job he is supposed to accomplish. Second, information about what we call *trustworthiness* henceforth is supposed to encompass all social and reciprocal preferences by the candidate. We consider these two measures of a worker's traits the most relevant skills in our setting. Hence, we expect that information about these skills matters for firms. In a situation characterized by moral hazard, we expect both elements to play an important part in the effort decision of the agent and hence for the outcome for the principal: Controlling for social preferences, an agent who is more productive at accomplishing a certain task will produce a higher outcome for the principal. Similarly, for given productivity, a reciprocal agent will put in more effort in response to a "generous" wage offer leading to a higher outcome for the principal.

In this chapter, we concentrate on a contracting situation where information about a worker stems from sources external to the firm-worker relationship. In contrast to e.g. Brown et al. (2004) or Bartling et al. (2011), we abstract from information about the worker that arises endogenously in a repeated relationship and can be used for firms to adapt contracts over time. We focus on the trade-off between two pieces of information and their impact on contracting behavior by both principals and agents in a one-shot interaction. The main research question we have in mind is to explore how and if at all these two pieces of information are conditioned upon when writing contracts and to what extent they can be used to predict behavior. Eventually, we evaluate how the presence of certain skill sets and available information about them shapes labor market outcomes under moral hazard. The high degree of control makes the laboratory an ideal setting to address these questions.

**Design** Our experiment consists of two parts, which are presented sequentially to subjects such that they do not know what will be the content of the next part. Subjects know in advance, however, that decisions in earlier parts may have an impact on later parts. In the first part, agents work on a real effort task under a piece rate contract. We use their score in this piece rate task as our measure of productivity. Subsequently, agents are presented with a binary, neutrally framed, trust game which we use to proxy for social and reciprocal concerns. In the second part, half of the players are randomly assigned to be employers and the other half to be employees. Subjects play a gift exchange game where the employer first offers the employee a flat wage and the employee thereafter performs the real effort task from the first part under standard gift-exchange incentives. Before making their wage offers, principals are presented with the information about workers from the elicitation tasks. The amount and the degree of information provided to firms is our treatment variable. In our main treatment, employers are presented the *productivity* and the *trustworthiness* measure in a *binary* way (hereafter treatment **PTB**) before submitting wage offers. To control for strategic behavior in the elicitation phases, we run two control treatments where only one piece of information is made available to firms. In treatment "Productivity"(hereafter **P**) they are only presented the productivity measure, and in treatment "Trustworthiness" (hereafter **T**) they are only presented the trustworthiness measure. By comparing our treatments P and T to the PTB treatment, we can control if the information revelation in the final phase distorts the elicited measures in phases 1 and 2.[1] To assess the impact of the mode of information presentation on behavior, we conduct a further control treatment (hereafter **PT**) where we vary the precision of the productivity information.

[1]In P (T) it is communicated to subjects that in the second part only information from the elicitation of productivity (trustworthiness) possibly made available in later parts of the experiment, whereas in PTB this applies for both measures.

**Results** We have four main findings. 1) Contracts offered by principals systematically vary with the information they have about the agent. Principals tailor their wage offers to employee types, offering more generous contracts to more productive and more trustworthy subjects. The wage premium for "better" agents is higher with respect to productivity than for trustworthiness. 2) We find a positive wage-effort relation for all worker types, i.e., gift-exchange is robust across all types of workers, but is most effective with trustworthy workers. 3) Worker characteristics affect firms' profit levels, such that optimal wages are made contingent on the worker type. Only for trustworthy workers the wage-effort relation is steep enough that an increase in the wage induces a significant profit increase for firms. 4) Subtle differences in the information presentation to firms induce an endogenously different distribution of wage offers. If given access to the precise level of worker productivity instead of the binary measure from PTB, firms are too focused on productivity than on trustworthiness.

**Related Literature** An extensive experimental literature documents incentives and behavior in gift exchange games, see e.g. Akerlof (1982), Fehr et al. (1993), Fehr et al. (1997). This protocol has proven to be a valuable paradigm that captures incentives on real world labor markets in the laboratory. As a major finding of this literature, preferences for fairness and reciprocity serve as a powerful source of motivation to overcome the informational asymmetry between principals and agents on labor markets[2]. These laboratory studies have also been validated in the field; see e.g. Falk (2007) or Bellemare and Shearer (2009). It is now also widely acknowledged that social preferences like reciprocity or inequity aversion potentially do not only shape market outcomes or the result of bilateral bargaining, but have an important effect on the design of optimal incentive schemes as well, see Englmaier and Wambach (2010) for a theoretical treatment and Fehr et al. (2007) for empirical evidence. One additional important empirical finding from both field and laboratory data (e.g. Dohmen et al. (2009)) is that there is substantial heterogeneity with respect to the prevalence of reciprocal inclinations and social preferences among the population, see Fehr and Schmidt (1999) or Fischbacher et al. (2001).

Recent theoretical and experimental work suggests that there are complementarities from

[2]For references see Fehr and Gächter (2000), Fehr and Schmidt (2003) or Fehr and Falk (2008).

matching incentive structures to worker types; e.g. Ichniowski et al. (1997); Englmaier and Leider (2011) or Bartling et al. (2011). However, little work has been done that tests how worker characteristics (amongst them social preferences, if measured) relate to behavior across games, see Englmaier and Leider (2009) for an exception, and, more to our point, how worker types interact with incentives. One recent important exception is a paper by Cabrales et al. (2010) who design an experiment where in the first phase, all subjects choose a payoff vector and play a self-chosen effort game. From these choices their preference parameters in terms of both outcome preferences and reciprocal inclination are estimated, assuming preferences à la Charness and Rabin (2002). In the second phase, it is documented that these estimated preferences predict behavior in a gift exchange game conditional on contract offers. Moreover, contract offers vary systematically with estimated preferences of principals. However, Cabrales et al. (2010) do neither use a real effort task (and hence they do not elicit measures about productivity), nor is information about workers presented to the principals prior to their contract offers. We consider this last feature essential for our understanding of the functioning of real world labor markets. Most closely related to our study, Dohmen and Falk (2011) design a laboratory experiment where they elicit worker characteristics to explain sorting behavior of subjects into variable or fixed-payment incentive schemes. They find strong evidence for worker sorting along multiple dimensions, but claim that ”many of the discussed worker attributes are typically unobservable in the hiring process“ (p.558). While this is certainly the case for some attributes that are difficult to observe, we argue that proxies for the most important skills of a worker are well available to firms before hiring a worker, e.g. in the form of a curriculum vitae or the results from hiring tests. We therefore complement their analysis by showing how the presence of information about these attributes interacts with incentives on the labor market.[3]

This chapter adds to another strand of literature that assesses the effects of the availability of potentially costly information about an interaction partner on subsequent strategic behavior. Kurzban and DeScioli (2008) show that subjects in public goods game buy

[3] A somewhat related literature on cognitive and non-cognitive skills has mainly focused on the relation between these two skill sets and their interdependence, see e.g. Heckman et al. (2006), Borghans et al. (2008b), and their relationship to balor market outcomes, see e.g. Murnane et al. (1995), Borghans et al. (2008a), Heineck and Anger (2010).

information about behavior of others in previous round to adjust their behavior. More recently, Eckel and Petrie (2011) give subjects the possibility to purchase a picture of the interaction partner in a trust game before deciding about trust and trustworthiness. They find that there is informational value in a counterpart's face since many subjects do purchase the picture at nonzero costs. Evidence from the field suggests that firms use the acquired or available information about workers and applicants for screening purposes and to tailor incentive schemes in the presence of moral hazard, see e.g. Ichniowski et al. (1997); Huang and Cappelli (2010). To the best of our knowledge, we are the first to focus on the pure effects of the availability of information about interaction partners on contracting outcomes in a one-shot moral hazard situation. A controlled laboratory study allows us to exogenously vary the information structure and eliminate effects of worker competition.

The remainder of the chapter is structured as follows. Section 2.2 presents the experimental design. In section 2.3 we lay out our hypotheses and section 2.4 presents the results of the experiments. Section 2.5 discusses our findings and concludes.

## 2.2 Experimental Design

The experiment consists of two parts, which are presented sequentially to subjects such that they do not know what will be the content of the next part. Subjects knew in advance, however, that their decisions earlier on may potentially be disclosed to other subjects later on of the experiment. Overall, we ran 14 sessions with a total of 336 subjects in June and July 2011 at the MELESSA laboratory at LMU Munich. The subjects were invited via ORSEE (Greiner, 2004), and the experiment was implemented with zTree (Fischbacher, 2007). Subjects earned experimental points (EP) during the experiment. The exchange rate from EP to Euros was 1EP = 0.0125 €. The experiment lasted about 60 minutes and subjects earned on average 11.8 €.

### 2.2.1 Elicitation of Productivity and Trustworthiness

In a first part of the experiment, measures of both productivity and trustworthiness are elicited from all subjects. We proxy productivity with a measure from a real effort task that consists in matching words and a four-digit code from a list.[4] Subjects perform this task for 90 seconds and are paid a piece-rate per correct answer of 10 EP that is paid out at the end of the experiment.[5] There is no particular training required for fulfilling this task and we assume that all subjects put in full effort under the piece rate scheme such that our measure of productivity is as closely related to underlying ability as possible. In the remainder of the analysis, we therefore refer to ”productivity“ as the number of *correct answers* in this task. The corresponding payoff from part one is calculated as follows according to the number of correct answers:

$$10 \text{ EP} * (\#\text{correct answers})$$

Subjects are presented three screens of 30 seconds each one after another, i.e., a total of 90 seconds, with randomly generated words and codes for every new screen. We conjecture that intrinsic costs for the task are linear over the interval of 90 seconds, i.e., there are no effects from fatigue or boredom. The resolution of the number of correct answers is given to subjects only at the end of the experiment.

Furthermore, subjects play a standard binary trust game in neutral framing to provide a measure of trust and trustworthiness at the individual level.[6] We make use of the strategy method to get data on both trusting behavior and trustworthiness. Subjects take both decisions for both roles and at the end of the experiment they are matched with another subject, roles are randomly determined and payoffs are realized according to the decisions taken in the respective roles. Behavior in the trust game can be seen as indicative whether individual preferences are characterized by high or low levels of trust and trustworthiness

[4]A screenshot of the experimental screen of the coding task can be found in Figure A2.1 in appendix A2.3. We conduct a trial period before the elicitation task to familiarize subjects with the computer programme.

[5]To discourage guessing there is a penalty for every wrong answer of 10 EP, which is known to subjects. Our measure of productivity therefore consists in the number of correctly matched codes after subtracting all wrong answers.

[6]See Figure A2.2 in appendix A2.3 for the precise amounts used in the trust game.

when engaging in an interaction with another person. We focus on the trustworthiness of subjects as this appears as a more relevant proxy of social and reciprocal concerns in the gift exchange game than the initial trusting decision.

Importantly, there was no feedback given to subjects about the elicitation procedures (and the resulting payoffs) until the very end of the experiment such that subjects' subsequent behavior in the experiment was not affected. We also elicit all subjects' detailed expectations about productivity and trustworthiness in the population. Since they turn out not to matter for the subsequent experiment, we relegate the description of the experimental protocol and the results on expectations to appendix A2.1.

### 2.2.2 Gift-Exchange Game

In the second phase of the experiment, an experimental gift exchange game is implemented in which the task to be fulfilled is identical to the real effort task in the first part of the experiment. Subjects are randomly allocated to be either a firm or a worker. Overall, there are 12 workers and 12 firms per session. We employ the strategy method, i.e., firms have to submit a binding wage offer for each of the 12 workers such that we obtain the full wage profile firms are submitting for all workers in their market.[7] After all wage offers have been submitted, every worker is matched randomly to a single firm, i.e., every firm hires only one worker. Workers learn only the wage offer that their matched firm has determined for them before they start working for their firm. There is no possibility for workers to be influenced by offers that the firm has submitted for other workers or by offers that other firms have submitted to them. Subsequently, workers perform the same real effort task from the first part for 90 seconds. The interaction is one-shot to preclude any effects of repetition over time and to focus in the cleanest possible way on the effects of information on contracting behavior.

Workers' performance then determines the payout to the firm according to the following formula

$$\text{firm payoff} = 10 \text{ EP} \cdot (\#\text{correct answers}) - \text{wage}$$

[7] A screenshot of the wage setting screen can be found in Figure A2.3 in the appendix A2.3.

where # correct answers is given by all solved matches minus all wrong matches. Workers are paid their predetermined fixed wage and have non-monetary costs of effort from solving the task:

$$\text{worker payoff} = \text{wage}$$

To avoid that agents can ruin firms by deliberately giving wrong answers we impose a lower limit for the payoff to the firm from the task at 0. This does not preclude firms from making losses if the wage exceeds the revenues generated by their worker. Losses had to be paid from earnings in other parts of the experiment. Given the nature of the task and the fact that new words and codes are randomly generated for every screen, there should be virtually no learning possibilities from doing the task a second time. After the real effort task is completed, there is feedback about the number of correct answers and the payoff to the firm and the worker. Both firms and workers learn only the details from their interaction, but not from the interaction between any other firm-worker pair.

**Treatments**

Our treatment variation consists in the pieces of information elicited in the first part from the experiment that are made available to firms when submitting their wage profiles. In our main treatment **PTB**, information about *productivity* and *trustworthiness* is available in a *binary* way. Information about productivity is given to firms in the form of whether a worker has achieved a productivity score in the coding task which is higher than the mean of all subjects in the respective session, or below the mean. Information about trustworthiness is given in the form of the binary decision as trustee in the trust game, i.e., either whether a subject returned trust or not. To preclude framing effects, both pieces of information were given in a neutral way, i.e., in the trust game the actual information was labeled “left” or “right” depending on whether subjects opted for the left or the right branch of the game tree. For the productivity measure, subjects were divided into two groups labeled “blue” or “yellow” which was explained to subjects.

To control for strategic effects in the elicitation phases of our two measures, we conduct two control treatments where we make only one piece of information accessible to firms. In treatment **P**, information about *productivity* only is available and in treatment **T** information about *trustworthiness* only is available.[8]

In an additional control treatment, we explore the effects of increasing the precision of information given to subjects. In treatment **PT**, we give firms access to the exact measure of worker productivity (i.e., the number of correct answers) rather than the binary categorization into workers above or below the productivity mean as in PTB.

We conduct 4 sessions of the main PTB treatment (96 subjects), 2 sessions each (48 subjects) of the two control treatments P and T and 6 sessions (144 subjects) of the PT treatment.

## 2.3 Hypotheses

In this section, we sketch a simple agency model for workers who are heterogeneous with respect to productivity and trustworthiness (assuming stable preferences and productivity types), which is based on a simplified version of Englmaier and Leider (2011). A firm hires a worker for a fixed wage $w$ which is binding. The interaction is one shot, effort $e$ is not contractible and there are no contingent contracts. The firm relies on gift exchange to elicit performance. Exerting effort has convex costs of effort $c(e)$ for the agent. Output $\pi$ is not contractible and assumed to accrue deterministically according to $e \cdot p$ where $p$ is the worker's productivity. Firm profits are then $\pi - w = e \cdot p - w$ and the worker's utility $u(w)$ is given by

$$\begin{aligned} u(w) &= w + \eta(w - o)\,(e \cdot p - w) - c(e) \\ &= w + \eta \cdot w \cdot (e \cdot p - w) - c(e) \\ &= w + \eta \cdot w \cdot e \cdot p - \eta \cdot w^2 - c(e) \end{aligned}$$

where $\eta$ captures the worker's reciprocal inclination, and $o$ is the outside option which we normalize to 0. We abstract from explicitly modeling feelings of negative reciprocity but focus on positive reciprocity between firms and workers, i.e., $\eta \geq 0$ since in our setting as

[8]Since the P treatment serves only to exclude strategic concerns in the elicitation phases, we do not classify subjects into high or low productivity subjects as we do in PTB, but give firms full information about workers in the contracting phase.

a worker there are no possibilities for punishing the firm other than shirking ($e = 0$).

In order to elicit a positive effort response the worker has to receive a wage "gift", i.e., a wage exceeding his outside option. When the reciprocal worker receives a positive wage gift his utility increases in the firm's profit. From this, and assuming that the first order condition is necessary and sufficient for an optimal response, we can determine the worker's best response $e^*$:

$$\begin{aligned}\frac{\partial u(w)}{\partial e} &= \eta \cdot w \cdot p - c'(e) = 0 \\ c'(e^*) &= \eta \cdot w \cdot p\end{aligned}$$

The first order condition implicitly defines $e^*$ and we immediately see that $e^*$ increases in $w$, $p$ and $\eta$. The two elicited measures in our experiment are proxying $p$ through the productivity measure and $\eta$ through behavior in the trust game. As in treatment PTB proxies for $p$ and $\eta$ are available to firms, we expect according effects on offered wages.

**Hypothesis 2.1** *In treatment PTB, we expect wage offers and performance in the gift exchange game to be higher for more productive subjects and for more trustworthy subjects.*

For the gift-exchange game, we next look at predictions about the wage-effort relation from our model for workers classified in the productivity and trustworthiness dimension. To do so, we make the further assumption that $c(e) = \frac{1}{2}e^2$ and can explicitly solve for the individually rational effort level $e^*$:

$$e^* = \eta \cdot w \cdot p.$$

Now we see that $\forall \eta, p > 0$ it holds that $\frac{\partial e}{\partial w} > 0$; for a given level of $\eta$ and $p$, a higher wage induces workers to provide high levels of effort. In particular $\frac{\partial^2 e}{\partial w \partial p} > 0$ and $\frac{\partial^2 e}{\partial w \partial \eta} > 0$, i.e., the wage-effort relation is steeper for more productive and more trustworthy individuals. Moreover, $\frac{\partial^3 e}{\partial w \partial p \partial \eta} > 0$ implies that productivity and trustworthiness are complementary in enhancing the efficacy of gift exchange.

**Hypothesis 2.2** *We expect a positive wage-effort relation that is steeper for more productive and more trustworthy subjects.*

As a smaller wage gift is needed to elicit any level of effort, efforts are cheaper to implement for firms if workers have high $\eta$ and $p$. Hence, we expect firm profits to increase in those worker traits and to be highest when interacting with a worker that displays both traits.

**Hypothesis 2.3** *We expect firm profits to increase if interacting with more productive and more trustworthy subjects and to be highest in interactions with subjects that are productive and trustworthy.*

## 2.4 Experimental Results

### 2.4.1 Part One: Elicitation of Productivity and Trustworthiness

We start by reporting summary statistics for the coding task performance in part one from treatments PTB, P and T. Subjects receive three screens with 15 matches each such that the maximum attainable is 45 correct answers. Only one out of 192 subjects succeeded in giving all 45 answers within 90 seconds correctly such that time was indeed the limiting factor and the way productivity was measured does not harm high productivity subjects. The average number of correct answers given was slightly below 29 with a standard deviation of about 7 answers. Table 2.1 shows summary statistics and illustrates

Table 2.1: Descriptive Statistics (Real Effort Task)

| Treatment | Mean | S.D. | N | Median | Min | Max |
|---|---|---|---|---|---|---|
| PTB | 28.64 | 6.31 | 96 | 28.0 | 9 | 44 |
| P | 28.89 | 8.08 | 48 | 29.5 | 0 | 45 |
| T | 29.46 | 6.06 | 48 | 29.5 | 14 | 41 |
| Total | 28.91 | 6.71 | 192 | 29 | 0 | 45 |

Descriptive statistics for coding task performance in part one

that there are no differences across treatments with all treatments being almost identical in terms of the main statistics. We particularly do not find any evidence that subjects in the two control treatments (P and T) behave differently than in our main treatment PTB. We can therefore exclude that workers behave strategically in the elicitation phases to signal something to potential future employers as a model of career concerns would predict. The distribution of correct answers in all treatments is symmetric around the mean, but normality is rejected by all conventional tests, see Figure A2.4 in appendix A2.3. As a further robustness check, we regress the coding outcome on a number of socio-demographics to see whether there is explanatory power from gender, age, subject of study or the treatment. We also control for five character traits in the framework of the Big Five Personality Test that we elicited in a control questionnaire at the end of the experiment. Table A2.3 in appendix A2.2 clearly indicates that there is no effect from gender nor from a quantitative orientation in the subject of study (economics, mathematics, natural sciences) on the performance in the task. Apart from some negative effect of age and the character trait "agreeableness", there is no significant effect from the four other elicited personality traits from the Big Five Index on coding performance either. More importantly as the treatment P/T dummies are insignificant, the regressions confirm that there is no distortion from strategic concerns between the treatments (and hence different levels of information disclosure) on the outcome of the productivity task.

For the second dimension of information, we let subjects play a binary trust game presented to them in a neutral frame. Since we employ the strategy method, we have data on choice behavior in both roles of the trust game for every subject. Table 2.2 displays the percentage of subjects' behavior in the trust game per treatment. According to our

Table 2.2: Descriptive Statistics (Trustworthiness)

| Treatment | no trust<br>no returntrust | trust<br>no returntrust | no trust<br>returntrust | trust<br>returntrust | Total |
|---|---|---|---|---|---|
| PTB | 43.8% | 13.5% | 8.3% | 34.4% | 100% |
| P | 35.4% | 14.6% | 6.3% | 43.8% | 100% |
| T | 45.8% | 16.7% | 18.8% | 18.8% | 100% |
| Total | 42.2% | 14.6% | 10.4% | 32.8% | 100% |

Percentage of subjects trusting and returning trust per treatment

measure, about 40 % of all subjects can be considered selfish in the sense that they neither trust others nor do they return trust as a trustee. In a similar vein, about one third of subjects appears to have other-regarding concerns such that they both trust and return trust. The remaining quarter either trusts but does not return trust or vice versa. Although there is some variation across the treatments, these patterns are quite stable in all three treatments. Subject to the population averages, not to trust is indeed optimal for selfish subjects and cannot be rationalized by choosing to trust for strategic reasons.

What is important to note is that about half of all subjects trusted and also about 40 % of all subjects returned trust while about 60 % did not return trust. That is to say that our design succeeds in creating variation across subjects which makes information about other subjects valuable for the contracting phase. If in the population our binary measure of social and reciprocal concerns was distributed less symmetric, the value of the information would clearly decline - if not vanish - when certain character traits were only to show up in small minorities of the underlying population.

Since we deem the decision of returning trust as more indicative of an individual's concern for reciprocity in the gift-exchange game, we focus in the remainder of the analysis mainly on the behavior of subjects as a second mover in the described trust game. We control for the impact of the same socio-demographics on trustworthiness in a probit regression which can be found in Table A2.4 in appendix A2.2. A similar picture to above emerges with the absence of a gender, age and treatment effects. A quantitative subject of study decreases the probability of returning trust and all personality traits are insignificant with the exception of "conscientiousness". We also control for the number of correct answers in part one on the propensity to reciprocate trust in part two, but do not find any effect which confirms that there is no relationship between our measure of productivity in part one and reciprocal behavior in part two, which we summarize in our first result.[9]

**Result 2.1** *There are no differences across treatments in personal traits for individual productivity and reciprocal concerns. The two measures quantify two distinct dimensions of a person's characteristics.*

### 2.4.2 Part Two: Gift Exchange Game

We subsequently present the results from the contracting phase, where initially, i.e., before the measures are elicited, all workers were told which set of information would be disclosed to firms. This set of information consisted in

- worker productivity from part one (in treatments PTB and P) and

[9]The independence of our two measures is also confirmed non-parametrically with a highly insignificant Spearman rank correlation between the two measures at a significance level of $p = 0.57$.

- the decision whether to return trust or not in part two (in treatments PTB and T).

Firms do not have any experience or knowledge on how workers behave such that we consider firms' wage policies as the cleanest possible measure of their preferences for information about workers.

In the analysis of the data from the gift-exchange game, we will focus on the PTB sessions where both measures about workers were revealed to firms in a binary way. We will also comment on the results of the two control treatments P and T, but since the set of information firms could condition their wage policy on is smaller, we refrain from directly comparing decisions in P or T with PTB.

**Firm Behavior**

We begin by looking at wage offers received by workers. Wages were bounded to be not negative and not above 250 such that the surplus split under maximum efficiency constituted an interior solution.[10] Every worker obtained one offer from each of the 12 firms, but just received and saw the relevant wage level for him, which was randomly chosen. As a consequence, we can analyze all 12 wage offers directed to a worker through the strategy method, i.e., we have $48 * 12 = 576$ observations in the PTB treatment. Average wages that were submitted to one single worker are given in the first column of Table 2.3, whereas the second column lists the average of the actually randomly determined *relevant* wage offer. As provided levels of performance by agents are significantly lower than the

Table 2.3: Summary Statistics for Workers

| Treatment | Average wage offer | Actual wage offer | Average performance | Average productivity | Average returntrust |
|---|---|---|---|---|---|
| PTB | 82.3 | 82.5 | 20.1 | 27.8 | 43.8% |
| P | 108.4 | 109.2 | 23.0 | 29.2 | 37.5% |
| T | 89.0 | 87.1 | 19.4 | 29.8 | 20.8% |
| Total | 90.5 | 90.3 | 20.6 | 28.6 | 36.4% |
| Obs | 1152 | 96 | 96 | 96 | 96 |

Summary statistics for workers. The first columns show all submitted wage offers to a specific worker, the second column only the randomly selected relevant wage. The last column lists the percentage of workers that returned trust in part two.

[10] We refer to an outcome where the worker gives the maximum of 45 correct answers and receives a wage of 225, which would yield a payoff of 225 to both the firm and the worker.

elicited productivity measures in part one (Wilcoxon-signed-ranks-test, $p < 0.01$), there is evidence that agents do not put in unconditionally full effort levels in the contracting phase.[11] Furthermore, we see that the random attribution of roles to workers and firms has not distorted our two measures in the sense that the sample means of a session (24 subjects) lie close to the means of the workers (12 subjects). To get a deeper understanding on how firms set wages in PTB, we run a series of firm fixed effect regressions on the wage offer to a specific worker that we report in Table 2.4. We find a positive impact of

Table 2.4: Determinants of Wage Offers in PTB

| **Dep. Var.:** Wage offer | **I** PTB | **II** PTB |
|---|---|---|
| 1 if high | 27.681*** | |
| | (2.597) | |
| 1 if trustworthy | 13.243*** | |
| | (2.182) | |
| 1 if high and trustworthy | | 41.748*** |
| | | (3.412) |
| 1 if high and not trustworthy | | 26.164*** |
| | | (3.084) |
| 1 if low and trustworthy | | 11.949** |
| | | (2.246) |
| Constant | 66.100*** | 66.706*** |
| | (1.352) | (1.491) |
| Obs | 576 | 576 |
| $R^2$ | 0.14 | 0.15 |

Panel Regressions include firm fixed effects on wage offers. Standard errors (clustered on the session level) in brackets. *** represents significance at p=0.01, ** at p=0.05 and * at p=0.10.

being a high productivity and a trustworthy worker on the offered wage level, which is highly significant for both traits in specification I. When comparing the relative sizes, it is immediate to see that employers provide a higher wage premium for the productivity measure than for the trustworthiness measure. The wage premium is roughly double for the productivity measure compared to the trustworthiness measure. In specification II, we divide all workers into four categories. To do so, we classify a worker as being of "high productivity" or "low productivity" as well as being "trustworthy" or "not trustworthy". The left out category is a worker who is neither of high productivity nor trustworthy. We confirm the findings of specification I, i.e., the presence of information about each dimension of a worker increases the wage a firm offers in the first place. All three co-

[11] In what follows, we term "performance" the number of correct answers given by workers in the gift exchange relation with firms, to draw a clear semantic distinction to the measure of "productivity" in part one.

efficients are significantly different from each other (t-test between two coefficients, all three tests below $p = 0.024$) indicating that there is little substitution of the wage premia between the two dimensions. If a worker moves from the lower to the higher category in one dimension, this yields a constant wage premium regardless of her position in the other category.[12] The results support Hypothesis 2.1 from section 2.3 for the impact of information on wage setting behavior by firms.

**Result 2.2** *Firms are willing to pay a significant wage premium for both characteristics. The premium for being a high productivity worker amounts to roughly double the premium for being trustworthy.*

In the two control treatments P and T, information about one dimension of worker characteristics is not available to firms. When we run identical firm fixed effects regressions for these two treatments, we find a significant wage premium for productivity in the P treatment, but not for trustworthiness. In the T treatment, there is a positive premium for being trustworthy, but this is not significant. As expected, there is no wage premium for productivity in T. We take this as further evidence that information about workers matters for wage offers, but more so for the productivity dimension than for the measure of trustworthiness.[13] We next look at worker behavior in terms of effort provision.

### Worker Behavior

With respect to effort levels, the question arises what influences workers most in their decision to provide effort. Along the lines of the gift exchange literature, one can argue that the main driving force will be a high wage offer such that agents reciprocate by exerting high levels of effort. This notwithstanding, the characteristics of a person in terms of productivity and intrinsic willingness to perform well at a given task can similarly affect actual effort levels.

[12]We also explore what is the driving force of firms' wage offers in a regression on the average offered wage, but all firm characteristics (own coding and behavior in the trust game, expectations, gender, age, field of study, risk proxy, Big Five Index) turn out insignificant for the average level of wage offered for all 12 workers by the firm. We particularly do not find an effect of both measures of elicited expectations on the wage setting behavior.

[13]See Table A2.5 in appendix A2.2 for the detailed regression results.

Table 2.5: OLS Regression on Effort

| Dep. Var.: # Correct Choices | **I** PTB | **II** PTB |
|---|---|---|
| Wage | 0.183*** | 0.148*** |
| | (0.048) | (0.058) |
| 1 if productive | -1.065 | -1.805 |
| | (3.508) | (3.908) |
| 1 if trustworthy | 3.284 | 2.188 |
| | (3.032) | (3.333) |
| Female | | -2.411 |
| | | (3.885) |
| Age | | 0.038 |
| | | (0.376) |
| Quant | | -5.497 |
| | | (3.850) |
| Big Five Measures | | ✓ |
| Constant | 3.951 | 2.463 |
| | (3.864) | (18.788) |
| Obs | 48 | 48 |
| $R^2$ | 0.32 | 0.41 |

OLS Regression on the number of correct choices in the gift exchange game. Standard errors (clustered on the session level) in brackets. *** represents significance at p=0.01, ** at p=0.05 and * at p=0.10

To find out more about which of these rationales helps to explain worker behavior in the contracting phase, we regress the number of correct answers on the offered wage, both measures of worker characteristics and a set of controls which are reported in Table 2.5. We find that the offered wage has a highly significant positive impact on the amount of effort, which we take as a clear sign that gift-exchange considerations play a role in our real-effort experiment. Controlling for the wage which already includes information about worker types, a higher productivity measure from part one does not increase the performance in the interaction between firms and workers. Surprisingly, reciprocal concerns among workers are not predictive for the effort decision. A worker that has returned trust in part two of the experiment, gives on average the same amount of correct answers more than a worker who has not returned trust. When we control for interaction effects between wages and worker types, there are no additional effects on the effort decision either. This is indicative of wages being set optimally by firms with all informational value from types already incorporated in the wage offer. There is no additional effect of individual worker characteristics on effort that is not yet captured in the wage offer. When we control for other socio-demographic characteristics of workers, we do not find significant gender, age and quantitative dummies in specification II. All Big Five measures are insignificant. There is no evidence for a relationship between performance and characteristics in *excess* of the wage, what we summarize in our next result.

**Result 2.3** *Only the wage offer has a significant impact on performance in the gift exchange game. We do not find an additional positive impact from worker's characteristics on effort choices in excess of the one already embodied in the wage offer.*

### Performance of Contracts in the Gift Exchange Game

Putting the wage setting decision by firms and the worker's effort decision together, we turn to the analysis of the profitability of firm's wage policies given the information they have about workers. We pool all contracts concluded in the full information treatment PTB and allocate all workers into the four above mentioned broad categories.[14] Table 2.6

Table 2.6: Descriptives for Contracts in PTB

| **Worker Type** | **PTB treatment** | | | |
|---|---|---|---|---|
| | **Effort** | **Wage** | **Firm's Profits** | **Obs** |
| high productivity & trustworthy | 28.9 | 117.1 | 171.4 | 7 |
| high productivity & not trustworthy | 19.7 | 96.4 | 100.9 | 11 |
| low productivity & trustworthy | 19.4 | 69.6 | 123.9 | 14 |
| low productivity & not trustworthy | 17.1 | 69.0 | 102.3 | 16 |
| **Total** | **20.1** | **82.5** | **118.35** | **48** |

Statistics of all contracts for the four different worker types in the PTB treatment.

shows the key summary statistics for all 48 concluded contracts in the PTB treatment. We find that particularly the interaction of both worker characteristics produces high levels of effort from workers (i.e., the number of correct choices given in the gift-exchange relation). These workers give close to 10 correct answers more than all other workers which leads to higher profits for firms only when they interact with a worker of this type. Wages are increasing in worker characteristics and significantly different according to the worker type identified (Kruskal-Wallis-Test on session averages, $p = 0.01$) whereas there is no difference in firms' profits (Kruskal-Wallis-Test on session averages, $p = 0.15$). The two

[14] All results in this section are identical if we were to include the data from the two control session P and T, but we want to exclude any possible effect that the absence of one piece of information might have both on firms and workers in what follows.

characteristics act as complements to produce highest levels of efficiency and effort only if they both are present at the same time. We find further evidence for the importance of the trustworthiness characteristic by constructing a variable that compares the amount of worker effort (i.e., number of correct answers) with worker productivity from part one and abstracts from the wage. This measure takes on values of 71 % and 76 % for high and low productivity workers, but increases to 83 % for trustworthy workers and decreases to 67 % for not trustworthy workers. Trustworthy workers provide the highest levels of effort compared to their measure of productivity in part one.

**Result 2.4** *Workers that have both characteristics - high productivity and trustworthiness - provide significantly higher levels of effort which leads to higher firm profits. In the absence of at least one of the character trait, efforts and firms' profits are substantially lower in a firm-worker interaction.*

We end this section by presenting evidence about the interaction of the worker types with the profitability of employment strategies of firms in PTB. We have seen that firms' profits are considerably increased when they interact with a high productivity and trustworthy worker. Since firms cannot actively choose their worker in our design, but are allocated a worker at random and can only offer different wages, we next proceed to analyzing how the different types of workers react to an offered wage. We begin with the effects of different levels of trustworthiness among workers on contracting behavior between firms and workers before we proceed to our productivity measure.

Not surprisingly and in line with the literature, we find a significantly positive relationship between the wage and the provided effort level.[15] The slope can be interpreted as the wage increment required to induce the worker to provide one additional unit of effort. Both slopes are significantly positive at 0.19 (trustworthy workers) and 0.16 (not trustworthy workers). Since the efficiency factor of effort in our design is exogenously fixed at 10, the wage effort relation must be significantly larger than 0.1 to make it profitable for firms to pay high wages. Only then additional wage costs are outweighed by an increase in revenue and hence lead to an increase in profits. For trustworthy workers, the coefficient

[15] Figure A2.5 in appendix A2.3 displays the scatterplot of the wage-effort relationship for trustworthy and not trustworthy workers.

in the wage-effort regression is significantly different from 0.1 at $p = 0.03$ (t-test), but not so for not trustworthy workers at $p = 0.44$ (t-test). It pays to increase the wage for firms only to workers that are trustworthy, since profits do not increase significantly when interacting with a not trustworthy worker. In Figure 2.1 we plot a firm's profit against the implemented wage separately for trustworthy and not trustworthy workers in the PTB treatment. Firm profits are increasing with both worker types, but only

Figure 2.1: PTB: Wage-Profit Relation (Trustworthiness Measure)


Scatterplot of the wage-profit relation of firms for trustworthy workers (left panel) and not trustworthy workers (right panel). The shaded areas indicate the 95 % confidence interval around the linear regression line.

significantly so for trustworthy workers. In terms of profits from the gift-exchange with trustworthy subjects, a wage increment of 1 EP leads to an increase of 0.9 EP in profits, which is significantly different from zero. When firms interact with a worker that did not return trust, however, there is a flatter relation between the offered wage and firms' profits. An additional increment in the wage offer leads to an increase of 0.3 EP in profits, which is not different from zero. Spearman rank correlation coefficients between firm's profit and the wage confirm this finding: $p = 0.03$ for trustworthy types and $p = 0.25$ for not trustworthy types. In addition, we find a lower profit variance for trustworthy types which seems prima facie desirable.

We next look at the interaction of the wage and the measure for worker productivity. There is again a significantly positive relationship between wages and effort for high productivity workers (0.20) and low productivity workers (0.18), both at $p < 0.001$ in an OLS regression. Both types of workers react to the wage incentive in a positive way. The slope for both types of productivity workers is positive but not significantly bigger

Figure 2.2: PTB: Wage-Profit Relation (Productivity Measure)


Scatterplot of the wage-profit relation of firms for high productivity workers (left panel) and low productivity workers (right panel). The shaded areas indicate the 95 % confidence interval around the linear regression line.

than 0.1, the threshold needed for a higher wage to induce higher profits for firms (t-test, 0.24 and 0.15 respectively). We confirm this by Spearman rank correlation coefficients between wages and profit levels that are positive, but not significantly so ($p = 0.10$ for not productive types and $p = 0.33$ for productive types). Figure 2.2 reports the wage-profit relationship in a similar presentation as above, but now for the productivity measure. Taking the results for both elicited measures together, we find evidence in favor of Hypothesis 2.2 with a particular focus on the trustworthiness measure.

The relationship between profits and wages is important to determine optimal wages for firms that want to maximize their profits. The mapping of wages into profits must not be linear and will depend on the type of the worker in the presence of information about types. From the relationship between firm profits $\pi$ and initially offered wages $w$, we estimate the following relationship on the basis of the data from the contracting phase:

$$\pi = \alpha \cdot w + \beta \cdot w^2$$

We force the wage-profit curve through the origin by omitting the constant and allow for a quadratic relationship as the simplest non-linear functional form. Standard OLS regressions for each of the four types separately produce our estimates of $\alpha$ and $\beta$. After solving for the implied optimal wage offer $w^*$, we calculate implied profits $\pi^*$ and compare this to the actual averages we have from our data. We find that firms optimally should

Table 2.7: Observed and Implied Optimal Wages in PTB

| **Worker Type** | **Observed Wage** | **Implied Optimal Wage** | **Observed Profits** | **Implied Profits** |
|---|---|---|---|---|
| high productivity & trustworthy | 117 | 152 | 171 | 185 |
| high productivity & not trustworthy | 96 | 130 | 101 | 118 |
| low productivity & trustworthy | 70 | 105 | 124 | 164 |
| low productivity & not trustworthy | 69 | 0 | 102 | 0 |

offer wages that are generally higher than those observed in the experiment. Firms could increase their profits by increasing the wage.[16] Optimal wages differ with respect to the type such that firms should condition their wage offer on the worker type. The qualitative order of optimal wages is identical to that observed in the experiment with highest wages being best for trustworthy workers with high productivity - leading to maximal profits. A corner solution for the lowest worker category yields implied wages and profits of zero, i.e., the theory would suggest that it does not pay for firms to hire these types. The last column in Table 2.7 provides evidence in favor of Hypothesis 2.3 and confirms the positive impact particularly of the trustworthiness measure for firm profits.

[16]The wage setting of firms naturally also interacts with firms' individual characteristics e.g. risk aversion, since there is always the danger that a worker shirks completely. Risk aversion could make observed wage levels optimal.

Taken together, the results from the contracting phase suggest that gift-exchange is robust across all identified types of workers from the positive wage-effort relations. Regardless of the information about workers available to firms in the contracting phase, they can expect on average a positive wage-effort relation from the interaction with workers. However, only with trustworthy types gift-exchange is so effective that a higher wage leads to a significant increase in firm profits. Worker types affect firm profits and are rightly taken into account by firms when writing contracts. Only if information about worker types is available, firms can increase their profits by tailoring the corresponding wage level to a specific worker. Exploiting complementarities between the incentive scheme and worker types therefore crucially depends on the amount of information accessible to firms before writing contracts.

**Result 2.5** *Gift exchange is present for all types, but is strongest for trustworthy types. Optimal wages depend on worker characteristics such that information about workers is valuable for firms to tailor incentives to worker types.*

#### Effects of the Mode of Information Presentation

We have seen above that the availability of information helps firms to adapt contracts to workers and shapes outcomes in a gift-exchange interaction. In this chapter, we examine if the exact presentation of information has an impact on contracting and final outcomes.

To do so, we run a series of sessions in an additional treatment (PT) where the information about productivity is given to firms in a quasi-continuous manner. Instead of the binary information in PTB that a worker has a higher or lower productivity than the mean, firms are now given access to the exact number of correct answers in part one as a measure of worker productivity. Everything else (including the binary nature of the trustworthiness measure) is kept identical to the PTB treatment from above. Note that this modification is not a different framing of the same piece of information, since firms have access to more information which is more precise compared to the binary case. We run 6 additional sessions of this PT treatment with an additional 144 subjects, i.e., 864 contract offers.

For the elicitation phases of productivity and trustworthiness as well as expectations, there are no differences to the PTB treatment.[17] Similar to above, we first look at the wage setting behavior of firms. Table 2.8 reports a panel regression with firm fixed effects where worker productivity is given in a continuous manner. Wage offers are now much

Table 2.8: Determinants of Wage Offers in PT

| **Dep. Var.:** | **I** | **II** | **III** |
|---|---|---|---|
| Wage offer | Treatment PT | PT+P+T | PT+P+T |
| Worker productivity | 3.020*** | 2.584*** | 2.347*** |
| | (0.205) | (0.385) | (0.497) |
| 1 if trustworthy | 3.923 | 3.933 | 1.283 |
| | (2.868) | (2.381) | (2.633) |
| Worker productivity * treatment P | | | 0.854 |
| | | | (0.515) |
| 1 if trustworthy * treatment T | | | 10.892 |
| | | | (7.369) |
| Constant | 4.270 | 18.479 | 20.849 |
| | (5.752) | (11.000) | (11.748) |
| obs | 864 | 1440 | 1440 |
| $R^2$ | 0.10 | 0.12 | 0.12 |

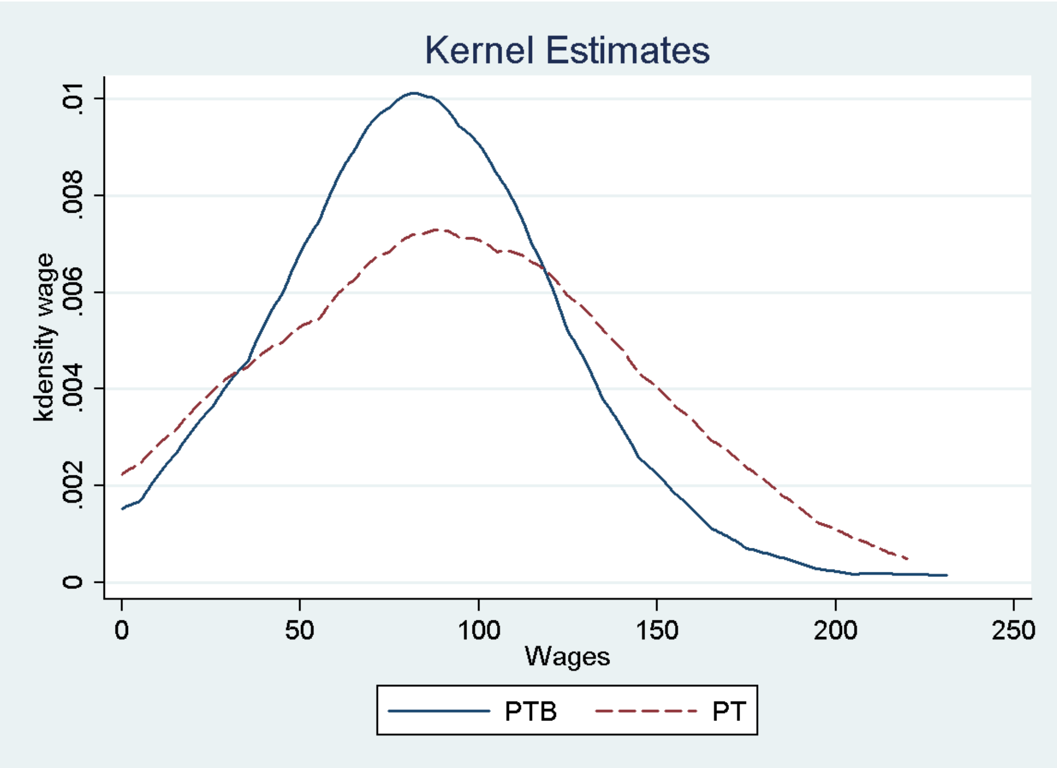
Panel Regressions include firm fixed effects on wage offers in treatments PT, P and T. Standard errors (clustered on the session level) in brackets. *** represents significance at p=0.01, ** at p=0.05 and * at p=0.10.

more focused on information about worker productivity rather than trustworthiness. For every correct answer in part one, firms are willing to pay a highly significant wage premium of about 3 EP. We find no significant wage premium for trustworthiness anymore in PT. We are able to compare the results in PT to our two control treatments P and T since the precision of information is identical in these three treatments, but pooling yields identical results. The increased precision of the productivity measure appears to be a much more important concern for firms in the wage setting phase compared to the binary information about trustworthiness. As a consequence, a different distribution of wage offers arises endogenously from a different presentation of information.

Figure 2.3 documents this difference and plots the estimated distributions of all wage offers in PT and PTB. The binary presentation of information in PTB leads to a decrease in the mean and the variance of wage offers compared to the PT treatment where firms have access to details about worker productivity. Both distributions are significantly different from each other (Kolmogorov-Smirnov-test, $p < 0.01$). Workers subsequently respond to this new wage distribution in a different way, when we look at effort choices

[17]Since the instructions for the first two parts of the experiment were identical to the PTB treatment, this is not surprising at all. Thus we omit a detailed comparison of these results between PTB and PT here. They are available upon request from the authors.

Figure 2.3: Distribution of All Offered Wages




Distribution of all wage offers submitted by firms in treatments PT and PTB.

of workers in PT. We document a significantly positive wage-effort relation across all workers, but in contrast to the findings in PTB, there is no positive effect on firm profits from the interaction with a trustworthy type in PT. With the wage distribution in PT being generated under a pronounced focus on the productivity measure by firms, outcomes in the gift-exchange games change considerably compared to our main PTB treatment.

**Result 2.6** *Subtle differences in the presentation of worker information to firms matters a lot for contracting behavior and subsequent outcomes in the gift-exchange game.*

## 2.5 Discussion and Conclusion

We present evidence from a laboratory gift-exchange experiment indicating that firms condition their wage policies on available information about worker productivity and worker's trustworthiness. Firms offer more generous wages to workers who are, according to the elicited measures, more productive and more trustworthy. Our results suggest that workers with better productivity skill sets and high levels of trustworthiness earn wage premia on the labor market.

The channel through which the availability of information has an impact on final outcomes lies in the possibility for firms to adapt contract offers to specific worker types in order to make use of complementarities between wages and types. Our results show that gift-exchange considerations play a role across all types, but they are strongest for trustworthy workers. Only for trustworthy types wage increases translate into an increase in profits for firms. Optimal wages for firms differ with respect to the worker type and induce maximal firm profits for high productivity workers that are trustworthy which illustrates the importance of complementarities also with respect to different worker characteristics. We show that subtle differences in the information presentation to firms induce an endogenously different distribution of wage offers in our control treatment PT. If given access to the precise level of worker productivity, firms are much more focused on the productivity measure than on trustworthiness.

While moral hazard is an important friction that governs contracts and incentives on labor markets, we argue in this chapter that there is an important interaction between worker heterogeneity and the contractual incompleteness of labor markets. The role of a worker's individual productivity for contracting outcomes has long been acknowledged. We show that if a firm takes heterogeneity with respect to trustworthiness into account, it affects the effectiveness of gift-exchange and hence its profit levels. Since the interaction between worker types and the solution to the moral hazard problem matters, the existence of information about worker characteristics contains an economic value for firms. Given the resources spent on information acquisition by firms for the hiring decision, we consider the role and the acquisition of information on labor markets a field of high economic relevance and a promising topic for future research to foster our understanding of the functioning of labor markets under incomplete contracts.

## 2.6 Appendix A2

### Appendix A2.1: Elicitation of Expectations

In this appendix, we describe in details the elicitation of subjects' expectations about the characteristics in the population in an incentivized manner. After the elicitation of the productivity and the trustworthiness measure, all participants are asked to estimate the number of correct answers in the productivity task reported before. A correct guess of the session average is rewarded with a prize of 100 EP, from which 10 points are deducted for every correct answer that the guess was away from the true value. If the difference between the guess and the true value exceeded 10 answers, subjects earned at worst 0 EP from this part.[18] In a similar vein, we ask subjects how many of the 24 subjects in their session have chosen to reciprocate trust. Subjects are rewarded for the precision of their guesses with a prize of 100 EP if their estimate was correct, and 20 points were deducted for every subject that their guess was away from the true value. Hence, if their guess was more than 5 subjects away from the true value, earnings were 0 EP from this part. Expectations were elicited referring to the current session (24 subjects) of the experiment, which we consider sufficiently large that subjects perceive their impact on the session average small enough to enter their expectation about the whole population.

With no feedback about the choices of other participants, individual expectations about the population are a likely candidate to explain the decision on contract choices later on. The resolution of this part took also place at the end of the experiment, such that subjects entered the gift exchange without any information about the behavior of other subjects in the experiment.

Table A2.1: Descriptives of Expectations

| Treatment | # Correct Choices realized | # Correct Choices expected | # Returntrust realized (per session) | # Returntrust expected (per session) |
|---|---|---|---|---|
| PTB | 28.6 | 23.8 | 10.3 | 9.9 |
| P | 28.9 | 23.9 | 12 | 10.5 |
| T | 29.5 | 26.0 | 9 | 9.8 |
| Total | 28.9 | 24.4 | 10.4 | 10.0 |

Subjects' expectations vs. realizations of two elicited measures for both elicited measures.

[18]Subjects could only enter integer guesses and the average was rounded to an integer.

We report in Table A2.1 both subjects' expectations and actual realizations in the three treatments. As far as the number of correct choices in part one is concerned, subjects underestimate the average number of correct outcomes by about 4 answers compared to the true average value. This is confirmed by a highly significant Wilcoxon-signed-rank-test between the own performance and the guess of the average productivity measure ($p < 0.01$). Indeed, out of all 192 subjects, 146 give a lower expectation of the average than their own coding performance in part one, 36 a higher one and 10 subjects consider themselves to be average. Note however that at this point of the experiment, subjects do not know their performance in part one explicitly but only implicitly from remembering how many correct answers they gave.

Table A2.2: OLS Regression on Expectations

| | **Dep. Var.:** Expected Average of Correct Choices | **Dep. Var.:** Expected # Subjects Returning Trust |
|---|---|---|
| # Correct Choices | 0.431*** | -0.101* |
| | (0.058) | (0.060) |
| 1 if trust | 0.291 | 2.528*** |
| | (0.838) | (0.868) |
| 1 if trustworthy | -0.007 | 7.074*** |
| | (0.845) | (0.876) |
| Female | 0.749 | -0.016 |
| | (0.739) | (0.766) |
| Age | 0.036 | -0.145 |
| | (0.093) | (0.096) |
| Treatment P | 0.208 | -0.171 |
| | (1.089) | (1.129) |
| Treatment T | 1.875** | 0.697 |
| | (0.837) | (0.868) |
| Constant | 10.064*** | 12.055*** |
| | (3.098) | (3.211) |
| Obs | 168 | 168 |
| $R^2$ | 0.28 | 0.45 |

OLS Regression on elicited expectations. Coefficients show effects relative to answers in the PTB treatment. trust and returntrust are dummy variables for behavior in part two. Due to a server breakdown at the end of one P session, we are lacking the socio-demographic variables of one session (24 subjects). Standard errors in brackets. *** represents significance at p=0.01, ** at p=0.05 and * at p=0.10.

When asked about the number of subjects that returned trust in their session, guesses are more accurate and subjects correctly predict that a bit less than half of the participants chose to return trust. To gain a deeper understanding of what drives the formation of expectations, we regress expectations on behavior in part one and include the control treatments P and T. From Table A2.2 it is immediate to see that subjects are strongly (positively) influenced by their past behavior, which is suggestive evidence for the false consensus effect. For estimating the average number of correct choices in the elicitation task, the own result is highly significant. In a similar manner, having trusted and returned trust oneself increases one's expectation of the number of subjects that return trust within a session significantly. Socio-demographics do not matter for expectations and we find that there are some negative effects from the coding task on expectations of returning trust, i.e., subjects with a high productivity in part one are susceptible to adapt their expectations about reciprocal inclinations downward, but not vice versa.

## Appendix A2.2: Tables

Table A2.3: OLS Regression on Productivity Measure

| Dep. Var.: # Correct Choices | **I** | **II** |
|---|---|---|
| Female | -0.308 | -0.548 |
| | (1.021) | (1.203) |
| Age | -0.298** | -0.277** |
| | (0.123) | (0.126) |
| Quant | -0.546 | -0.286 |
| | (1.141) | (1.150) |
| Treatment P | 0.948 | 0.748 |
| | (1.449) | (1.462) |
| Treatment T | 0.881 | 0.391 |
| | (1.112) | (1.135) |
| Big Five (Extraversion) | | 0.588 |
| | | (0.426) |
| Big Five (Agreableness) | | 0.840* |
| | | (0.506) |
| Big Five (Conscientiousness) | | -0.587 |
| | | (0.455) |
| Big Five (Emotional Stability) | | -0.221 |
| | | (0.446) |
| Big Five (Openness) | | -0.169 |
| | | (0.541) |
| Constant | 36.019*** | 34.261*** |
| | (3.154) | (4.817) |
| Obs | 168 | 168 |
| $R^2$ | 0.04 | 0.07 |

OLS Regression on the number of correct answers in part one. Coefficients show effects relative to answers in the PTB treatment. Quant is a dummy for quantitative orientation of studies. All Big Five measures are on a scale from 1 to 7 indicating the strength of the individual personality trait. Due to a server breakdown at the end of one P session, we are lacking the socio-demographic variables of one session (24 subjects). Standard Errors in brackets. *** represents significance at p=0.01, ** at p=0.05 and * at p=0.10.

Table A2.4: Probit Regression on Trustworthiness

| Dep. Var.: 1 if Trustworthy | **I** | **II** | **III** |
|---|---|---|---|
| Female | 0.109 | 0.109 | 0.295 |
| | (0.210) | (0.210) | (0.255) |
| Age | 0.006 | 0.007 | 0.009 |
| | (0.025) | (0.025) | (0.026) |
| Quant | -0.849*** | -0.848*** | -0.884*** |
| | (0.249) | (0.249) | (0.254) |
| Treatment P | 0.410 | 0.409 | 0.513 |
| | (0.301) | (0.301) | (0.312) |
| Treatment T | -0.194 | -0.195 | -0.256 |
| | (0.232) | (0.232) | (0.239) |
| Big Five (Extraversion) | | | -0.069 |
| | | | (0.092) |
| Big Five (Agreableness) | | | 0.096** |
| | | | (0.107) |
| Big Five (Conscientiousness) | | | -0.229 |
| | | | (0.098) |
| Big Five (Emotional Stability) | | | 0.141 |
| | | | (0.096) |
| Big Five (Openness) | | | 0.069 |
| | | | (0.115) |
| # Correct Choices | | 0.001 | -0.001 |
| | | (0.016) | (0.017) |
| Constant | -0.188 | -0.243 | -0.249 |
| | (0.641) | (0.879) | (1.170) |
| Obs | 168 | 168 | 168 |
| Pseudo $R^2$ | 0.07 | 0.07 | 0.10 |

Probit Regression on trustworthiness. Coefficients show effects relative to answers in the PTB treatment. Quant is a dummy for quantitative orientation of studies. All Big Five measures are on a scale from 1 to 7 indicating the strength of the individual personality trait with 1 being very weak and 7 being very strong. Due to a server breakdown at the end of one P session, we are lacking the socio-demographic variables of one session (24 subjects). Standard Errors in brackets. *** represents significance at p=0.01, ** at p=0.05 and * at p=0.10.

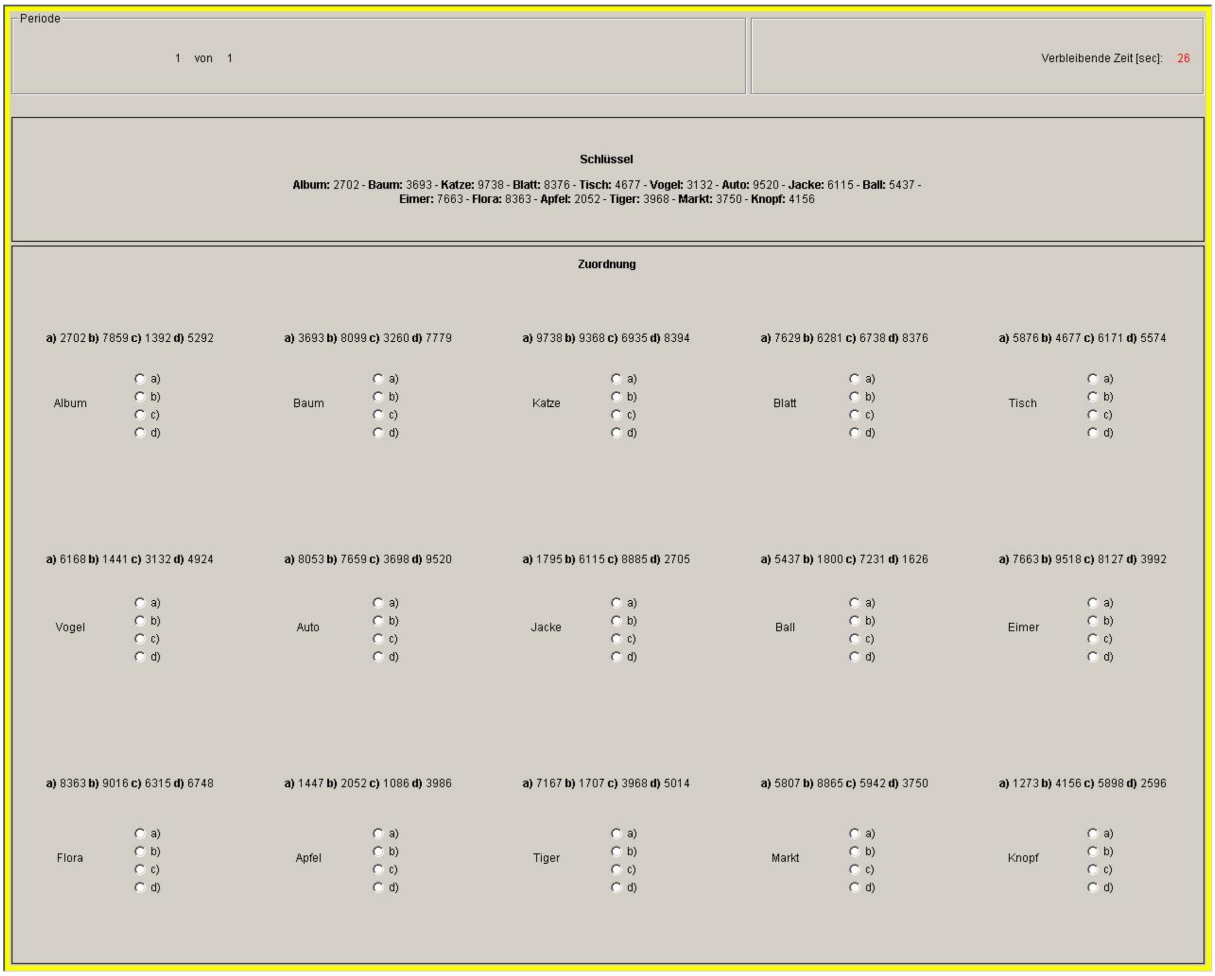
Table A2.5: Determinants of the Wage Offer - Control Treatments P and T

| **Dep. Var.:** | **I** | **II** |
|---|---|---|
| Wage offer | Treatment P | Treatment T |
| Worker Productivity | 3.204*** | 0.109 |
| | (0.021) | (0.022) |
| Worker's Trustworthiness | 1.068 | 18.223 |
| | (1.835) | (11.048) |
| Constant | 14.431* | 81.953** |
| | (1.314) | (2.986) |
| Obs | 288 | 288 |
| $R^2$ | 0.17 | 0.04 |

Panel Regressions include firm fixed effects on wage offers. Productivity from part one is displayed in a continuous manner to firms. Return trust is a dummy variable for worker behavior in part two. Standard errors (clustered on the session level) in brackets. *** represents significance at p=0.01, ** at p=0.05 and * at p=0.10.

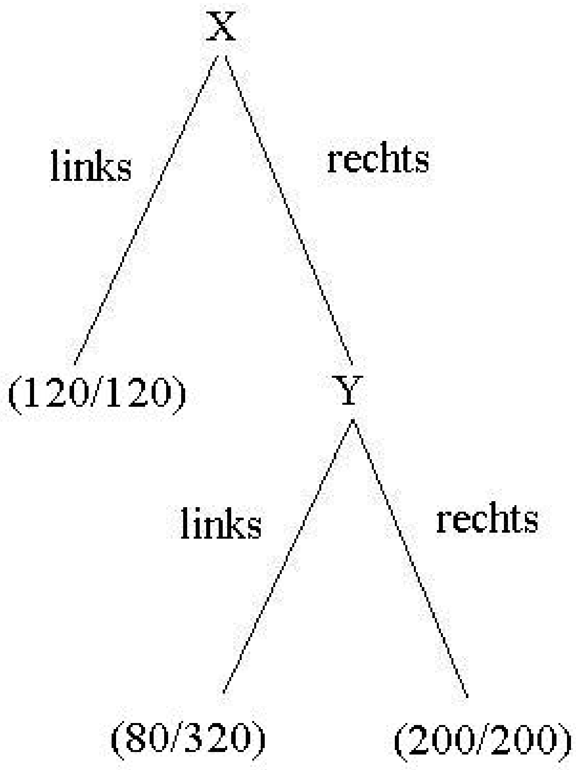
## Appendix A2.3: Figures

Figure A2.1: Screenshot: Real-effort Task




Screenshot of the real-effort task: The key is shown in the upper half of the screen, the matching is done in the lower half of the screen. Subjects had 30 seconds for each of the three screens.

Figure A2.2: Trust Game

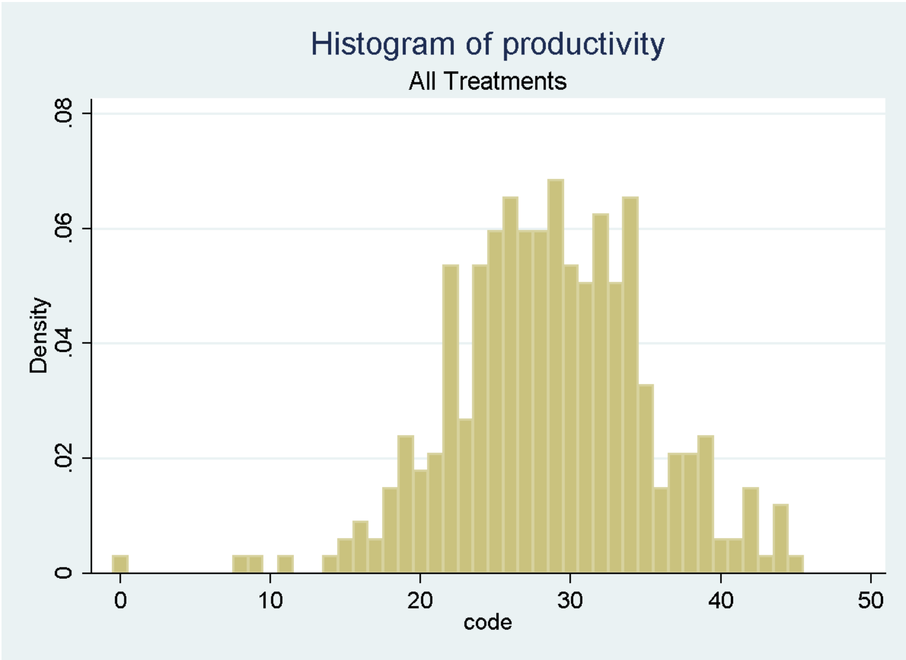



Presentation of the trust game to subjects. Subjects had to choose as person X (first mover) and as person Y (second mover), where at each point they could choose between "left" and "right". The corresponding payoffs are given in experimental points (EP) with the first mover's payoff listed first.

Figure A2.3: Screenshot: Wage Entry Screen for Firms


Wage entry screen for firms: Worker Characteristics are shown in brackets; firms enter one wage for every worker.

Figure A2.4: Histogram of Coding Task




Histogram of the elicitation stage for productivity for all treatments (PT, P, T, PTB) with density of the number of correct answers.

Figure A2.5: PTB: Wage-effort Relation (Trustworthiness Measure)


Scatterplot of the wage-effort relation for trustworthy types (left panel) and not trustworthy types (right panel) in PTB. The shaded areas indicate the 95 % confidence interval around the linear regression line.

Figure A2.6: PTB: Wage-effort Relation (Productivity Measure)


Scatterplot of the wage-effort relation for high productivity types (left panel) and low productivity types (right panel) in PTB. The shaded areas indicate the 95 % confidence interval around the linear regression line.

# Appendix A2.4: Instructions (translated from German)

PTB/P/T/PT

This experiment serves the investigation of economic decision making. In the experiment you and other participants[1] of the experiments are asked to make decisions. You can thereby earn money. Your decisions as well as the decisions of other participants determine your earnings from the experiment according to the rules explained below.

The whole experiment approximately lasts **1 hour and 15 minutes and consists of four parts**. First of all, you receive the instructions for part I. Instructions for parts II to IV are handed out to you at the beginning of the respective parts. For each part you are asked to enter your decisions into the computer. The parts are not independent of each other. This implies that decisions taken in one part of the experiment may sometimes (not always) affect other parts of the experiment.

Please raise your hand if you have any questions after reading through the instructions of during the experiment. One of the experimenters will then come to you and answer your questions in private.

While making your decisions, there is a clock counting down in the right upper corner of your computer screen. This clock serves as a guide for how much time it should take you. You may, of course, exceed the time limits. Once time has run out, it is only the pure information screens which will be dismissed as they do not ask you to make any decisions.

## Payment

At the beginning of the experiment you receive **4 Euro** for arriving on time. During the experiment you can earn more money by collecting **points**. At the end of the experiment, the points get converted into Euro at the exchange rate of

**1 Point = 0.0125 Euro (1.25 EUROCENT)**
that is **1 Euro = 80 Points**

**At the end of the experiment the amount of money you earned during the experiment as well as your 4 Euro starting balance will be paid to you in cash.**

## Anonymity

**At no point during or after the experiment you will find out with whom you interact and the identity of other participants. In turn, other participants will not find out your identity and your earnings at any point during or after the experiment.** There is a possibility that decisions you took in [PTB/PT: parts I and II] [P: part I] [T: part II] are made public to other participants in later parts of the experiment. Please note that your identity remains secret all the same.

**It is strictly prohibited to communicate with other participants during the experiment.** Furthermore, please note that you may only use the functions of the computer that are part of the experiment. Communication or playing around with the computer results in exclusion from the experiment.

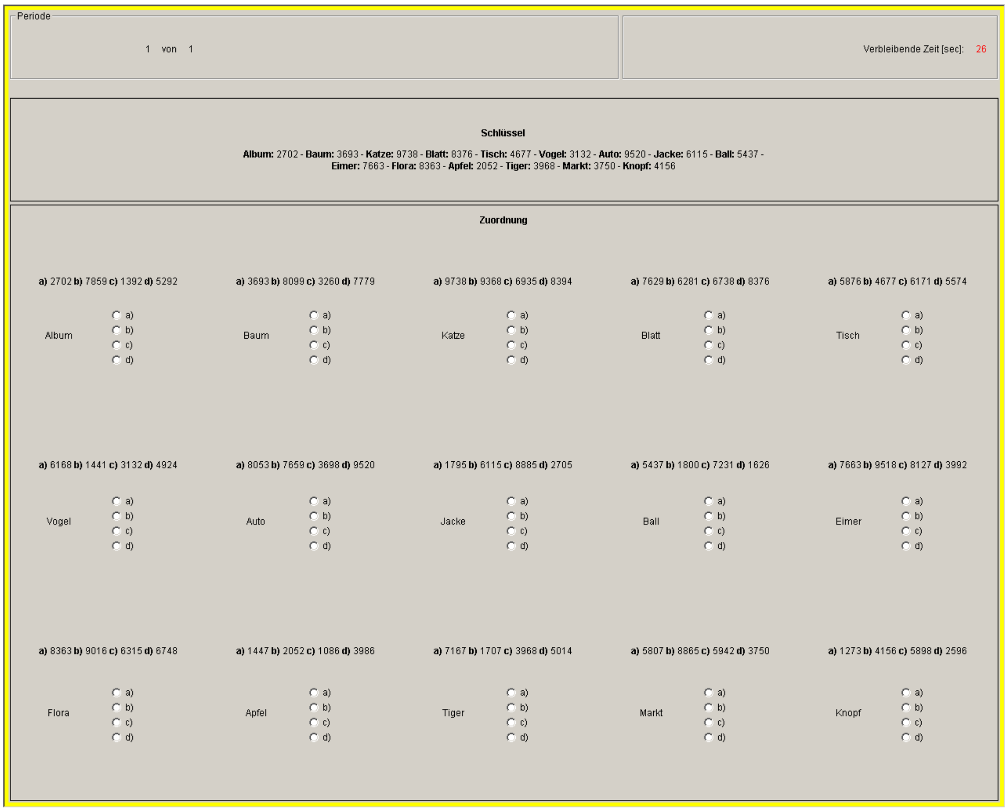
[1] For convenience, we only use male terms in the instructions. They should be considered as being gender neutral.

Page 1 of 8

PTB/P/T/PT

## Part I

During the first part of the experiment you are asked to link terms to the numerical codes corresponding to them. The screen illustrating a representative scenario is displayed below:




The upper part of the screen shows a code key that links specific terms to specific codes. The numerical code always consists of four digits. In the lower part of the screen, you have to assign terms to their respective numerical codes. For each term there are four possible codes, displayed as options a) to d), but **only a single one code among the listed one is correct**. Please click on the correct numerical code for each of the terms. The order of terms as shown in the key code is identical to the order of terms in the assignment task.

There are always 15 terms per screen and you are given 30 seconds per screen. This means that **after 30 seconds there is a new screen that pops up and contains 15 new terms and codes**. In total, you are given 90 seconds for the numerical code tasks, i.e. **three different screens** pop up one after the other.
The order in which you assign terms to their corresponding codes does not play a role. You may skip terms and you may go back to change your old decisions. All terms that you were not able to assign before the screen disappears after 30 seconds do not count for your final payment determination.

For every correct answer you receive **10 points**. For each wrong answer you get **a deduction of 10 points**. You may not run into a loss however, i.e. it's not possible to get minus points and **in the worst case your earnings amount to 0 points** in this part of the experiment. The difference between correct and wrong answers is called correct assignments.

Page 2 of 8

PTB/P/T/PT

Correct assignments = # correct answers – # wrong answers

You will only find out about your performance and thus about the amount of points you earned in this part of the experiment at the very end of the experiment. **Your earnings from this part of the experiment correspond to the sum of all points that you earned by giving correct answers reduced by the points that got deducted for each wrong answer.**

*<u>Example 1:</u> You achieve 26 correct answers and 2 wrong answers. Your earnings amount to (26-2) * 10 = 240 points.*

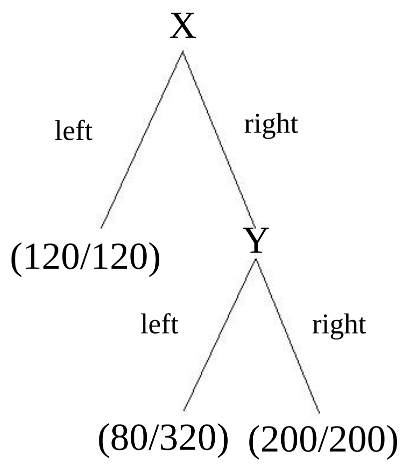
*<u>Example 2:</u> You achieve 8 correct answers and 12 wrong answers. Your earnings are 0 points.*

There will be a 60 second trial run of the numerical code task before the start of the experiment in order to get familiar with the computer program. The trail run is not part of the experiment and does not influence your final payments.

## Part II

*(parts and instructions were presented sequentially to subjects)*

During the second part of the experiment you are asked to make two decisions, both of which refer to the following situation. Numbers correspond to the earnings in points from this part, and they are labelled in a way such that person X is always referred to first:




Person X chooses between "**left**" and "**right**". If he decides for "**left**", person X himself and person Y receive 120 points respectively from this part of the experiment. If he decides on "**right**", it is person Y who decides on the final earning points in this part. If person Y chooses "**left**", person X receives 80 points and person Y receives 320 points. If he chooses "**right**", person X and person Y receive 200 points respectively from this part of the experiment.

**You do not know whether you are person X or person Y. The decision is made by the computer at the end of the experiment only.** You thus have to make two decisions: The first decision is implemented if you end up becoming person X ("left" or "right"). The second decision is implemented if you end up becoming person Y ("left" or "right"). At the end of the experiment the type of person (X or Y) is randomly assigned to you. Also, there is another participant who is randomly assigned to you and who takes on the respective other type of person.

Page 3 of 8

PTB/P/T/PT

**It is only your decision of your randomly assigned person type that is relevant for your final earnings. This means that if you end up being person X (or Y), the decision that you took as person X (or Y) is relevant only.** The final earnings from this part of the experiments are therefore not found out until the end of the experiment.

Please type your decisions into the computer and confirm by clicking OK. As long as you haven't used the OK button yet, you may change your decisions.

## Part III

During the third part of the experiment you are asked for your assessment concerning all participants' past behaviour in parts I and II. **The better your estimates for the average outcomes in the first two parts, the more money you earn in part III.** You are asked to estimate two outcomes.

Estimate 1: How did the participants in part I of this experiment perform **on average**? As a reminder, part I dealt with the numerical code task. **Please enter your estimate for the average number of correct assignments of all participants in part I on your screen.** As defined in part I, the number of „correct assignments" refers to the number of correct answers net of the number of wrong answers.

**If your estimate is correct, you receive 100 points. Points will be reduced in case your estimate deviates from the true value.** 10 points are taken away for each assignment that your estimate deviates from the true average of all participants. For example, in case your estimate deviates from the true value by one correct assignment, you receive 90 points. In case your estimate deviates from the true value by two correct assignments, you receive 80 points, etc. In case your estimate deviates by 10 or more correct assignments, you receive 0 points.
You are informed on the true average and on your earnings from this part of the experiment only at the end of the experiment.

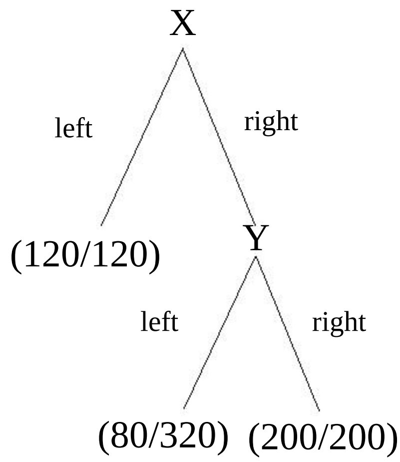
*Example 1: You estimate the average to be 20 correct assignments. The true average is 17 correct assignments. Your estimate thus deviates from the true average by 3 correct assignments. Consequently, you earn 100 – 3*10 = 70 points.*

*Example 2: You estimate the average to be 9 correct assignments. The true average is 22 correct assignments. Your estimate thus deviates from the true average by 13 correct assignments. Consequently, you earn 0 points.*

Estimation 2: How many of the 24 participants of this experiment chose **"right" being person Y** in part II? Please enter your estimate for the average number of correct assignments of all participants in part I on your screen. **If your estimate is correct, you receive 100 points**. 20 points are taken away for each participant that your estimate deviates from the true number of participants choosing "right". For example, in case your estimate deviates from the true number of participants by one participant, you receive 80 points. In case your estimate deviates from the true number by two participants, you receive 60 points, etc. In case your estimate deviates by 5 or more participants, you receive 0 points.
You are informed on the true number of participants and on your earnings from this part of the experiment only at the end of the experiment.

The following figure serves as a reminder of the situation faced in part II:

PTB/P/T/PT




*<u>Example 1</u>: You estimate the number of participants that chose "right" being person Y to be 16. The true number is 14 participants. Your estimate thus deviates from the true number by 2 participants. Consequently, you earn 100 – 2*20 = 60 points.*

*<u>Example 2</u>: You estimate the number of participants that chose "right" being person Y to be 5. The true number is 12 participants. Your estimate thus deviates from the true number by 7 participants. Consequently, you earn 0 points.*

**Please note:** The decision that participants took being person X does not play a role for this part. Your estimate merely concerns the decision that all 24 participants took being **person Y**. The type of person that is randomly assigned to the participants at the end of the experiment is irrelevant to this part. Your estimate should refer to all participants of this session.

## Part IV

During the fourth part of the experiment, the computer randomly assigns a type of person to you. **There are two types of persons, employers and employees.**

**<u>Brief overview of part IV of the experiment:</u>**
Part IV of the experiment consists of two stages. The stages are structured as following:

1. **Employers and employees sign an employment contract.** In the first stage employers thus state which wage level they are willing to pay to which employee.
2. In the second stage of this part, each employer is randomly assigned to an employee who, once again, is given 90 seconds to solve the numerical code task of part I for the employer. The number of correct answers determines the earnings of the employer. The employee is paid the wage by the employer.

**<u>Detailed procedure:</u>**
There are 24 participants in this room, i.e. there are exactly 12 employers and 12 employees. On the first screen of this part you are told which type of person you are (employer or employee).

Page 5 of 8

PTB/P/T/PT

1. Determination of wages

In a first step, employers state which wage level they are willing to pay to which of the 12 employees in return for them solving the numerical code task in stage 2 of this part. For this purpose, employers get two pieces of information about **each employee**: The number of correct assignments (**# correct answers – # wrong answers) from part I** of the experiment, and the **decision of the employees that he took being person Y** in the decision situation of part II of the experiment.

Employers state a wage level for each of the 12 employees. Wage levels are entered in a table that looks like the following:


The order of listed employees is random. **Information on each employee's behaviour in [PTB/PT: parts I and II] [P: part I] [T: part II] of the experiment are provided in brackets to the employers.**
**[PTB: If an employee achieved more correct assignments than the average in part one, he belongs to "group yellow". If he achieved less than the average in part one, he belongs to "group blue".] [PT: The number of correct assignments in part one and the decision in part two is given in brackets] [P: The number of correct assignments in part one is given in brackets] [T: The decision in part two is given in brackets]**

For example, the employee [PTB: **(group yellow, right),** PT: **(5, right),** P: **(5),** T: **(right)]** [PTB: achieved more correct assignments than the average in part I] [PT/P: achieved 5 correct assignments in part I] and [PT/T decided on "right" in the decision scenario of part II]. Correspondingly, the employee [PTB: **(group blue, left),** PT: **(13, left),** P: **(13),** T: **(left)]** [PTB: achieved less correct assignments than the average in part I] [PT/P: achieved 13 correct assignments in part I] and [PT/T decided on "left" in the decision scenario of part II].

Page 6 of 8

PTB/P/T/PT

Wage levels should be entered in the box labelled “Your wage offer“. The offer may not be smaller than 0 and exceed 250:

$$\mathbf{0 \leq wage\ offer \leq 250}$$

Employers may enter a different wage level in each box or the same wage level for everyone or for some of the employees. **Employers have to fill in every box, i.e. they are required to make a wage offer to every employee.**
Employers find out which employee is allocated to them only in the second stage of this part. The allocation is done by the computer. An employer can get allocated to any employee.

*<u>Example</u>: An employer offers a wage of 70 to the employee [PTB: (group yellow, right), PT: (5, right), P: (5), T: (right)] and a wage of 130 to the employee [PTB: (group blue, left), PT: (13, left), P: (13), T: (left)]. In case the employer gets allocated to the employee [PTB: (group yellow, right), PT: (5, right), P: (5), T: (right)], the employer is required to pay him a wage of 70 in return for the employee performing the numerical code task.. In case the employer gets allocated to the employee [PTB: (group blue, left), PT: (13, left), P: (13), T: (left)], the employer is required to pay a wage of 130.*

While employers enter their wage offers, employees are asked to state their wage expectations and how certain they are on their expectations.

2. <u>Task stage</u>

**One employee gets allocated to one employer for each task stage. The employee receives a wage in return for performing the numerical code task. Again, employees are given 90 seconds to this end.**

**In this setting, the wage corresponds to the one level offered by an employer to the respective employee in the previous stage.** Earnings of an employer are determined by the number of correct answers net of wrong answers which is achieved by the employee in the numerical code task.

Employees do not find out to which employer they are allocated. Before starting the numerical code task, employees only find out the wage that they get paid.
Employers are shown the information on the employees that got allocated to them on their screens.

While employees work on the numerical code task, employers are asked to state their expectations on the number of correct answers of their employee and how certain they are on their expectations.

## How are earnings determined?

**Earnings of employers:**

- Earnings of employers depend on the number of correct answers of their respective employee (net of wrong answers) as well as on the wage they offered to pay to their employee. Earnings are determined in the following way:

**Earnings of an employer =**
**10 points** * (# correct answers – # wrong answers) – wage

The employer receives 10 points for each correct answer of their employee (net of wrong answers). He is required to pay the wage to the employee from this revenue. Earnings of employers are thus higher the more correct assignments their employee scores. Earnings of employers are lower the fewer correct assignments their employee score.

Page 7 of 8

In case the employee has given more wrong answers than correct ones, revenues of the employer amount to 0 points in the worst case. But in any case, the employer is required to pay the wage to the employee.

**Earnings of employees**

- Earnings of employees are the wages that they receive from their respective employer. Earnings are determined in the following way:

| **Earnings of an employee = wage** |
|---|

Earnings of employees are thus independent of the number of correct and wrong answers in the numerical code task.

Earnings of all employers and employees are determined in the same way. **Consequently, every employee is able to compute the earnings of the employer he works for.**
Please note, that, in principle, it is possible to **incur losses**. You are required to settle losses using your show-up fee or earnings from other parts of the experiment.

*<u>Example 1:</u> An employer offers a wage of 110 points to an employee and is allocated to this particular employee. If the employee achieves 21 correct and 2 wrong answers, the employer's earnings amount to:*

$$10 * (21 – 2) – 110 = 10 * 19 – 110 = 190 – 110 = 80$$

*The employee receives his wage of 110 points.*

*<u>Example 2</u> An employer offers a wage of 80 points to an employee and is allocated to this particular employee. If the employee achieves 35 correct and 0 wrong answers, the employer's earnings amount to:*

$$10 * (35 – 0) – 80 = 10 * 35 – 80 = 350 – 80 = 270$$

*The employee receives his wage of 80 points.*

*<u>Example 3</u>: An employer offers a wage of 200 points to an employee and is allocated to this particular employee. If the employee achieves 15 correct and 4 wrong answers, the employer's earnings amount to:*

$$10 * (15 – 4) – 200 = 10 * 11 – 200 = 110 – 200 = – 90$$

*This loss has to be settled with earnings from other parts of the experiment or the starting balance. The employee receives his wage of 200 points*

You are informed on your earnings as well as the earnings of your partner at the end of part IV on a particular **screen showing your earnings**:

This screen contains the following information:

- Information on the employee from parts I and II of the experiment
- Number of correct assignments (# of correct answers – # wrong answers) of the employee
- Earnings of the employee (wage)
- Earnings of the employer

The experiment does not start until all participants have become familiar with the exact calculations of earnings. For this purpose, we kindly ask you to solve several practice exercises on your screens beforehand. Please raise your hand in case you have any questions.

At the very end of the experiment, the computer calculates your final earnings from parts I to IV and provides you with detailed earning information for each part of the experiment on the screen.

# Chapter 3

# Competition, Cooperation, and Preference Heterogeneity

## 3.1 Introduction

Imagine you are a researcher in economics and about to enter the job market in the near future. You are currently working on two projects and have to decide between the two which one to work on. Both papers still need some work, but you do not have the time to work on both of them at the moment. The first one, however, is co-authored with a colleague that you could work together and the second one is a single-authored paper. Which paper do you decide to work on?

The trade-off you are facing in this decision is to allocate resources between a cooperative and a non-cooperative activity (i.e., putting effort in your joint project or in your own project) given that you are measured against others in a competition with the product of the activity you undertake (i.e., the imminent job market or academic output in general when applying for positions). The cooperative activity in a group is characterized by both your effort having a positive externality on others and their effort having a positive externality on you. This leads to the classical problem of a social dilemma where it is individually rational not to provide effort but socially optimal to do so. In the example, this would be the co-authored paper where it is individually rational for you to let your co-author do all the work, but which would end up being a much better paper if both of you actively collaborated on it. Alternatively, you can allocate resources to a non-cooperative activity, which amounts to working on your own paper.

The proceeds of any activity you undertake (cooperative or not) determine your position in a competition to other people who are in a similar situation. In particular, the proceeds of the (non-)cooperative activity decide about whether you win a prize as the winner of a tournament with other persons that are facing the same trade-off. Referring to our initial example, the final papers will be evaluated against all other papers in your field and only if they are sufficiently well published, you will end up obtaining the position you applied for.

The underlying conflict is one between cooperation (i.e., to defect or not to defect) and competition at the individual level. Competition provides an additional incentive to cooperate if you think that the others in your group will cooperate as well, but keeps freeriding still as a dominant strategy. The marginal incentive to contribute under competition is smaller since contributing now also lowers your chances to win the tournament. The decision in the aforementioned situation ultimately also depends on your general willingness to engage in the cooperative activity (in the presence of competition), i.e., your preferences over outcomes to members in your group and to those not in your group against whom you are competing. Since preferences are heterogenous in the population, it is unclear whether the impact of competition is the same for every individual.

In this chapter, we directly test the impact of competition on cooperative behavior by making two distinct contributions. Our first contribution is to assess the impact of individual competition on behavior and outcomes in social dilemma situations. We thereby refer to a situation where the individual incentives of a social dilemma situation are augmented by the possibility to win an exogenously fixed prize if the product of cooperation exceeds that of a competitor. To the best of our knowledge, little is known about the effects of individual tournament incentives on cooperative behavior and whether this fosters or inhibits cooperation. Individual competition may have ambiguous effects on cooperative behavior: On the one hand, competition may result in an in-group effect (or team spirit) that fosters collaboration in a group and leads to higher contribution levels than in the absence of competition. On the other hand, the marginal incentive to contribute is smaller in the presence of competition since every contribution ceteris paribus deteriorates one's position in a tournament against a competitor outside the group. Our second contribution is to analyze the impact of competition on the distribution of different types of preferences in the population. There is substantial evidence that individuals differ systematically in their contribution preference in social dilemma situations. Our hypothesis is that different types of contributors will react differently to the presence of competition in the form of a rank-order tournament. The experiments which we conduct allow us to assess the reactions of the different types to competition both in one-shot and in repeated interactions. With respect to the latter we are particularly interested in the effects of competition on the incentive to build a strategic reputation.

The workhorse of our analysis is a standard linear public goods game. We extend it by the possibility to win an exogenously fixed prize if an individual's payoff exceeds the payoff of another player from a different group.[1] In a first experiment, subjects are classified according to the methodology from Fischbacher et al. (2001) with respect to their cooperative preferences and interact *repeatedly* in a linear public goods game. Their cumulated earnings determine if they win a tournament against somebody from another group. In a second experiment on *one-shot* behavior, the classification exercise from Fischbacher et al. (2001) itself is augmented by the possibility to win a bonus if an individual's payoff exceeds that of another subject from a different group. In both experiments, we draw special attention to the identification of contribution types and the impact of the tournament on changes in contribution behavior for the respective types identified.

There are numerous attempts to identify different types of contribution behavior in public good environments, from e.g. actual behavior in the game (Weimann, 1994), a test of "Social Value Orientation" (Liebrand, 1984; Liebrand and McClintock, 1988; Offerman et al., 1996; Park, 2000; van Dijk et al., 2002) and the contribution function approach (Brandts and Schram, 2001). (Fischbacher et al., 2001) introduce the concept of conditional cooperation which is now standard in the literature. They employ the strategy method for every possible level of average contribution by all other group members. It is now widely acknowledged that heterogeneity in contribution preferences is a persistent phenomenon and hence has attracted a lot of attention in the literature, especially to explain the typ-

[1] Ledyard (1995), Zelmer (2003) and Chaudhuri (2011) provide surveys on the major findings from the experimental public goods literature.

ical decay in contributions under repetition.[2] Fischbacher and Gächter (2010) study how preference heterogeneity and beliefs interact to explain the widely observed decay in contributions over several rounds by subjects acting as ”imperfect conditional cooperators“. Closest to our study is a paper by Fischbacher and Gächter (2009) who study the validity of the strategy method in public good experiments. They establish a direct link between the preferences elicited in a conditional cooperation exercise and the subsequent behavior in a repeated social dilemma situation to find strong evidence for heterogeneity of types and high levels of consistency of behavior within subjects. We extend their analysis by looking at the effects of strategic behavior in a partner setting and by looking at the (possibly different) reactions of the different types identified on a treatment variation.
All of these studies, however, take preference heterogeneity as given and report the consequences on aggregate outcomes, but draw little attention to how contribution behavior is affected by different institutions and incentives. Furthermore and to the best of our knowledge, there is no paper that explicitly examines the impact of a treatment variation on the specific contribution types identified, i.e., whether the various types identified actually react differently to an exogenously imposed change in the strategic environment. This link between types and institutions has important consequences for setting the right incentives in organizations. If the same incentives have different effects on different types, eliciting information especially about other-regarding concerns at the individual level would be a crucial ingredient for optimal incentive design, see Englmaier and Wambach (2010) and Englmaier et al. (2011). In this chapter, we concentrate on the introduction of competition as the change in the strategic environment since we consider competition an important and omnipresent feature in the interaction of two or more individuals.
The interaction of social preferences and their role in competitive environments has attracted a lot of attention in the literature, see e.g. Schmidt (2010) and Dufwenberg et al. (2011). It is argued that the strategic environment in general and more precisely to what extent individuals can affect the final allocation of payoffs is a crucial ingredient of whether other-regarding preferences have an impact on final outcomes or not, see Fehr and Schmidt (1999). What remains unclear, however, is how individuals with heterogeneous

[2]See e.g. Burlando and Guala (2005); Kurzban and Houser (2005); Gunnthorsdottir and Rapoport (2006); Bardsley and Moffatt (2007); Kocher et al. (2008); Muller et al. (2008); Herrmann and Thöni (2009); Ambrus and Pathak (2011); Kosfeld and von Siemens (2011).

other-regarding concerns react in different strategic environments. Particularly assuming that competition has identical effects across all identified types of other-regarding preferences has a clearly ad-hoc flavor. In the context of a social dilemma, we thus expose individuals to an individual rank-order tournament to assess how different types individually cope with the trade-off between cooperation and competition.

In the framework of public goods games, there are numerous studies that focus on the effects of competition between teams on cooperation and efforts[3]. Competition between sales teams or research groups are examples of tournaments between groups to elicit high levels of efforts and cooperation within the team. Common to these tournaments is that the winning team is rewarded with a prize which is then shared between all members. There are numerous instances, however, where individuals while being in a social dilemma within a group are subject to a competition *as individuals themselves* with the proceeds from the team work. They have to decide how many resources they allocate to an individual and a cooperative activity knowing that the proceeds from collaboration are higher than those from the individual activity if the other group members collaborate and lower if others fail to cooperate.

Despite the prevalence of competitive incentives through rank-order tournaments (Lazear and Rosen, 1981) in the real world for individuals (e.g. promotion tournaments in firms, grade performance in class), the effects of competition at the individual level have not received much attention in the public goods literature up to now. Experimental methods seem particularly appropriate for the analysis, as clean empirical evidence which is able to exogenously control for the presence of competition is scarce and hard to obtain in the field.

We have two main results: In a repeated setting, we find that competition has a detrimental effect on the incentives of freeriders to build a strategic reputation, but has no impact on the behavior of conditional cooperators. For aggregate levels compared to a standard social dilemma, competition between individuals does not foster or inhibit cooperation unlike in settings with competition within teams. We present evidence that competition in a one-shot public goods game à la Fischbacher et al. (2001) leads to a shift of the un-

[3] See e.g. Bornstein et al. (1990); Bornstein (1992); Erev et al. (1993); Bornstein and Ben-Yossef (1994); Bornstein et al. (2002); Bornstein (2003); Tan and Bolle (2007); Sutter and Strassmair (2009); Reuben and Tyran (2010).

derlying distribution of types from conditional cooperators towards more freeriders, but leaves the self-serving bias of the remaining conditional cooperators unaffected.
The chapter is structured as follows: Section 3.2 outlines the experimental design before section 3.3 lays out the hypotheses. Section 3.4 contains the main results and section 3.5 concludes the chapter.

## 3.2 Experimental Design

The core mechanism in our experiments is a standard linear public goods game played in a group of four subjects. Each subject has an intial endowment $E$ that she can decide to either contribute to the public good or keep for herself. The monetary payoff from the baseline public goods game (PGG) $\pi_i^b$ to subject $i$ is given by

$$\pi_i^b = E - c_i + 0.5 \sum_{j=1}^{4} c_j$$

where $c_i$ is the individual contribution. The return of every token contributed to the public good is 0.5, thus leading to a clear prediction of free riding. Under standard assumptions, rational subjects put the entire endowment on the private account. The social dilemma arises from full contribution levels being socially efficient.

### 3.2.1 Implementation of Competition

The way we model competition in both experiments is in the form of a two-player tournament with an exogenously fixed prize. Two subjects from different groups compete for this prize and the competition is implemented through a comparison of their final wealths after having played the public good game described above. If both subjects are tied, a coin flip determines the winner of the prize. The payoff in a one-shot interaction under competition $\pi_i^c$ is then given by

$$\pi_i^c = \underbrace{E - c_i + 0.5\sum_{j=1}^{4} c_j}_{\text{standard PGG}} + \underbrace{p(c_i, \sum_{j=1}^{4} c_j, c_k, \sum_{l=1}^{4} c_l)B}_{\text{incentive from competition}}$$

where $p(\cdot)$ is the probability of obtaining the bonus $B$ and $c_k$ the contribution of the competitor from the other group with aggregate contribution of $\sum_{l=1}^{4} c_l$. As a consequence, subjects are exposed to a "twofold“ incentive in addition to the standard set of incentives arising from the public good game: On the one hand, the marginal incentive to contribute in a symmetric equilibrium with selfish agents holding other's contributions fixed is decreased since every token invested into the public good decreases the individual likelihood of obtaining the bonus. More formally, the marginal incentive to contribute to the public good is given by

$$\frac{\partial \pi_i^c}{\partial c_i} = -0.5 + \underbrace{\frac{\partial p}{\partial c_i}}_{<0} B < \frac{\partial \pi_i^b}{\partial c_i}$$

On the other hand, the benefits of high levels of cooperation in the group are increased since winning the bonus becomes more likely the higher one's final wealth in a group that cooperates well. Furthermore, if subjects have other-regarding concerns towards members of their own group and a team spirit prevails, the presence of the bonus induces contributions which improve the chances of members of one's own group to be awarded the prize. Compared to the standard public goods game with selfish agents, the tournament against outside group members yields an additional mechanism to sustain high levels of cooperation within a group in the presence of non-selfish preferences. A team spirit emerges when other-regarding preferences relate only towards the own group.

It is important to note that the competitor is always member of *another* group such that a subject's contribution to the public account within her own group does not directly increase the payoff of her competitor and hence does not increase the competitor's chance to obtain the bonus per se. The procedure is common knowledge and it is explicitly pointed out in the instructions that every competitor is taking decisions under the exact same conditions than oneself. It is this trade-off between the (possibly) weakening or strengthening effects of competition that we are interested in. We identify cooperation types in the population through the methodology proposed by Fischbacher et al. (2001). Two different experiments are designed to assess the impact of competition both on the (strategic) behavior in a repeated game setting (Experiment 1) and on the preferences elicited under the method of conditional cooperation by Fischbacher et al. (2001) (Experiment 2). We describe the design of both experiments in the next two subsections.

### 3.2.2 Experiment 1: Competition & Strategic Behavior

The experiment consists of two parts, where subjects receive instructions for the second part only after the first part is over and hence do not know what follows in the later part.

1. In the first part of Experiment 1, subjects' contribution types are elicited in groups of four along the protocol from Fischbacher et al. (2001). Subjects are endowed with 20 experimental tokens that they can allocate to either a public or a private account with the parameter specification from above, i.e., a marginal per capita return of 0.5.[4] Everybody in the group is asked to submit an unconditional contribution and thereafter a contribution table where entries have to be made conditional on the average rounded contribution of all three other group members.[5] Both contributions are made incentive compatible in the following way: one randomly selected subject in every group is given the role of a conditional contributor, the three other subjects in the group are allocated the role of unconditional contributors. The average contribution of the three latter subjects is used to determine the payoff relevant contribution of the one conditional contributor per group from her contribution tables. Subjects earn the payoff of the public goods game in part one, but do not learn the result from this part until the end of the experiment. As is standard in the literature, see Fischbacher et al. (2001), subjects are classified into four categories according to their pattern of conditional contribution:

    - **Freeriders:** These subjects never contribute anything to the public good, regardless of the contributions of other group members.

[4] One experimental token is converted at an exchange rate of 0.10 €.

[5] Since contributions are made from an endowment of 20 experimental tokens, 21 choices have to be made.

- **Conditional Cooperators:** These subjects display a positive correlation between their own contribution levels and the average contribution of others in their group.[6]
- **Hump-shaped (or Triangle) Contributors:** Subjects in this category (weakly) increase their contribution up to a certain level of average contributions from which on they decrease it.[7]
- **Others:** The residual category

The conditional cooperation exercise is particularly useful in eliciting underlying preferences from subjects in the absence of strategic uncertainty. Even in a one-shot interaction, an unconditional contribution of a subject is affected by both the preferences *and* the beliefs of a subject about other group member's contributions, see Fischbacher and Gächter (2010). The conditional contribution, however, completely abstracts from beliefs about behavior of others since the contribution is entered taking the others' action as given. As a consequence, the conditional contribution table gives a close picture of a subject's underlying preferences over final outcome distributions.

2. In the second part of the experiment, the same groups of four subjects repeatedly interact in a partner design for 10 rounds. In every round, subjects enter their unconditional contribution to the public good. Full feedback is given after every round, i.e., subjects learn about the contributions of the other group members, albeit in a way that precludes identification of subjects across rounds.

   We conduct 2 treatments, (**COMP**) and (**BASE**). In the competition treatment (**COMP**) the final wealth of subjects after 10 rounds decides about the allocation of the prize. More precisely, every subject is randomly matched with a subject from a different group such that no rivalry between two distinct groups emerges.

[6] We take the Spearman rank correlation coefficient as a criterion to classify a subject to be a conditional cooperator, i.e., positive and $p < 0.01$. For a robustness check, we also take the somewhat more restrictive "weak monotonicity rule" rule as a criterion according to which a subject is classified as a conditional cooperator if her contributions are weakly increasing (and never decreasing) in the contributions of all other team members. This yields identical results.

[7] Note that in the classification method through the Spearman rank correlation, it can happen that subjects get classified as conditional cooperators and hump-shaped contributors at the same time. In that case, we classify them in the former category, but this applied to only 4 out of 192 subjects.

After 10 rounds, the final wealths of the two subjects are compared with each other and the subject who has earned more in accumulated earnings is awarded a prize of 70 experimental tokens (equivalent to 7 €).[8] Subjects do not get any feedback about contributions or earnings of their competitor from the other group during the course of the experiment. In the baseline treatment (**BASE**), everything is identical to COMP except that the prize is awarded randomly through the public roll of a dice between two subjects from two different groups.

In addition to the contribution decision, we elicit subjects' expectations about their position in the competition against their competitor every round. To do so, subjects have to answer the following question every round by entering a percentage number on the screen: "How likely is it that you have higher accumulated earnings than your competitor *up to the current round?*". There is no feedback given about the performance of the competitor in the other group within rounds. Payoffs to subjects are made as follows:

$$\pi_t = \begin{cases} (\text{probability entered}) \cdot 40\text{EP} & \text{if cumulated earnings until t are higher} \\ (100 - \text{probability entered}) \cdot 40\text{EP} & \text{if cumulated earnings until t are lower} \end{cases}$$

One period is drawn at random for payoff and within that period, either the contribution decision or the expectation elicitation is paid out with a 50 % chance. If the contribution decision is selected, subjects earn the proceedings from the public good game in that round; if the expectation elicitation is selected in round $t$, correctness of expectations is rewarded according to the above mechanism. Hence, subjects can at most earn 40 experimental tokens if they enter 100 % (or 0 %) and they indeed have higher (lower) cumulated earnings than their competitor in that specific round. The way payments are implemented for part two is done to discourage subjects from hedging between the contribution decision and the elicitation of expectations. To keep everything identical between treatments, subjects are asked to indicate their expecations about the comparison to their matched subject also in BASE, even if the prize is allocated randomly at the end of the 10 periods. At the end of the ex-

[8]This is done to keep marginal incentives from competition high over all ten periods thus prohibiting a "dilution" effect of the bonus incentive if a smaller prize is attributed every period.

periment, subjects are told their earnings and whether they are awarded the bonus or not. In the BASE treatment, a dice is publicly rolled to determine the winner of the prize whereas in COMP cumulated payoffs of their competitor are shown to subjects on the computer screen.

It is important to note that the elicitation of types in part one is identical across the two treatments, since subjects learn only afterwards about part two. At the end of the experiment, subjects have to answer a socio-demographic questionnaire. Overall, 192 subjects participated in the experiments, i.e., 24 groups of four subjects in every treatment. The experiments were run between September and November 2010 in the MELESSA laboratory of the University of Munich and were computerized via z-Tree (Fischbacher, 2007). Subjects were recruited via ORSEE (Greiner, 2004). The sessions lasted roughly 45 minutes and subjects received 12.7 € on average including a show-up fee of 4 €.

### 3.2.3 Experiment 2: Competition & Conditional Cooperation

From part one of the experiments on repeated interactions (Experiment 1), we have a classification of participants into types according to the classification proposed by Fischbacher et al. (2001). This classification does not tell us anything about the impact of competition on the distribution of types since subjects learn only about the competition in part two of the experiment and can adjust their behavior. Hence, we take the classification of an individual in part one of Experiment 1 as the baseline behavior in the outlined public goods game and conduct an additional one-shot experiment to control for the impact of competition at the individual level on *conditional* cooperation. To do so, we expose participants to the same form of competition as in Experiment 1 also in the conditional cooperation exercise. Two subjects from two different groups are matched and their payoffs after the conditional contribution tables are compared to determine who is awarded the additional prize of 70 experimental tokens. If the payoff to one subject exceeds the payoff to her competitor, she is awarded the bonus. For the payoffs from the public good mechanism, one subject is chosen at random for whom the contribution tables are payoff relevant, whereas for the three others in a group the unconditional contribution is relevant. The overall payoff from the conditional cooperation exercise under competition is therefore given by

$$\pi_i = \begin{cases} 20 - c_i + 0.5\sum_{j=1}^{4} c_j & \text{if} \quad 20 - c_i + 0.5\sum_{j=1}^{4} c_j < 20 - c_k + 0.5\sum_{l=1}^{4} c_l \\ 20 - c_i + 0.5\sum_{j=1}^{4} c_j + 0.5 \cdot 70 & \text{if} \quad 20 - c_i + 0.5\sum_{j=1}^{4} c_j = 20 - c_k + 0.5\sum_{l=1}^{4} c_l \\ 20 - c_i + 0.5\sum_{j=1}^{4} c_j + 70 & \text{if} \quad 20 - c_i + 0.5\sum_{j=1}^{4} c_j > 20 - c_k + 0.5\sum_{l=1}^{4} c_l \end{cases}$$

where a coin flip decides about the allocation of the bonus in case of a tie. The procedure is again made common knowledge to all subjects such that we can compare the results to parts one of the experiments in the repeated game described above. An additional 48 subjects that earned on average 10.1 € (including a show-up fee of 4 €) participated in these experiments which took roughly 30 minutes.

## 3.3 Hypotheses

In this section, we assess theoretically how different types in the population react differently to the introduction of competition at the individual level. We sketch a very simple model of the effects from the introduction of a competitive rank-order tournament on two different types of subjects, selfish and non-selfish individuals. In line with the simple linear altruism model of Ledyard (1995), we assume an individual's utility $u_i$ to take the following form in the context of the one-shot standard linear public goods game outlined above:

$$u_i = \pi_i + \alpha g(\pi_j)$$

$\pi_i$ denotes the material payoff to player $i$ and $\alpha$ is a parameter that captures other-regarding concerns in the form of simple altruism. $\alpha$ is zero for selfish subjects and strictly positive for individuals with other-regarding concerns. $\pi_j$ is the representative monetary payoff of another group member j and $g(\cdot)$ some generic utility function over monetary payoffs of this member with $\frac{\partial g}{\partial \pi_j} > 0$. We consider this representative group member as the only social reference towards which a player exhibits other-regarding concerns and disregard such feelings for all subject outside the own group, including the competitor in the tournament. An interpretation would be some sort of team spirit within the own group. To assess the effect of competition, we compare the marginal incentive to contribute in both treatments. For $\alpha = 0$ and our parameters for the public goods game, we have in the BASE treatment

$$\frac{\partial u_i}{\partial c_i} = -0.5$$

In the competition treatment, we have

$$\frac{\partial u_i}{\partial c_i} = -0.5 + p_i'(c_i)B$$

where $p(\cdot)$ denotes the probability of obtaining the bonus of size $B$. The marginal incentive to contribute for an individual with selfish preferences is lower under competition since $p_i'(c_i) < 0$. For individuals with $\alpha > 0$, the marginal incentive to contribute in the BASE treatment is given by

$$\frac{\partial u_i}{\partial c_i} = -0.5 + \alpha \frac{\partial g}{\partial \pi_j} \frac{\partial \pi_j}{\partial c_i} = -0.5 + 0.5\alpha \frac{\partial g}{\partial \pi_j}$$

and in the COMP treatment

$$\frac{\partial u_i}{\partial c_i} = -0.5 + p_i'(c_i)B + \alpha \frac{\partial g}{\partial \pi_j} \frac{\partial \pi_j}{\partial c_i} = -0.5 + p_i'(c_i)B + \alpha(0.5 + p_j'(c_i)B) \frac{\partial g}{\partial \pi_j}$$

Ceteris paribus a contribution from player i decreases her probability to win the prize, but increases the probability of the representative group member as player j to obtain a bonus in his respective tournament, i.e., $p_j'(c_i) > 0$. We can now establish a condition for $\alpha$ where the decrease in the marginal incentive to contribute through competition is outweighed by other-regarding concerns.

$$\alpha^* = -\frac{p_i'(c_i)}{p_j'(c_i)\frac{\partial g}{\partial \pi_j}}$$

For $\alpha^*$, the marginal incentive to contribute for a subject with other-regarding preferences is identical between the two treatments. As an example, if $|p_i'(c_i)| = |p_j'(c_i)|$ an $\alpha$ of $\frac{1}{\frac{\partial g}{\partial \pi_j}}$ leaves a non-selfish subject to have identical marginal incentives to contribute in both treatments. Given the increase in utility from the chances a contribution creates for another group member to win the prize from the team spirit effect, we expect less of a difference in contribution levels for individuals with other-regarding concerns than for selfish subjects.

**Hypothesis 3.1 (General Effects of Competition)** *Competition has a detrimental effect on the marginal willingness to cooperate for subjects classified as selfish, but induces less change in the marginal incentive to contribute for non-selfish individuals.*

The outlined mechanism refers to a one-shot standard linear public goods game but is agnostic about both the incentives in a repeated setting (as in Experiment 1) and in a conditional cooperation exercise (as in Experiment 2). We therefore further elaborate on the relevant effects of competition in both of our Experiments 1 and 2 in turn.

With respect to repetition in a partner design as in Experiment 1, there is both theoretical and experimental evidence that especially selfish subjects have an incentive to mimic cooperative types by initially providing high levels of contributions before revealing their type towards the end of the game, see e.g. Kreps et al. (1982) and Roe and Wu (2009). So even if the argument laid out above would rationalize differences in the marginal willingness to contribute for selfish subjects in the presence or absence of competition, it still leaves freeriding the dominant strategy in one-shot games for selfish types, regardless of the strategic environment. If, however, the propensity to contribute is increased through a mechanism to build up a strategic reputation for selfish types, contribution levels may well be positive also for freeriders. Most importantly, if the incentive to contribute for strategic reasons is identical across the two treatments, we should observe higher contribution levels in the BASE treatment than in COMP for selfish individuals. However, we are agnostic about the size of the reputation effect such that we cannot make a prediction for non-selfish subjects about potential absolute levels of contributions. Since absolute contribution levels of non-selfish individuals depend on $\alpha$ and they are not in need of building up a reputation, we conjecture no difference in contribution behavior across treatments for them in Experiment 1.

**Hypothesis 3.2 (Competition and Strategic Behavior)** *Competition hampers the incentive to build up a strategic reputation for selfish individuals in a repeated interaction, but leaves contribution levels unchanged for non-selfish subjects.*

Concerning the predictions of our Experiment 2, we refer to an argument brought forward by Fehr and Schmidt (1999) on the effects of social preferences under competition. In line with their reasoning and as a general finding, other-regarding concerns of individuals matter if they have the possibility to influence the final payoff allocation, but play less of a role when the ultimate outcomes are beyond their control.[9] This explains why models of social preferences may yield powerful predictions when agents, e.g., bargain over an outcome, but less so when competition exempts subjects from the possibility to influence the ultimate payoff distribution as e.g. on competitive markets, see also Fehr and Schmidt (1999), Bolton and Ockenfels (2000) and Schmidt (2010).

As we have argued above and with respect to Experiment 2, the classic elicitation method for conditional cooperation measures underlying *preferences* independent of beliefs about behavior of others in the group. Adding competition, this does not hold anymore since conditional contributions now depend on the beliefs about what the competitor for the prize does and more generally what the contribution levels in the rival's group are. Put differently, the introduction of competition implies for a subject that by entering conditional contributions she does not have full control over final payoff distributions anymore as would be the case in the absence of competition. Our experiment can therefore be interpreted as an experimental test of this hypothesis with an identification of the effect what happens when decision makers lose the possibility of deciding over final outcomes with certainty. We therefore argue that this loss of control over final outcome allocations under competition promotes selfish behavior on behalf of subjects.

**Hypothesis 3.3 (Competition and Conditional Cooperation)** *Competition promotes selfish inclinations in the conditional cooperation exercise compared to the elicitation task from Fischbacher et al. (2001).*

We have little guidance from economic theory, however, as to how the increased selfishness under competition is documented at the individual level, which hence remains an empirical

[9] see Fehr and Schmidt (1999), p. 856

question that we tackle in Experiment 2. There are two possible ways competition can affect subjects: First, the distribution of types in the population can be affected by the presence of competition. More precisely, the chance of obtaining the prize may promote egoistic inclinations among subjects. Second, conditional cooperators may increase their self-serving bias that is usually found in experiments.[10]

## 3.4 Results

We first look at the results from the repeated interaction from Experiment 1, before moving to the effects of competition on conditional cooperation in Experiment 2.

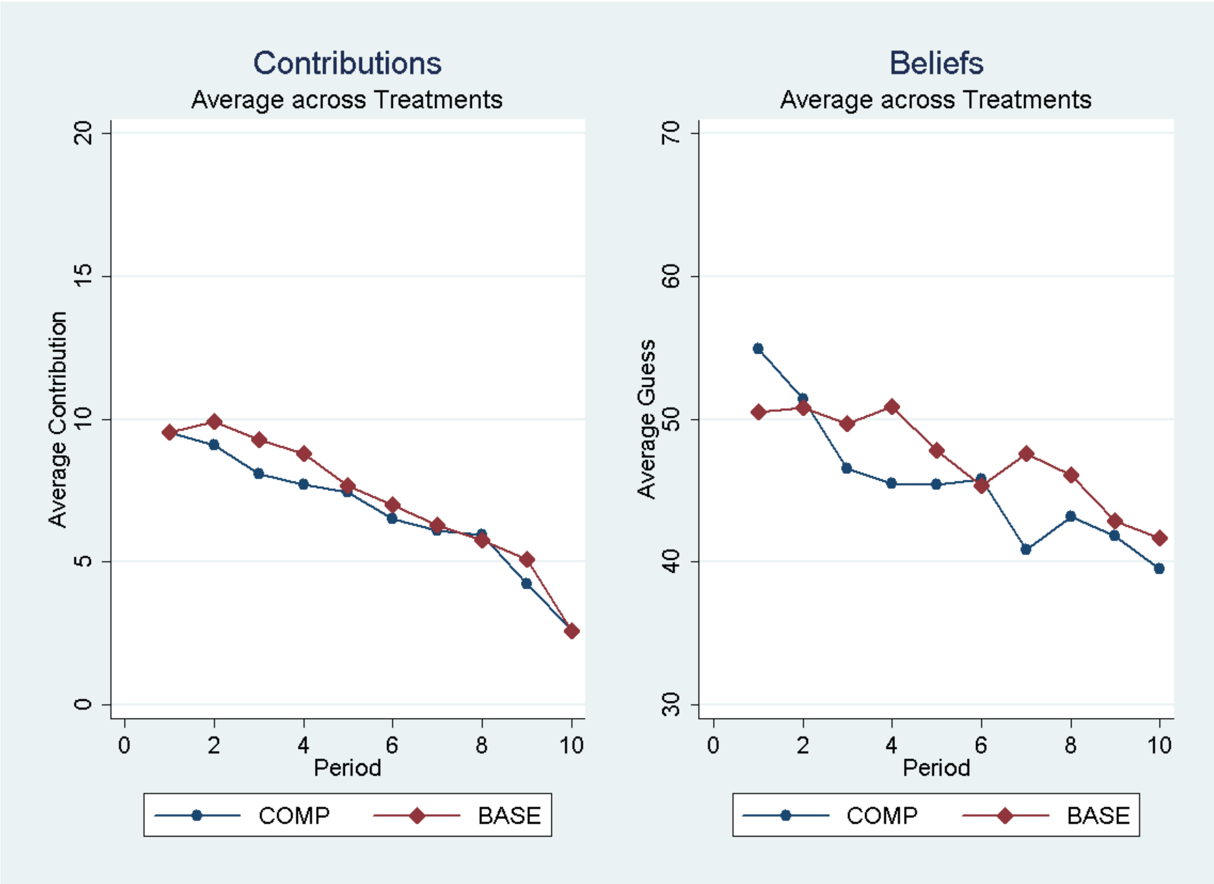
### 3.4.1 Competition & Strategic Behavior (Experiment 1)

**Aggregate Behavior** We start by reporting the aggregate results from the repeated public goods game. In terms of average contributions, the left panel of Figure 3.1 clearly shows that there are no discernible differences between the two treatments with and without competition. This is confirmed by non-parametric tests for averages across all periods and for contributions in each period separately (Mann-Whitney-U-test, $p > 0.33$). We observe the standard decay in contributions over the 10 periods in both treatments, BASE and COMP. Subjects learn about the institutional environment (competition vs. no competition) before they enter their contribution level in the first period, but there is clearly no difference in the contribution levels for period one on the subject level (Mann-Whitney-U-test, $p = 0.86$). Subjects start on average with contributing about 50 % of their endowment in both treatments, which is consistent with the findings in the literature[11]. On the aggregate level, the introduction of competition for the prize hence has neither a positive nor a negative effect on contribution levels. Generally speaking, the decrease in the marginal incentive to contribute seems to be outweighed by an increased incentive to behave cooperatively within the group to enhance every member's chances to win the prize. In contrast to competition between teams which generally increases

[10]The self-serving bias in this context refers to the fact that many subjects classified as conditional cooperators contribute a little less than the perfect match of the average contribution of others.

[11]See Marwell and Ames (1981), Isaac et al. (1984), Isaac and Walker (1988).

Figure 3.1: Contribution Levels and Beliefs in BASE and COMP




Average Contribution Levels and Beliefs in treatments BASE and COMP across all subjects

the level of cooperation (Bornstein and Ben-Yossef, 1994), competition at the individual level does not foster cooperation. As far as expectations are concerned, we ask subjects about the assessment of their chances in % whether they have a higher cumulated wealth than their competitor. The average guess is plotted in the right panel of Figure 3.1. Note that if beliefs are correct, the average of all beliefs should be exactly 50 %. All subjects' assessment starts out at roughly 50 % before their guess declines to about 40 % at the end of the 10 periods. This holds true for both treatments where the decline is highly significant over the 10 periods (Wilcoxon-signed-ranks-test, $p < 0.01$ for both treatments)[12]. The decline in expectations sheds an interesting light on the decay of contributions. Even if subjects are aware of the fact that the payoffs in their group decrease over time, they on average think that the decline in contributions is not happening in the group of their

[12]Group averages in period 1 and 10 as tested variable. All p-values in this chapter are two-sided unless otherwise indicated.

competitor and hence lower their beliefs, i.e., the competitive assessment. This piece of evidence characterizes a certain confession that a more cooperative strategy would have been possible in the group, but did not materialize. Since the presence of the prize does not affect aggregate contribution levels across treatments, the competitive assessments do not differ either (Mann-Whitney-U-test, $p = 0.58$), which leads us to our first result.

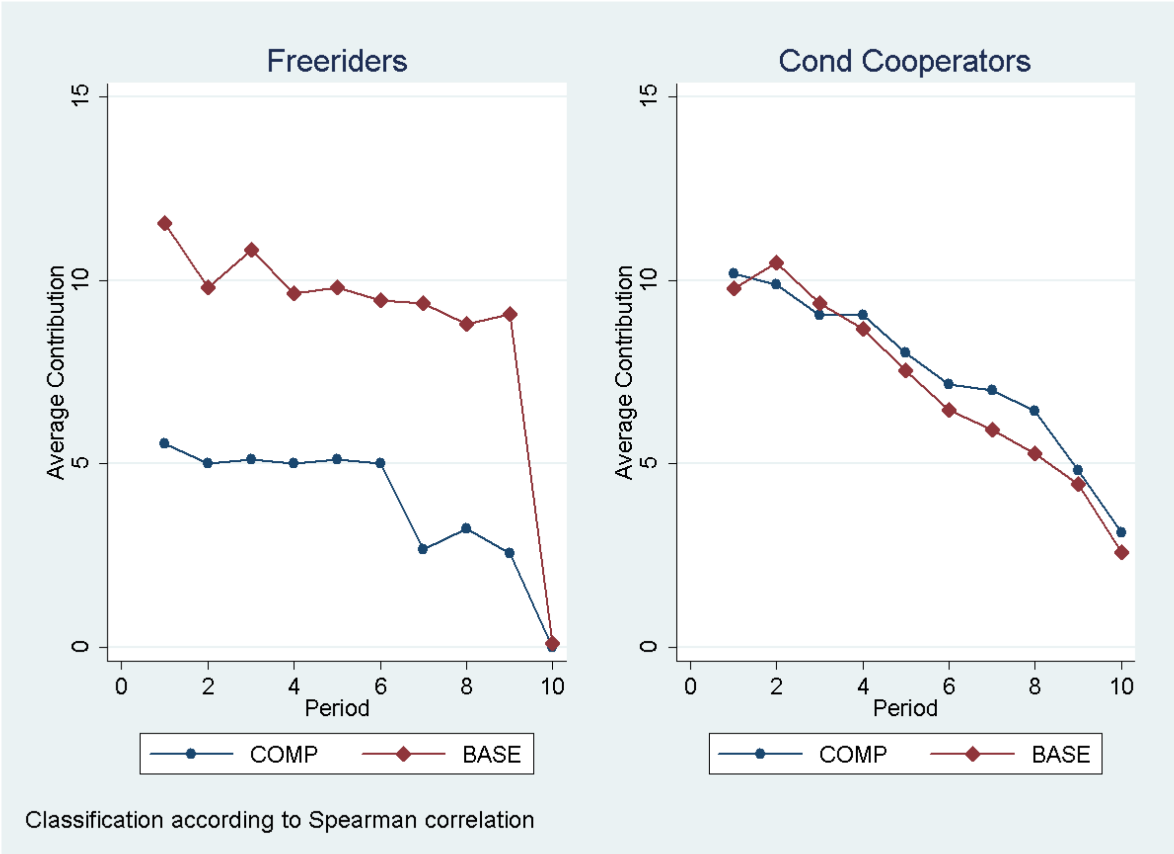
**Result 3.1** *Competition does not affect aggregate contribution levels and the competitive assessment of subjects.*

The results on aggregate behavior above rely on the behavior of groups of four subjects, such that they do not tell us anything about the different impact of competition on the various types of contribution types identified, which is what we turn to in the next section.

**Individual Behavior** We are particularly interested in the effects competition has on the different types of contributors identified. In section 3.3 we particularly conjectured that there is an asymmetric reaction to competition from the different types of subjects identified from the conditional cooperation exercise à la Fischbacher et al. (2001) conducted before the repeated interaction. When there is scope for strategic interaction over more than one period, the introduction of competition may have adverse effects on the different types through the mechanism to strategically build a reputation for selfish agents. We present contribution levels for freeriders and conditional cooperators in Figure 3.2.[13] Fischbacher and Gächter (2010) have shown that the preferences elicited in a conditional contribution exercise have high predictive power for the behavior in a repeated interaction of a standard public goods mechanism. What remains unclear is to what extent the different types alter their behavior when the strategic environment changes, i.e., competition for the prize is introduced and there is scope for reputation. For subjects classified as conditional cooperators in the right panel, contribution levels are almost identical between the two treatments. As far as freeriders are concerned (left

[13]We find about 74 % of all subjects to be conditional cooperators, about 10 % to be freeriders, 7 % to be triangle contributors and 9 % to be in the residual category. In absolute numbers across both treatments, we have 20 freeriders, 142 conditional cooperators, 13 triangle contributors and 17 subjects in the residual category.

Figure 3.2: Contribution Levels for Freeriders and Conditional Cooperators




Average Contribution levels for freeriders (left panel) and conditional cooperators (right panel) in treatments BASE and COMP

panel), competition leads to a severe drop in contribution levels by 50 %. There is also clear evidence for the presence of freeriders building a reputation of being cooperative as their contribution drops to zero in the last round, see Kreps et al. (1982). However, the incentive to build a cooperative reputation for freeriders is considerably reduced in the presence of competition which confirms Hypothesis 3.2.[14] This further qualifies the first result: Since the proportion of conditional cooperators is very high at 74 % and their behavior does not chance across the two treatments, we do not find an effect on aggregate contribution levels.

As conjectured in section 3.3, competition has thus different effects on the different types identified which is confirmed by a series of regressions in Table 3.1. We do not find any simple effect on the contribution level neither from being a conditional cooperator nor

[14]When we pool all subjects that are not classified as conditional cooperators (see Figure A3.4 in appendix A3.3) we equivalently observe decreased contributions under competition.

Table 3.1: Random Effects Panel Regression on Contributions

| **Dep. Var.:** Contribution | **I** | **II** | **III** | **IV** | **V** |
|---|---|---|---|---|---|
| Period | -0.713*** | -0.713*** | -0.661*** | -0.758*** | -0.703*** |
| | (0.064) | (0.063) | (0.068) | (0.069) | (0.075) |
| COMP | -0.429 | -2.839 | -2.699* | -4.923 | -4.256 |
| | (1.166) | (1.792) | (1.499) | (3.288) | (2.859) |
| Cond Cooperator | 1.167 | -0.504 | -0.412 | -1.791 | -1.424 |
| | (0.721) | (1.028) | (0.819) | (2.724) | (1.693) |
| COMP * Cond Cooperator | | 3.258** | 3.103*** | 5.342** | 4.664* |
| | | (1.304) | (1.112) | (2.724) | (2.393) |
| Profits (t-1) | | | 0.187*** | | 0.211*** |
| | | | (0.042) | | (0.047) |
| Beliefs (t-1) | | | -0.010 | | -0.008 |
| | | | (0.007) | | (0.008) |
| Constant | 10.225*** | 11.478*** | 6.496*** | 13.012*** | 7.011*** |
| | (1.252) | (1.529) | (1.281) | (1.282) | (1.994) |
| Obs | 1920 | 1920 | 1728 | 1620 | 1458 |
| $R^2$ | 0.10 | 0.11 | 0.19 | 0.12 | 0.22 |

Random Effects Panel Regression on contributions. COMP and Cond Cooperator are dummy variables, Profits (t-1) are taken from the previous period. Clustering on the group level. Specifications IV and V only for subjects classified as conditional cooperators and freeriders. Standard errors in brackets. *** represents significance at p=0.01, ** at p=0.05 and * at p=0.10.

from the competition treatment. The highly significant interaction term between the two illustrates the different effects competition has on the preference types identified. Conditional cooperators contribute about three tokens more than all other preference types under competition. This is robust to controlling for beliefs in the form of the competitive assessment and for past profits which have a significantly positive impact on the level of subsequent contributions.[15]. The significant interaction term in specifications II and III confirms the asymmetric impact of competition on the different types in the population with conditional cooperators contributing significantly more compared to all other subjects. In particular when we restrict the sample to only those subjects classified as either conditional cooperators or freeriders (specification IV and V), we find that the incentives for freeriders to build a strategic reputation of being non-selfish is negatively affected by the presence of competition.[16] When we look at first round contributions of freeriders and

[15]Even if freeriders have somewhat higher beliefs about their chances to win the bonus than conditional cooperators (COMP vs. BASE: 61 % vs 51 % for freeriders, 43 % vs 45 % for conditional cooperators), beliefs do not have a significant influence on contribution levels, see Figure A3.2 in appendix A3.3 for the evolution of beliefs of the two types per treatment.

[16]Panel tobit regressions yield similar results and are hence relegated to appendix A3.2.

compare them with the unconditional contribution level of the preference elicitation task in part one of the experiments, we can document this decreased incentive to build a reputation for freeriders also non-parametrically. There is no difference in contributions under competition (Wilcoxon-signed-ranks-test, $p = 0.32$), but we find a significant increase in first round contributions through the reputation mechanism in the BASE treatment (Wilcoxon-signed-ranks-test, $p = 0.03$).[17] We summarize these findings in our next result.

**Result 3.2** *Competition asymmetrically affects contribution types in the population. It particularly weakens the incentive to build a cooperative reputation for freeriders, but leaves conditional cooperators unaffected.*

We next take a look at whether contributing less for strategic reasons under competition is a profitable strategy for freeriders. Since all other group members react to a group member contributing little by also lowering their contributions, contributing less is not necessarily a profit increasing strategy. Figure A3.3 in appendix A3.3 reports that there are no different profit levels in the COMP treatment between types. In the BASE treatment, conditional cooperators earn somewhat less than all other subjects, which is significant at the 10 % level. This is confirmed by a random effects panel regression on profits in the public goods game (reported in Table A3.4 in appendix A3.2) where we do not find a treatment effect on profits and especially no significant effect on the profit levels of conditional cooperators in the COMP treatment. This holds similarly also for freeriders such that even though they reduce their contributions under competition, this does not have a significantly positive impact on their earnings. Accordingly, there is no signficant difference in overall efficiency in the public goods game between the two treatments (Mann-Whitney-U-test on cumulated average earnings per group, $p = 0.82$). In a probit regression on the likelihood to obtain the bonus, we do not find any evidence for differences between types. The treatment dummy, the conditional cooperator dummy and the interaction are jointly not significant ($p = 0.56$). When we restrict ourselves to only subjects that have been classified as conditional cooperators or freeriders, the

[17]When comparing contribution levels between subjects in round one between the different treatments, we clearly find no difference for conditional cooperators (Mann-Whitney-U-test, $p = 0.67$). For freeriders, we also fail to reach significance at $p = 0.15$, which is likely to be due to the small sample size for between subject comparisons for freeriders only.

interaction of the three dummies becomes marginally significant ($p = 0.06$) indicating a slightly higher chance for freeriders to be awarded the prize under competition.

To further illustrate our finding that the incentive to build a reputation is weakened for freeriders in the presence of competition, we display our results next to those of Fischbacher and Gächter (2009) who conduct an experiment similar to ours that aims at assessing the consistency of subjects between the preference elicitation task (conditional cooperation) and behavior in a repeated public goods game. They also elicit contribution types through the method by Fischbacher et al. (2001), but have a MPCR of 0.4 (compared to 0.5 in our experiments) and conduct the repeated game in a random matching setting (compared to a partner matching in our case).[18] Their main finding is a robust

Table 3.2: Mean Contributions for Different Types

| | **Fischbacher and Gächter (2009)** (random matching, $MPCR = 0.4$, no competition) | | | **BASE** (partner matching, $MPCR = 0.5$, no competition) | | | **COMP** (partner matching, $MPCR = 0.5$, competition) | | |
|---|---|---|---|---|---|---|---|---|---|
| | Period 1 | Period 10 | All Periods | Period 1 | Period 10 | All Periods | Period 1 | Period 10 | All Periods |
| Freeriders | 4.88 | 0.88 | 2.49 | 11.54 | 0.09 | 8.84 | 5.56 | 0.00 | 3.92 |
| Conditional Cooperators | 8.61 | 2.81 | 5.64 | 9.79 | 2.58 | 7.05 | 10.19 | 3.12 | 7.47 |
| Hump-Shape | 9.06 | 1.29 | 4.88 | 4.00 | 4.00 | 5.36 | 10.80 | 0.80 | 6.07 |
| Others | 9.43 | 3.36 | 5.66 | 7.10 | 4.70 | 6.80 | 6.57 | 3.42 | 3.81 |

Mean Contributions for the four types identified in Period 1, Period 10 and across all periods. Listed for the results from Fischbacher and Gächter (2009) and the two treatments BASE and COMP.

relationship between the preferences elicited through the strategy method and subsequent behavior in the public goods game in the sense that freeriders contribute significantly less than conditional cooperators and all other types of contributors, see Table 3.2.[19] Since their experiments are run in a stranger design, reputational concerns are absent for subjects. When we compare their results with our BASE treatment, we immediately see the impact of the partner design for freeriders who contribute on average 8.84 tokens compared to 2.49 tokens in the Fischbacher and Gächter (2009) experiments. The behavior of conditional cooperators is not affected. Adding competition in COMP, the contribution

[18] As a further difference, we also disclose all contribution levels of every group member in every round, whereas Fischbacher and Gächter (2009) only report the group average every period.

[19] Brandts and Charness (2011) survey the literature and compare the strategy method and the direct response method to find that there are no differences between the results from both methods.

level of freeriders is decreased to a level very similar to that under the stranger treatment of Fischbacher and Gächter (2009). In addition, all four types under competition almost identically match the contribution behavior of the four types from the Fischbacher and Gächter (2009) experiment without reputation. We take this as a further piece of evidence that the introduction of competition has detrimental effects on the strategic incentives to build a reputation, which is particularly an issue for freeriders. The absence of reputation mechanisms in the framework of public goods has therefore similar effects as the introduction of competition on the different types.

As a consequence, the predictive power of the type elicitation procedure by employing the strategy method crucially depends on the strategic and institutional environment of players in the respective public goods game. As Fischbacher and Gächter (2009) have shown, in the absence of reputation type elicitation procedures have a high behavioral validity. If strategic concerns matter as in our BASE treatment, this validity vanishes completely, most notably for the distinction between freeriders and conditional cooperators. The comparison of our results and the results from Fischbacher and Gächter (2009) shows that in environments as in our COMP treatment with tournament incentives, the presence of strategic concerns affects the different types asymmetrically and produces behavior that is at first sight identical to a setting without reputational concerns. For certain types of preferences (like freeriders in our case), a strategic incentive may therefore be counterbalanced by the impact of the institutional environment.

As a consequence, when considering type heterogeneity for setting the right incentives, it is essential to take into account the interaction between identified types and the institutional framework to predict behavior. We sum this up in our next result:

**Result 3.3** *The predictive power of type elicitation procedures for observed behavior crucially hinges on types and incentives from institutional environments being substitutes or complements.*

The main elicitation procedure in the framework of social dilemmas to identify types and preferences is the outlined method by Fischbacher et al. (2001). We have seen in a repeated environment that competition has diverse effects on different types from this method, but are agnostic about the impact of competition when strategic considerations are absent, i.e., on the elicitation method itself. We therefore turn to the effects of competition on subjects' underlying preferences in the form of conditional cooperation, i.e., on behavior in a one-shot game in the absence of strategic concerns.

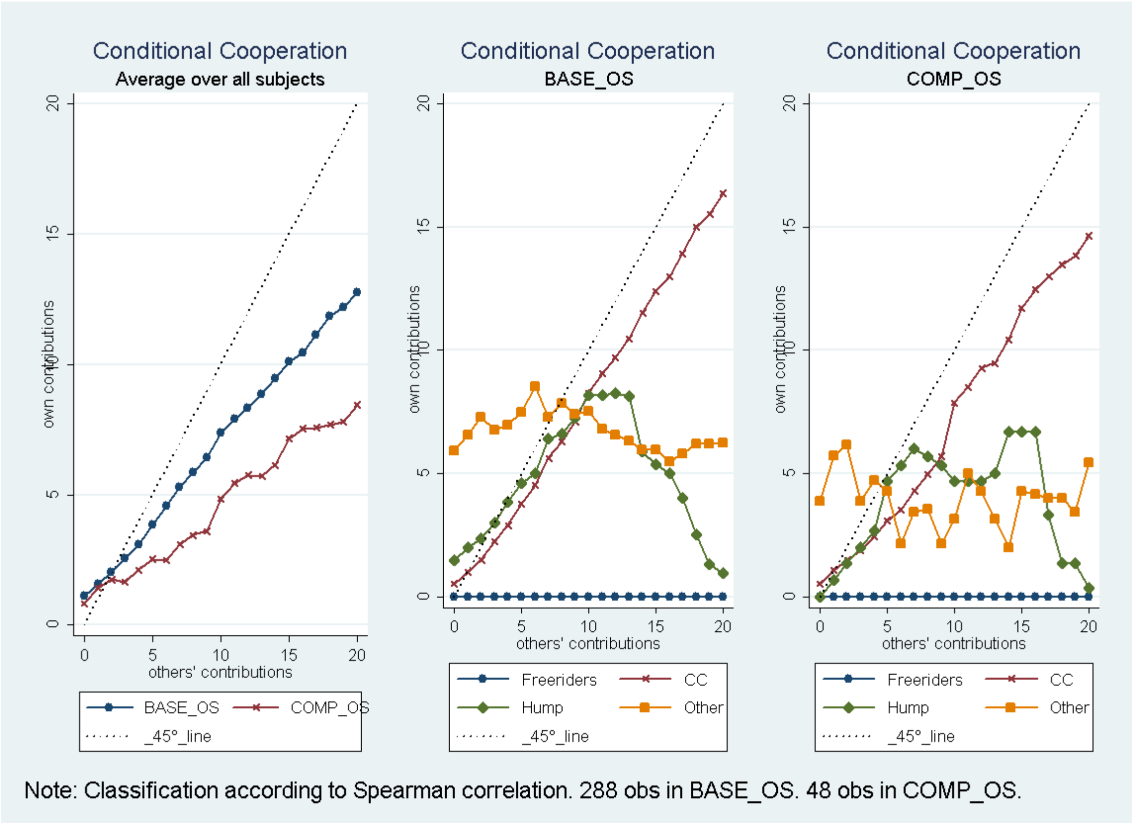
### 3.4.2 Competition & Conditional Cooperation (Experiment 2)

First, we analyze the results from the classification of subjects into types from the conditional cooperation exercise. To do so, we compare the contribution behavior of subjects in the conditional cooperation exercise with competition (from Experiment 2, treatment **COMP-OS** for one-shot) and without competition (part one from Experiment 1, treatment **BASE-OS**). For the analysis on one-shot interactions, we have 192 observations in BASE-OS from part one of Experiment 1 as described above and 96 observations from another experiment that was identical in the structure to Experiment 1 but did not have a modification until part two, such that we can use the data in exactly the same way as the data from Experiment 1.[20] In total, we thus have 288 independent observations in the BASE-OS treatment and 48 in the COMP-OS treatment.

**Preferences for types identified** When we look at aggregate contribution levels for all subjects, we find - not surprisingly and consistent with the literature - a positive slope of the conditional contribution function in the BASE-OS treatment (see left panel of Figure 3.3). In the COMP-OS treatment this slope is considerably reduced compared to the base treatment. For all average contribution levels by the other group members which are higher than 3 EP, we find significant differences between the two treatments ($p < 0.01$, Mann-Whitney-U-test). This difference can stem either from a different contribution behavior in every identified type group or from a shift in the distribution of types. To further explore the issue, we display the contribution functions for every type group separately for the two treatments in the middle and right panel of Figure 3.3. The classification of types is done as described in section 3.2. We compare contribution levels between the two treatments at the individual level with a Mann-Whitney-U-test for each

[20]Results from later parts of this additional experiment are not reported in this chapter. To exclude any effect from including the additional data in the analysis, we report in appendix A3.1 all analyses also without the additional 96 observations. In short, there are no differences in the reduced dataset compared to the results we report here.

Figure 3.3: Conditional Contribution




Aggregate levels of conditional contribution for all subjects (left panel) and contribution patterns for determined types in treatments BASE-OS and COMP-OS (middle and right panel)

type group. Table 3.3 lists the significance levels of these pairwise tests. For conditional cooperators, we cannot identify major changes in behavior and especially no increase or decrease of the self-serving bias at the 5% level.[21] In a similar vein, subjects classified as humpshape contributors or "others" do not contribute more or less under competition than in the control treatment.

The elicited contribution patterns within all four groups of different types identified are robust to the introduction of competition. The fact that contribution patterns of those subjects classified as conditional cooperators are not affected by the possibility to win the prize is strong evidence that the classification method yields indeed a stable picture of subjects' underlying preferences. We sum this up in the next result:

[21]We also do the same exercise for the classification of conditional cooperators according to the weak monotonicity rule and find identical results. We take this as evidence that our results do not rely on the way subjects are classified into different type groups.

Table 3.3: Differences in Conditional Contribution

| Others' Contribution | All Subjects | Conditional Cooperators | Humpshaped | Other |
|---|---|---|---|---|
| 0 | - | - | - | - |
| 1 | - | - | - | - |
| 2 | - | - | - | - |
| 3 | ** | - | - | - |
| 4 | *** | - | - | - |
| 5 | *** | - | - | - |
| 6 | *** | * | - | *** |
| 7 | *** | * | - | * |
| 8 | *** | * | - | ** |
| 9 | *** | * | - | ** |
| 10 | *** | - | - | * |
| 11 | *** | * | - | - |
| 12 | *** | - | - | * |
| 13 | *** | - | - | - |
| 14 | *** | - | - | * |
| 15 | *** | - | - | - |
| 16 | *** | - | - | - |
| 17 | *** | - | - | - |
| 18 | *** | - | - | - |
| 19 | *** | - | - | - |
| 20 | *** | - | - | - |
| obs Total | 336 | 237 | 20 | 38 |
| obs BASE-OS | 288 | 212 | 17 | 31 |
| obs COMP-OS | 48 | 25 | 3 | 7 |

Significance Levels of Mann-Whitney-U-tests on individual conditional contribution differences between BASE-OS and COMP-OS for every average contribution level of the other group members. *** represents significance at p=0.01, ** at p=0.05 and * at p=0.10. Classification of Conditional Cooperators according to the Spearman correlation.

**Result 3.4** *There is no change in preferences from competition within a classified contribution type group. Especially, competition does not amplify the self-serving bias of conditional cooperators.*

**Distribution of types** We find a decrease in average levels of conditional contribution across all subjects between the two treatments, but since we are not able to explain this difference with altered behavior within the identified groups of types, we next take a look at the frequencies of types in the population. Table 3.4 lists the absolute numbers and the frequencies for both treatments, BASE-OS and COMP-OS. In the BASE-OS treatment, there is a clear majority of about 74 % of the subjects (212 out of 288) classified as conditional cooperators with the Spearman method.[22] 10 % of the subjects (28 out of 288) do not contribute positive amounts and are classified as freeriders. The two remaining categories comprise 6 % (triangle contributors) and 11 % (not classifiable) respectively. The heterogeneity of types in the population is in line with the literature, the exact distribution of types, however, is somewhat different from the findings of Fischbacher et al. (2001) or Fischbacher and Gächter (2010) who find roughly 50 % of subjects to be conditional cooperators and up to 30 % to be freeriders. This is likely to be partly due to differences in the subject pool and partly to a different MPCR of 0.4 compared to 0.5 in our experiments.

Table 3.4: Distribution of Types

| | **Freerider** | **Conditional Cooperator** | **Humpshape** | **Other** | **Total** |
|---|---|---|---|---|---|
| COMP-OS | 13 | 25 | 3 | 7 | 48 |
| | 27.1% | 52.1% | 6.3% | 14.6% | 100% |
| BASE-OS | 28 | 212 | 17 | 31 | 288 |
| | 9.7% | 73.6% | 5.9% | 10.8% | 100% |
| Total | 41 | 237 | 20 | 38 | 336 |
| | 12.2% | 70.5% | 6.0% | 11.3% | 100% |

Number of observations and treatment frequencies of the observed types. Classification of Conditional Cooperators according to the Spearman correlation.

Introducing competition has a strong effect on the proportion of freeriders and conditional cooperators in the population. The share of freeriders increases from below 10 % to more than 27 % and the percentage of conditional cooperators is decreased from about 74 % to 52 %. The difference between the distribution of the four types in the population is highly significant (Fisher exact test and $\chi^2$ test, $p < 0.01$). Tests on different frequencies of each specific type in the two treatments confirm that the proportions of freeriders and conditional cooperators are significantly different ($\chi^2$ test, $p < 0.01$), but not for humpshape contributors and the residual category ($\chi^2$ test, $p > 0.46$).

The introduction of competition leads to a significant change of the distribution of contribution types according to Fischbacher et al. (2001) in the population towards more selfish subjects. We therefore find evidence for Hypothesis 3.3 that the loss of control over final payoff allocations through competition is at the heart of a weakening of other-regarding

[22] Using the stricter weak monotonicity rule as a classification methodology for conditional cooperators, 185 subjects (i.e., 64 %) fall into this category.

preference in the population. Our results show that a substantial part of the population is ”shifting“ from pro-social behavior in the form of conditional cooperation to the freeriding behavior. Rather than ”reducing“ their other-regarding concerns, subjects cease to behave pro-socially altogether what we summarize in our next result.[23]

**Result 3.5** *The introduction of competition leads to a significant shift in the distribution of the different contribution types away from conditional cooperators to more freeriders in the population.*

We can explain the decrease in average conditional contributions caused by the introduction of competition through a change in the distribution of underlying types in the population but not through changed preferences within an identified type of contribution behavior. Standard classification exercises of subjects into types may thus have only limited predictive power when the strategic environment becomes more competitive.

## 3.5 Conclusion

In this chapter, we investigate the impact of competition on cooperation behavior of subjects with heterogeneous other-regarding preferences. In the context of a social dilemma, we show that competition at the individual level alters the distribution of types in the population as measured in the conditional cooperation exercise and weakens incentives to build up a strategic reputation for selfish individuals in repeated interactions.

Our study is among the first to account for the differences in behavior between heterogeneous types in the population to an exogenous intervention. Our results shed an interesting light on the interpretation of experimental results on other-regarding preferences as they underline the importance of controlling for individual preferences when analyzing an aggregate treatment effect. The impact of an intervention through modified incentives may thus have a different impact on different types identified. For the implementation of e.g. a policy or a change in incentives within a firm, it is crucial to have information

[23]We do not have data on this, but the results suggest that in our setup the size of the bonus determines the extent to which conditional cooperators convert into freeriders, i.e., the share of the population that ”gives up“ social concerns to act as selfish.

about types ex-ante in order to induce the desired change in behavior ex-post. The implementation of behaviorally optimal contracts thus hinges on the availability of information on the precise nature of type heterogeneity and the knowledge about the effects of the policies on types.

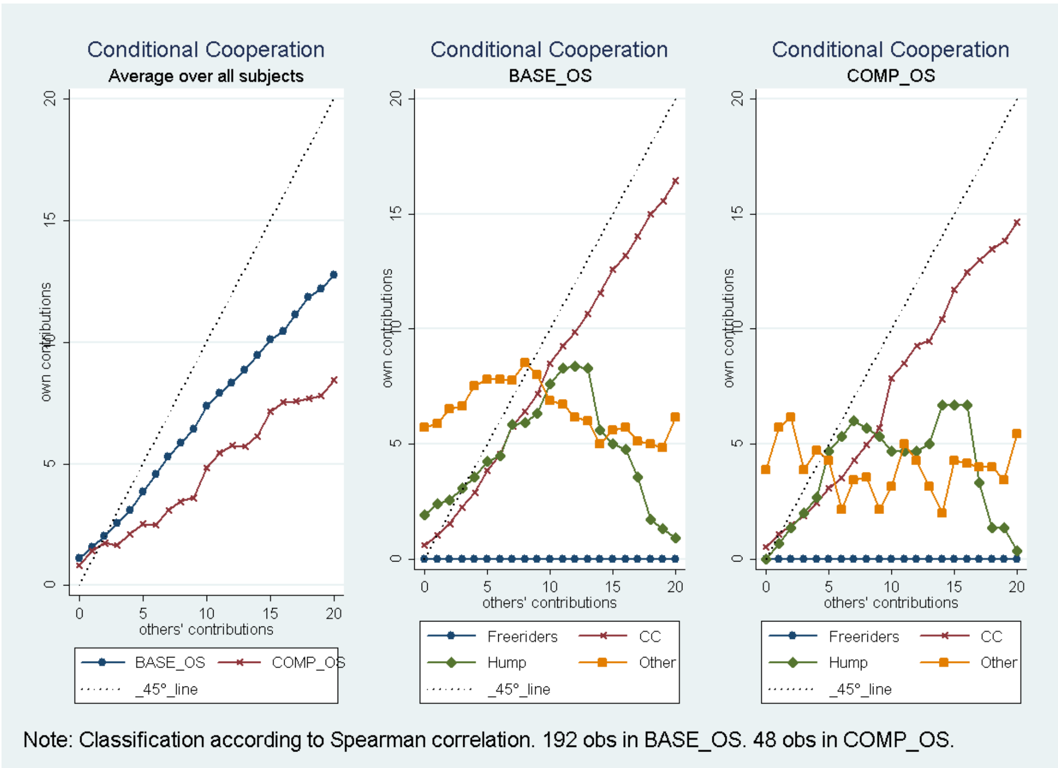
In this context it is important to note that our results may underestimate the effect of different incentive schemes on different people since there is no selection of individuals of similar preferences into groups in our laboratory experiment. If preferences are homogeneous within a group through selection into tasks or professions (Dohmen and Falk (2011)), the discrepancy between the change in behavior of two groups from an identical intervention may be considerable which constitutes a promising field of research that we leave for future investigation.

## 3.6 Appendix A3

### Appendix A3.1: Reduced Sample

In the text, we claim that the results do not change when we omit the additional 96 observations of subjects that are classified according to the methodology by Fischbacher et al. (2001). We report here all analyses from the text for the reduced sample to show that there is no effect from adding the 96 observations in the BASE treatment in the classification exercise. As one can see from the two left panels, there is no difference to those reported in the text. The contribution patterns within every type identified lie very

Figure A3.1: Conditional Contribution - Reduced Sample




Aggregate levels of conditional contribution for all subjects (left panel) and contribution patterns for determined types in treatments BASE-OS and COMP-OS (middle and right panel). Reduced dataset without additional data for BASE-OS treatment.

close to those from the complete dataset in the text. The same applies to the comparisons between the BASE and the COMP treatment for every type separately which is reported in the following table. Also in the reduced sample, we find the same significance levels for the Mann-Whitney-U-test that we reported in the text.

Table A3.1: Differences in Conditional Contribution - Reduced Sample

| Others' Contribution | All Subjects | Conditional Cooperators | Humpshaped | Other |
|---|---|---|---|---|
| 0 | - | - | - | - |
| 1 | - | - | - | - |
| 2 | - | - | - | - |
| 3 | ** | - | - | - |
| 4 | ** | - | - | - |
| 5 | *** | - | - | * |
| 6 | *** | * | - | *** |
| 7 | *** | ** | - | ** |
| 8 | *** | * | - | * |
| 9 | *** | * | - | *** |
| 10 | *** | - | - | * |
| 11 | *** | * | - | - |
| 12 | *** | - | - | * |
| 13 | *** | - | - | - |
| 14 | *** | - | - | * |
| 15 | *** | - | - | - |
| 16 | *** | - | - | - |
| 17 | *** | - | - | - |
| 18 | *** | - | - | - |
| 19 | *** | - | - | - |
| 20 | *** | - | - | - |
| obs Total | 240 | 167 | 16 | 24 |
| obs BASE-OS | 192 | 142 | 13 | 17 |
| obs COMP-OS | 48 | 25 | 3 | 7 |

Significance Levels of Mann-Whitney-U-tests on individual conditional contribution differences between BASE and CAMP for every average contribution level of the other group members. *** represents significance at p=0.01, ** at p=0.05 and * at p=0.10. Classification of Conditional Cooperators according to the Spearman correlation. Reduced dataset without additional data for BASE-OS treatment.

The distribution of types in the BASE treatment is also identical to that reported in the text from the full sample. All non-parametric tests that control for differences in the distributions between BASE and COMP produce identical levels of significance.

Table A3.2: Distribution of Types - Reduced Sample

| | Freerider | Conditional Cooperator | Humpshape | Other | Total |
|---|---|---|---|---|---|
| COMP-OS | 13 | 25 | 3 | 7 | 48 |
| | 27.1% | 52.1% | 6.3% | 14.6% | 100% |
| BASE-OS | 20 | 142 | 13 | 17 | 192 |
| | 10.4% | 74.0% | 6.8% | 8.9% | 100% |
| Total | 33 | 167 | 16 | 24 | 240 |
| | 13.8% | 69.6% | 6.7% | 10.0% | 100% |

Number of observations and treatment frequencies of the observed types. Classification of Conditional Cooperators according to Spearman correlation. Reduced dataset without additional data for BASE-OS.

## Appendix A3.2: Tables

We report below the results from Panel Tobit regressions on the level of contributions which yield similar results as the Random Effects regressions reported in the text.

Table A3.3: Panel Tobit Regressions on Contributions

| **Dep. Var.:** Contribution | **I** | **II** | **III** | **IV** |
|---|---|---|---|---|
| Period | -1.148*** | -1.148*** | -1.209*** | -1.209*** |
| | (0.058) | (0.058) | (0.062) | (0.062) |
| COMP | -0.715 | -4.393* | -0.593 | -9.342** |
| | (1.150) | (2.256) | (1.295) | (3.765) |
| Cond Cooperator | 2.632** | 0.100 | 2.960 | -1.453 |
| | (1.312) | (1.868) | (2.021) | (2.652) |
| COMP * Cond Cooperator | | 4.932* | | 9.884** |
| | | (2.613) | | (4.000) |
| Constant | 10.192*** | 12.106*** | 10.136*** | 13.990*** |
| | (1.312) | (1.649) | (2.005) | (2.497) |
| Obs | 1920 | 1920 | 1620 | 1620 |
| Wald $\chi^2$ | 400.6 | 403.6 | 381.3 | 386.0 |

Panel Tobit Regressions on contributions. COMP and Cond Cooperator are dummy variables. Specifications III and IV only for subjects classified as conditional cooperators and freeriders. Standard errors in brackets. *** represents significance at p=0.01, ** at p=0.05 and * at p=0.10.

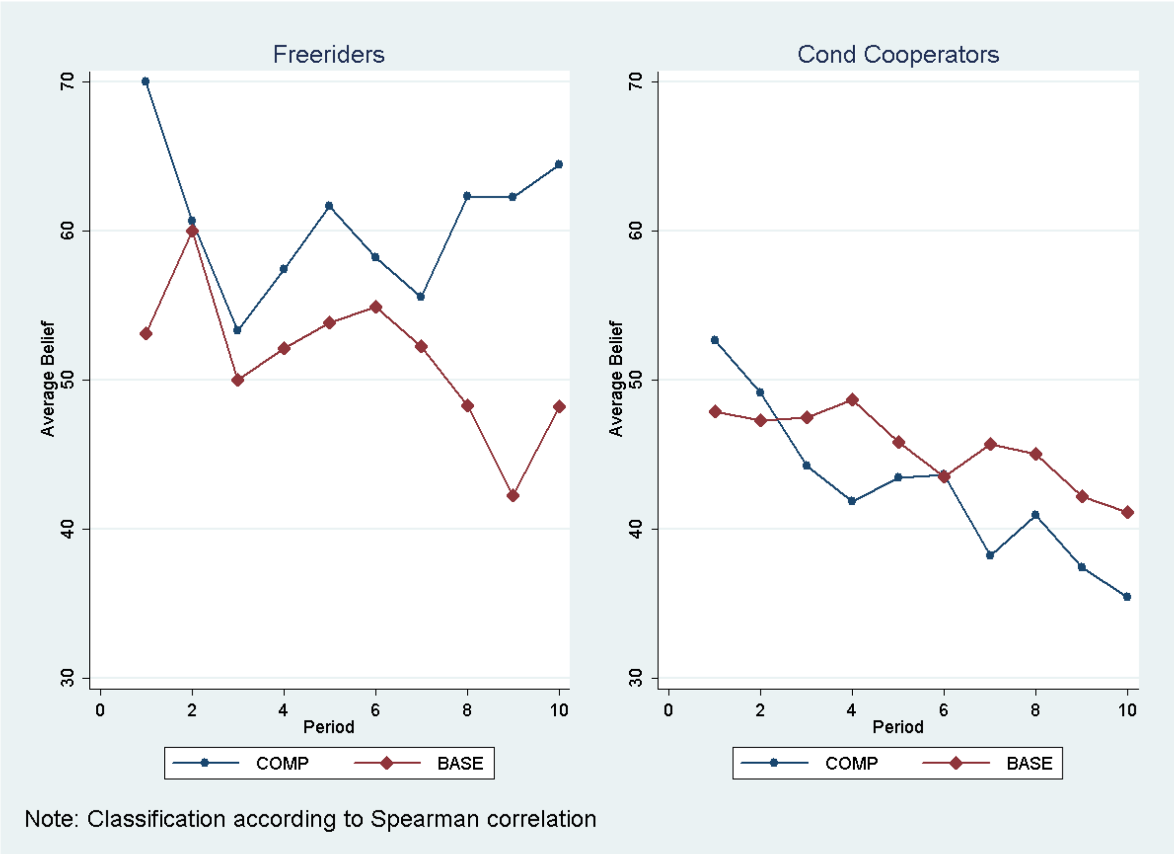
Table A3.4: Random Effects Panel Regression on Profits per Round

| **Dep. Var.:** Profits | **I** | **II** | **III** | **IV** |
|---|---|---|---|---|
| Period | -0.713*** | -0.713*** | -0.713*** | -0.689*** |
| | (0.064) | (0.064) | (0.064) | (0.069) |
| COMP | -0.453 | -0.472 | -1.696 | -1.913 |
| | (1.161) | (1.171) | (1.633) | (2.193) |
| Cond Cooperator | | -0.908 | -1.757* | -2.049 |
| | | (0.646) | (1.018) | (1.428) |
| COMP * Cond Cooperator | | | 1.654 | 1.871 |
| | | | (1.264) | (1.692) |
| constant | 31.100*** | 31.781*** | 32.418*** | 32.580*** |
| | (0.923) | (1.222) | (1.492) | (1.782) |
| Obs | 1920 | 1920 | 1920 | 1620 |
| $R^2$ | 0.10 | 0.10 | 0.10 | 0.10 |

Random Effects Panel Regression on profits per round. COMP and Cond Cooperator are dummy variables. Specification IV is restricted to subjects classified as conditional cooperators or freeriders. Standard errors in brackets. *** represents significance at p=0.01, ** at p=0.05 and * at p=0.10.

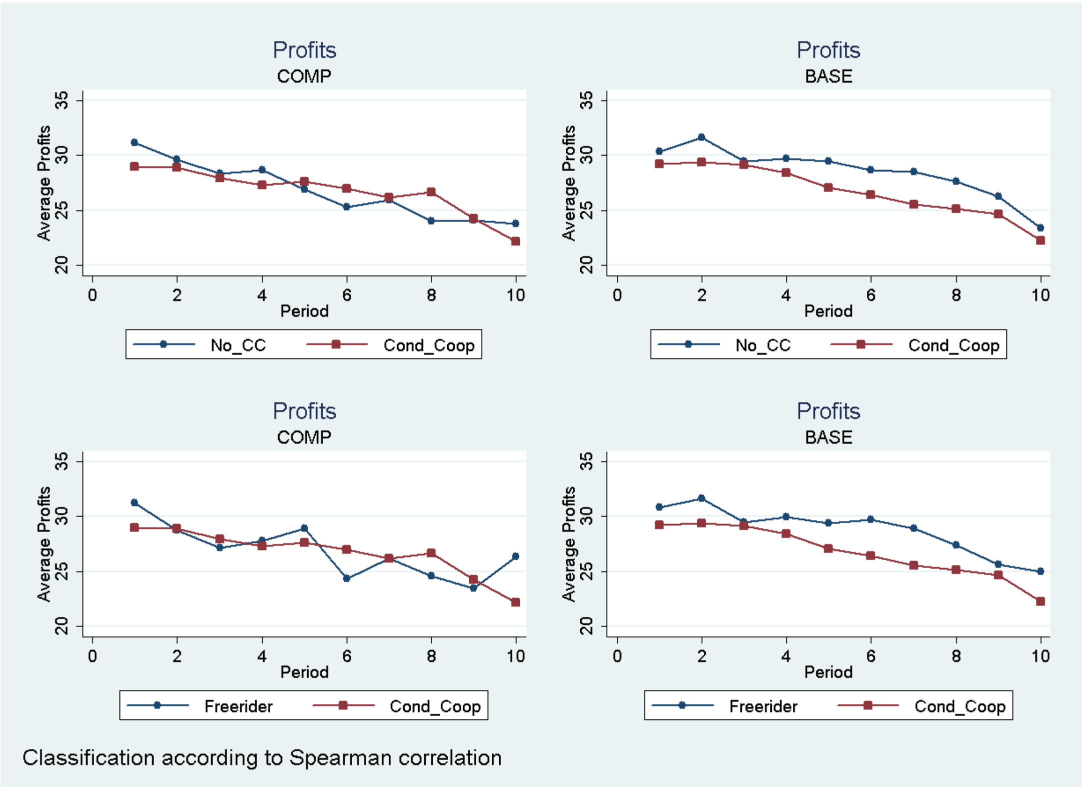
## Appendix A3.3: Figures

Figure A3.2: Average Beliefs



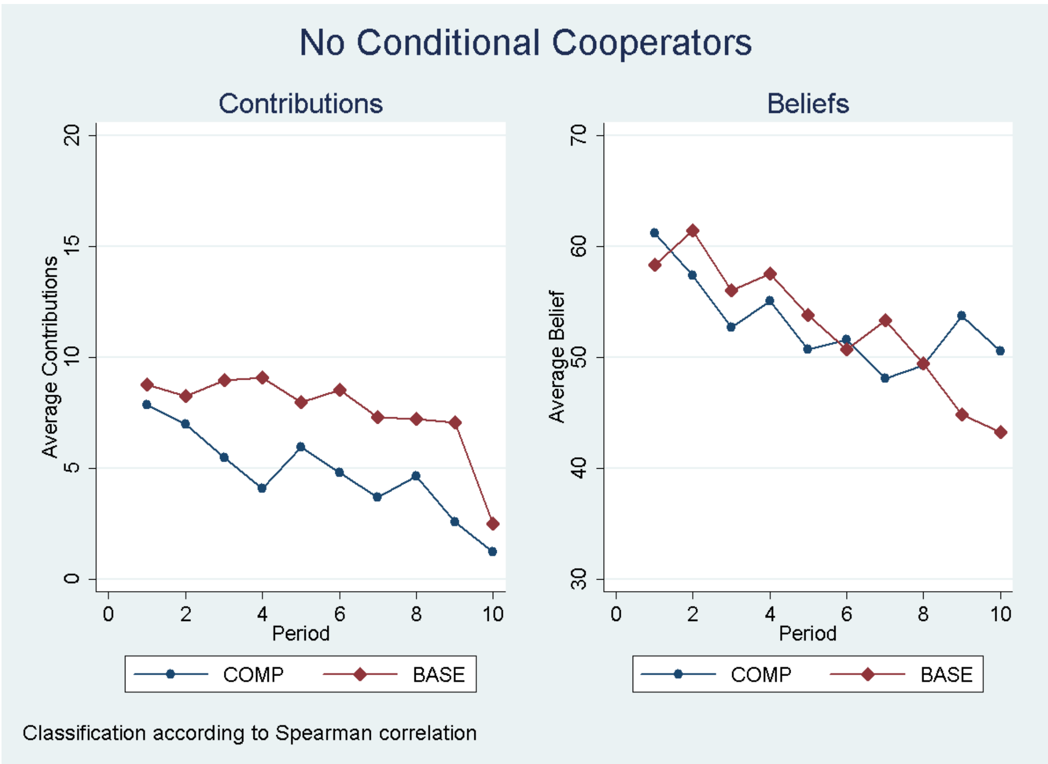

Beliefs about likelihood to have higher cumulative earnings up to current round for freeriders (left panel) and conditional cooperators (right panel) in treatments BASE and COMP

Figure A3.3: Profit Levels




Profit levels for freeriders, all subjects not classified as conditional cooperators (NoCC) and conditional cooperators in treatments COMP (left panels) and BASE (right panels)

Figure A3.4: Contributions Levels - No Conditional Cooperators




Contributions levels (left panel) and beliefs about likelihood to have higher cumulative earnings up to current round (right panel) for all subjects not classified as conditional cooperators in treatments COMP and BASE.

BASE / COMP

# Appendix A3.4: Instructions (translated from German)

The purpose of this experiment is to study decision making. In the course of the experiment, you as well as the other participants are going to make decisions. By making these decisions it is possible to earn some money. Your payoff is determined by your decisions as well as the decisions of the other participants according to the rules defined on the following pages.

**Anonymity**
During the experiment you will be interacting with other participants. Neither during nor after the experiment, will you be informed about the identity of the participants you interacted with. The other participants will neither during nor after the experiment be informed about your role or your earnings in the experiment. **During the whole experiment it is strictly forbidden to communicate.** Furthermore you are only allowed to use those functions of the computer that are designed for progressing with the experiment. Should you nonetheless try to communicate or tamper with the computer, you will be excluded from the experiment.

**Payoff**
For showing up on time you will receive **4 €, which you will keep in any case.** During the course of the experiment you can earn money according to the rules explained below. **After conclusion of the experiment** you will receive your guaranteed 4 € as well as the **amount of money**, that you earned during the course of the experiment. This will be paid **in private and in cash**. During the experiment we do not speak of Euros but of tokens. Every token you earned in the experiment will be converted to Euros according to the following rate:

**1 token = 0.10 Euro (10 tokens = 1 Euro)**

In your cubicle you will find a pen. Please transmit your decisions to the computer. While making your decisions there will be a countdown at the upper right of the computer screen. The purpose of this countdown is to give you an orientation about how much time you should need to come to a decision. However, the countdown will not be enforced in the case that you need more time to come to a decision. This will most probably be the case in the beginning. Only informational screens which do not require any decisions to be made will be removed after the countdown reaches zero.

**Time**
The duration of this three-part-experiment is approximately one hour and 15 minutes. You will receive the instructions for parts II and III after conclusion of part I. Should you have any questions or find anything obscure please raise your hand. One of the experimenters will come to your cubicle and answer your questions in private.

## Part I

At the beginning of this part participants will randomly be assigned to **groups of 4**. Neither during nor after the experiment will you learn about the identities of the other members of your group. Similarly the other participants will neither during nor after the experiment be informed with whom were they interacting in a group.

You have to make **two kinds of decisions** that are both made in the following setting:

First every participant receives an endowment of **20 tokens**. You can assign those 20 tokens to the two alternatives, **X** and **Y**:

1. You can assign 0 to 20 tokens to **box X**. The sum of all group members' contributions to box X will be multiplied by 2 of which one quarter will be allocated to every group member. Thus for every token in box X you will receive 0.5 (=2*1/4) tokens. If, for example, the sum of all tokens in box X is equal to 60, every group member receives 60*0.5 = 30 tokens out of box X. If the group members in sum allocate 10 tokens to box X, you and all other group members receive 10*0.5 = 5 tokens out of box X.

page 1 of 6

2. Of the 20 tokens, the amount you do not allocate to box X is automatically allocated to **box Y**. These tokens directly contribute to your payoff in this period. For example if you allocate 6 tokens to box Y, exactly 6 tokens out of box Y are added to the payoff of the period.

Your total payoff is the sum of your payoff out of box X and your payoff out of box Y.

Your payoff is thus calculated by:

$$\textbf{Your Payoff =} \quad \underbrace{(20-x)}_{\textit{Payoff out of box Y}} + \underbrace{(0{,}5 \cdot S)}_{\textit{Payoff out of box X}}$$

x = Your contribution to box X
S = Sum of the contributions of *<u>all</u>* group members to box X

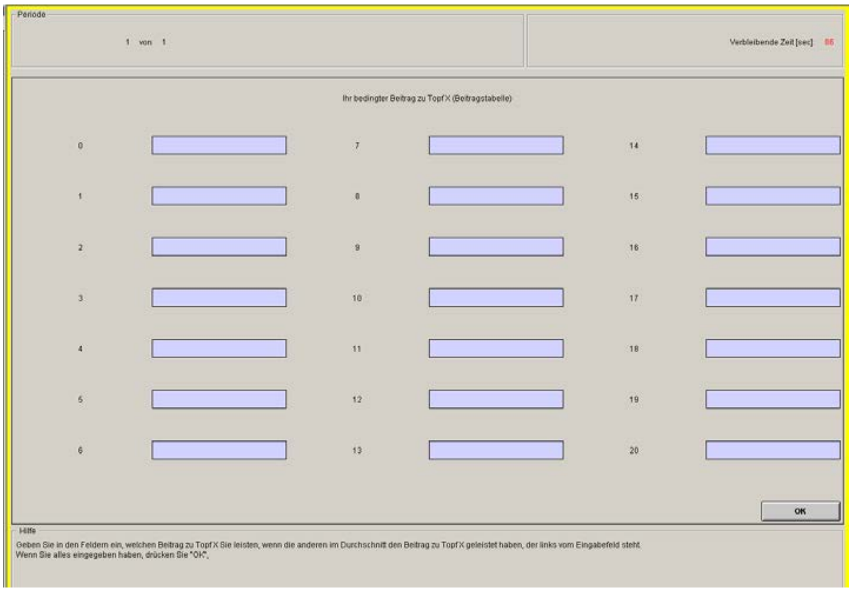
To clarify here is another example: You assign 12 tokens to box X and thus 8 tokens to box Y. The other group members' contributions are 4 tokens by participant A, 9 tokens by participant B and 15 tokens by participant C. The overall amount in box X thus is 12 + 4 + 9 + 15 = 40 tokens.

Your payoff then amounts to (20 – 12) + (0.5 * 40) = 8 + 20 = 28 tokens.
Participant A's payoff amounts to (20 – 4) + (0.5 * 40) = 16 + 20 = 36 tokens.
Participant B's payoff amounts to (20 – 9) + (0.5 * 40) = 11 + 20 = 31 tokens.
Participant C's payoff amounts to (20 – 15) + (0.5 * 40) = 5 + 20 = 25 tokens.

As mentioned above every group member makes **two kinds** of contribution decisions. We distinguish between the **unconditional contribution** and the **contribution table.**

- On the first screen you are asked to select your **unconditional** contribution to box X, i.e. you select how many of the 20 tokens you want to assign to box X. The remainder of the 20 tokens is automatically allocated to box Y. You can select any integer between 0 and 20 (including 0 and 20). Just type your unconditional contribution in the input box on your screen and confirm this amount by clicking „OK". As long as the OK button has not been clicked you can change the amount of your contribution.

- On the second screen you are asked to fill out a **contribution table**. In this contribution table you select for **every possible** (rounded) **average contribution of the other group members to box X, how many tokens <u>you</u> want to contribute to box X**. This means you select your contribution dependent on the average contribution of the others. The contribution table is illustrated in the picture below.



The numbers in the left columns are all possible (rounded) average contributions of the other group members to box X, i.e. the amount that every other group member on average contributes to box X. You type into every input box, how many tokens **you – given that the other group members on average contribute the respective amount –** want to contribute to box X. You determine how much you are going to contribute if the other group members contribute on average 0 tokens to box X; how much you contribute if the other group members contribute on average 1, or 2, or 3 tokens. Into every input box you can type any integer between 0 and 20 (i.e., you can also type in the same amount into several or all input boxes). **You have to type a number into every box.** Once you have filled out every box on the screen please click the "OK"-button. As long as the "OK"-button has not been clicked you can still revise your decision.

After all participants have decided about their unconditional contribution and filled out their contribution table the computer chooses one group member out of every group by a random draw. For the **randomly chosen group member** only the filled out contribution table is relevant for calculating your final payoff. For the **remaining three group members** only the **unconditional contribution** is relevant for calculating the payoff. While you decide on your unconditional contribution and fill out your contribution table you do not know whether you will be randomly chosen or not. Thus **you should contemplate both contribution decisions** as both can become relevant to your payoff.

*<u>Example 1:</u> Assume that **you have been randomly chosen**, **i.e. your contribution table is payoff relevant**. For the other three group members thus the unconditional contribution decision is relevant. Assume these to be 4, 9 and 16 tokens. The rounded average contribution of the other group members thus is 10 tokens ([4+9+16]/3=9.66).*
*If you specified in your contribution table that you contribute 13 tokens to box X if the others contribute on average 10 tokens, the total contribution to box X amounts to 4+9+13+16 tokens = 42 tokens.*
*Your payoff in part I then amounts to (20-13)+0.5*42 = 7+21 = 28 tokens.*

*<u>Example 2:</u> Assume that **you have not been randomly chosen, i.e. for you and two other group members the unconditional contribution decision is payoff relevant**. Assume your unconditional contribution to box X to be 8 tokens, the others' contributions to be 10 and 16 tokens. The average unconditional contribution of the three of you then amounts to 11 tokens ([8+10+16]/3=11,33).*
*If the randomly chosen group member specified that she contributes 15 tokens to box X if the average contribution of the other three group members amounts to 11 tokens, the total contribution to box X is 8+10+15+16=49 tokens.*
*Your payoff in part I then amounts to (20-8)+0.5*49=12+24.5=36.5 tokens.*

The result of part I will be disclosed after conclusion of the whole experiment. You will then be informed which decision (the unconditional contribution or the contribution table) is relevant to your payoff, the contributions of the other group members and your resulting payoff of part I. You will be paid at the end of the experiment.

## Part II

*(parts and instructions were presented sequentially to subjects)*

### Time structure

This part consists of **10 identical periods**. In every period you are again a member of a **group of 4**. This is exactly the same group as in part I. **The group stays the same over the whole course of part II, i.e. you are interacting with the same persons over all periods.**
**Additionally you will be randomly assigned to another participant in this room, who is <u>not</u> a member of your 4-person-group whom you have to assess in part II. In part III a prize of 70 tokens will be given either to you or this assigned participant from another group. Whether you or your partner receives the prize depends on the decisions in part II.**

In every period you will make two decisions, a **contribution decision** and an **assessment decision**. The contribution decision is equivalent to the **unconditional contribution decision in part I**. The assessment decision asks you to **assess whether the participant from a different group who has been randomly assigned to you has earned more or less than you up to (including) this period**.

BASE / COMP

Both decisions determine your payoff in part II: After conclusion of the experiment one of the 10 periods is randomly drawn; this period is relevant to your payoff. Another random draw determines whether your contribution decision or your assessment decision is relevant to your payoff.

As you do not know which period and which decision will be relevant to your payoff, it is in your own interest to consider every decision carefully as every decision in every period could be relevant to your payoff.

## 1. Contribution decisions in every period

The **setting is identical to period I**. You are endowed with 20 tokens in every period and decide how many of these 20 tokens you are contributing to box X and how many you are contributing to box Y. The payoff is calculated identical to part I.
The only difference is that you only decide on your **unconditional contribution** and do **not fill out a contribution table.**

The periodical earnings of all group members are **calculated the same way, depending on the unconditional contributions of all four group members. On the screen you will be asked how many tokens you want to contribute to box X.** The remainder of the 20 tokens automatically is contributed to box Y. Saving tokens for the next periods is not possible. You can only specify integers between 0 and 20 (including 0 and 20). Please type your contribution to box X into the input box and confirm your decision by clicking „OK“. As long as you have not clicked the “OK”-button you can still revise your decision.

**Then you are informed about the other group members’ contributions, the total contribution of all group members to box X and your payoff in this period.** The order in which the contributions of the other group members appear is random. Thus you cannot derive a specific group member’s contributions over several periods.

**After you have made contribution decisions in every of the 10 periods, the computer – after conclusion of the experiment – randomly chooses one payoff-relevant period and whether your contribution decision or your assessment decision in the chosen period is payoff-relevant.**

## 2. Assessment decisions in every period

Concluding each period you have to **assess the participant from a different group that has been randomly assigned to you**. You can earn up to **40 tokens**. The better your assessment, the higher your payoff.
In every period you specify an **integer probability between 0 and 100 which you believe is the probability that you have higher total earnings than the participant that has been randomly assigned to you. The total sum of earnings that have been accumulated by you or the participant assigned to you in part II, including this period,** are relevant. Note that the randomly assigned participant makes his decision subject to exactly the same rules as you. Which period is later randomly chosen to be payoff relevant is not important; every period up to and including the current period are relevant to the assessment decision. The earnings in part I are irrelevant.

**If the assessment decision is randomly determined to be payoff-relevant,** your payoff in period II depends on whether you or the participant randomly assigned to you accumulated a higher total amount of earning up to and including the randomly chosen period.

**Case 1 (Your accumulated earnings are higher):**

Your Payoff = ((Your Probability in %) / 100 ) * 40 Tokens

**Case 2 (The accumulated earnings of your randomly assigned participant are higher):**

Your Payoff = ((100 - Your Probability in %) / 100) * 40 Tokens

page 4 of 6

BASE / COMP

If the accumulated earnings of both of you are equivalent the computer randomly chooses with probability 50% whose earnings are to be considered as higher.

**The stronger your belief that your earnings are higher, the higher is the probability you should specify. The more you think that your earnings are lower, the lower the probability you should specify.**

*<u>Example 1:</u> Your relevant assessment for part II specified that with 70% probability your earnings are higher than those of the participant randomly assigned to you. There are three possible cases: If your total earnings are indeed higher, you receive 70/100 * 40 = 28 tokens. If the total earnings of the participant randomly assigned to you are higher, you receive (100 – 70)/100 * 40 = 12 tokens. If your earnings are equivalent the computer randomly determines with a probability of 50% if you receive 28 tokens (as if your earnings were higher) or 12 tokens (as if your earnings were lower).*

*<u>Example 2:</u> Your relevant assessment for part II specified that with 20% probability your earnings are higher than those of the participant randomly assigned to you. There are again three possible cases: If your total earnings are indeed higher, you receive 20/100 * 40 = 8 tokens. If the total earnings of the participant randomly assigned to you are higher, you receive (100 – 20)/100 * 40 =32 tokens. If your earnings are equivalent the computer randomly determines with a probability of 50% if you receive 8 tokens (as if your earnings were higher) or 32 tokens (as if your earnings were lower).*

Please type your assessment in % between 0 and 100 into the input box and confirm your assessment by clicking the „OK“-button. As long as the “OK”-button has not been clicked you can still revise your decision.

Afterwards the next period commences. **After 10 identical periods this part of the experiment ends.**

**Your payoff is determined after completion of the experiment. A period is randomly determined; another random draw determines whether the contribution decision or the assessment decision is relevant to your payoff.** While you either decide on your unconditional contribution or assess the participant randomly assigned to you, you neither know whether this period nor whether the contribution decision or the assessment decision is relevant to your payoff. **Thus you should carefully consider both decisions in every period** as both can become relevant to your payoff.

## Part III

In this part of the experiment you can – as mentioned above – in addition to your payoff in part I and part II receive a prize. **The prize amounts to 70 tokens**. You are guaranteed to keep your payoff from both parts; you cannot suffer any losses in part III.

[COMP: **Whether you win the prize of 70 tokens depends on the <u>total sum of earnings you and the participant randomly assigned to you accumulated through the contribution decisions in part II over the whole 10 periods.</u> Your total earnings from the contribution decisions in all periods as well as the total earnings from the contribution decisions in all periods of the participant randomly assigned to you determine who of you two receives the prize.** Neither the assessment decision nor your earnings from part I are relevant.

**To determine who receives the prize of 70 tokens your total sum of earnings you accumulated through the contribution decisions in part II over the whole 10 periods is compared to the total sum of earnings the participant randomly assigned to you accumulated through the contribution decisions in part II over the whole 10 periods. The person who accumulated a higher total sum of earnings receives the prize of 70 tokens. If you both accumulated the exact same total sum of earnings the computer randomly draws a winner with probability of 70 tokens.** Some examples for clarification:

page 5 of 6

BASE / COMP

*<u>Example 1:</u> Your earnings from **all contribution decisions in all periods of part II** amount to 310 tokens and the earnings from **all contribution decisions in all periods of part II** of the participant randomly assigned to you amount to 240 tokens. Thus you receive the prize. The participant randomly assigned to you does not receive any payoff from part III.*

*<u>Example 2</u>: Your earnings from **all contribution decisions in all periods of part II** amount to 240 tokens and the earnings from **all contribution decisions in all periods of part II** of the participant randomly assigned to you amount to 280 tokens. Thus the participant randomly assigned to you receives the prize. You do not receive any payoff from part III.*

*<u>Example 3</u>: Your earnings from **all contribution decisions in all periods of part II** amount to 270 tokens and the earnings from **all contribution decisions in all periods of part II** of the participant randomly assigned to you amount to 270 tokens. Thus the computer randomly determines with 50% probability who receives the prize.*

**Your earnings from all contribution decisions in all periods of part II as well as the earnings from all contribution decisions in all periods of part II of the participant randomly assigned to you are shown to you on the screen. You are then informed whether you or the participant randomly assigned to you receive the prize.** The payoff from part I is irrelevant to this, as is the actual payoff from part II (contribution decision or assessment decision).***]***

[BASE: **Whether you win the prize of 70 tokens depends on the throw of a dice. Your chances are 50 %. The computer allocates "Lucky numbers" between 1 and 6 to both you and the participant randomly assigned to you.** The computer randomly decides about whether you obtain the low (1-3) and the low (4-6) figures.

To determine who wins the prize in part III, a dice is publicly rolled to determine a "Lucky Number" between 1 and 6. If this "Lucky Number" is part of your lucky numbers, you win the prize of 70 tokens. If it is not, the participant randomly allocated to you obtains the prize.

<u>Please note</u>: The above described procedure ensures that every figure between 1 and 6 is chosen with equal probability, i.e. your chance to win the prize is exactly 50 %. Who is allocated the high or the low "Lucky Numbers" does not matter.]

Your payoff from part III is shown to you at the end of part III on the screen. You do not have to make any decision in part II.

At the end of the experiment your payoff from part I and part II is determined and are shown to you in conjunction with your total payoff from the whole experiment on the screen.

# Chapter 4

# Responsibility Effects in Decision Making under Risk[0]

## 4.1 Motivation

Economic situations in which an agent takes decisions that affect others' outcomes as well as her own constitute a common class of phenomena. For instance, they represent situations in which a decision maker's choices affect not only her own outcomes, but those of her family as well. Another common instance of such decision problems is the one of financial agency contracts in which the incentive structure of the agent coincides with the one of the principal. An example may be the one of executives that are compensated through company shares, or the one of a stock broker whose payoffs are determined by the outcomes of the investments she undertakes.

There is an extensive literature on individual decision making under risk and uncertainty (Abdellaoui et al., 2011; Post et al., 2008), as well as a substantial literature on risk attitude in agency problems and how to influence it through performance-contingent pay (Wiseman and Gomez-Mejia, 1998). What is missing, however, is a direct comparison of risk attitudes when decisions are individual to situations of responsibility. Indeed, to the extent that decisions under responsibility differ from decisions commonly found in the individual decision making literature, findings from the latter will only constitute an

[0] This chapter is based on joint work with Julius Pahlke and Ferdinand Vieider.

imperfect predictor of attitudes under responsibility. Given that the latter constitute an economically important class of decision situations-and indeed one that in its economic importance may even surpass individual decisions-additional evidence on any differences can provide important insights for descriptive as well as prescriptive and policy purposes. We thus explore the difference in risk attitudes between situations of decision-making for oneself and in situations of responsibility, i.e., situations in which the decision maker decides for others as well as herself. We explore such decisions for situations in which an anonymous other (the recipient) is affected by any outcomes in exactly the same way as the decision maker herself. This allows us to study possible changes in behavior in a clean way, excluding issues deriving e.g. from preferences over outcome distributions that may cause inequality concerns. Also, by making both the decision maker's outcome and the recipient's outcome dependent on the decision makers' choice, the latter will bear an actual cost in terms of her own preferences by accommodating any presumed preferences of the recipient or by following some social norm. Any findings should thus constitute a lower bound on the effects we want to investigate.

To our best knowledge, the only paper that reports results about this issue under equal payoff assumptions in a non-strategic setting is Bolton and Ockenfels (2010), although the authors report these results as an afterthought to their main results about inequality concerns and do not find statistically significant results due to their small sample size. They also discuss only the case of decisions in the gain domain. We explore the issue systematically for risky choices in the gain domain, the loss domain, and the mixed domain. Individual risk attitudes have been found to differ systematically in the different domains (Abdellaoui, 2000; Booij et al., 2010; Schoemaker, 1990). To the extent that individual risk attitudes have been found to differ systematically across the probability and outcome spaces, responsibility may well have different effects across those dimensions. While we adopt a theory-neutral approach in our exploratory efforts, the inclusion of different decision domains will allow us to capture any richness in behavior as predicted by descriptively more complex theories such as prospect theory.

We find that in the gain domain, being responsible for others as well as oneself does indeed increase risk aversion for medium to large probabilities, thus showing that the intuition of Bolton and Ockenfels (2010) was correct. In addition, we show that for pure loss prospects,

subjects become more risk seeking when responsible for others. Loss aversion on the other hand, being already very strong in individual decisions, does not seem to increase when subjects are responsible for others. In a second experiment aimed at exploring social norms on risk taking in the gain domain in more detail, we replicate the finding that risk aversion increases under responsibility for large probabilities. When choices regard small probability prospects, however, we find increased risk seeking under conditions of responsibility. Overall thus our results point to an accentuation of the fourfold pattern of risk attitudes typically found in individual decision making when subjects are responsible.
The chapter proceeds as follows. Section 4.2 discusses risk attitudes and how they may be influenced by social contexts. Section 4.3 describes the first experiment, with section 4.3.1 describing the methodology and section 4.3.2 presenting the results; section 4.3.3 discusses the result of experiment 1 and derives hypotheses for experiment 2. Section 4.4 introduces experiment 2, with section 4.4.1 describing the methodology and section 4.4.2 presenting the results. Section 4.4.3 discusses the results of experiment 2 as well as the overall results. Section 4.5 concludes this chapter.

## 4.2 Risk Attitudes in Social Contexts

In recent years, there has been a growing interest by economists in how social factors may influence decision making under risk (Bohnet et al., 2008; Bolton and Ockenfels, 2010; Goeree and Yariv, 2007). Such "social factors" could take various forms, ranging from whether a decision is observed by somebody else or whether the decision maker observes somebody else's decision, to whether one's outcome depends on somebody else or whether one's decision influences the outcome of somebody else (Trautmann and Vieider, 2010). We are interested in the latter category: do preferences over risky choices change when the decision influences somebody else's outcomes as well as the ones of the decision maker? And if so, how?
To date there is very little evidence on this issue, with the existing evidence appearing inconclusive. Bolton and Ockenfels (2010) hypothesize that risk aversion will increase under responsibility. However, their result fails to reach statistical significance. Indeed, their main results concern the effect of social comparison, so that they mainly examine

choice behavior when outcomes may differ between the decision maker and the recipient. They find an increase in risk taking under conditions of responsibility when the safe option yields unequal payoffs, and particularly when such payoff asymmetry is unfavorable to the decision maker. In contrast, they find that under responsibility risk taking does not depend on whether the risky option yields unequal payoffs.

In a somewhat related study from the game-theoretic literature, Charness and Jackson (2009) have subjects play Rousseau's stag hunting game against each other. They compare conditions in which one subject simply plays against another, to one in which a second, passive, subject depends on each player. They find that under responsibility for someone else the efficient equilibrium obtains less frequently. While this may again be an indication for increased risk aversion under responsibility, it is not clear where such a risk may actually come from since it is not in the interest of any of the players to deviate from the efficient equilibrium unless they think the other player may deviate. Furthermore, the setup of the study again creates issues of inequality aversion. Even if the recipient obtains the same payoffs as the decision maker, the strategic nature of the game implies that the decision maker can influence the payoffs of her opponent and the latter's passive recipient, which may affect her choices ex-ante.

We aim to specifically exclude inequality concerns to filter out the pure effect of being responsible for somebody else's payoffs. In order to achieve this, the exact choice that determines the decision maker's payoff also determines the recipient's payoff, resulting in exactly the same outcome for the decision maker and the recipient. This design thus allows us to isolate the effect of being responsible for somebody else as well as for oneself from any distributional issues (Rohde and Rohde, 2010). Furthermore, there are costs for the decision maker in adapting her preferences under conditions of responsibility in terms of sacrificing her own preferences. In this sense, we believe that our design constitutes a lower bound on any effects of responsibility that could be found employing alternative designs, such as e.g. salaried agents.

Given the lack of conclusive evidence to date, we propose to systematically explore the effect of responsibility on risk preferences throughout the outcome and probability domains. In order to facilitate that task, in what follows we will adopt a behavioral, and hence theory-neutral, definition of risk aversion. A decision maker will be defined as risk averse whenever she prefers the expected value of a prospect to the prospect itself; conversely, she will be defined as risk seeking whenever she prefers the prospect to a sure amount equivalent to the prospect in terms of expected value (Wakker, 2010, p.52). Risk aversion and risk seeking are thus relative terms, such that a decrease in risk aversion can be seen as equivalent to an increase in risk seeking, regardless of absolute levels of risk taking.[1] In our presentation of the results we recur to prospect theory-the prevalent descriptive theory of choice under risk and uncertainty today (Starmer, 2000; Wakker, 2010). Under prospect theory, risk attitudes are described by utility curvature, loss aversion, and probability weighting (Köbberling and Wakker, 2005). Since prospect theory is more general than other theories of decisions under risk such as expected utility theory, we can thus capture richer risk attitudes if present, without however imposing a theory on our data a priori. In individual decision making under risk, the typical finding is a fourfold pattern of risk attitudes: risk aversion for medium to large probabilities of gains; risk seeking for small probability gains; risk aversion for small probability losses; and risk seeking for medium to large probability losses (Abdellaoui, 2000; Abdellaoui et al., 2010; Bleichrodt and Pinto, 2000; Tversky and Kahneman, 1992). In addition to this fourfold pattern, for mixed prospects involving both gains and losses, risk attitudes are significantly influenced by loss aversion-the phenomenon according to which monetary losses are usually attributed greater weights than equivalent monetary gains (Abdellaoui et al., 2007; Schmidt and Zank, 2005; Tversky and Kahneman, 1992).

The question of whether and how being responsible for others changes choice behavior also raises interesting questions about rationality concepts, social norms on risk taking and the perceived acceptability of attitudes towards risks. This has implications for 'debiasing', or simply changing risk attitudes in ways that may seem socially desirable. By comparing situations of individual decisions to situations of responsibility for different probabilities and in different domains, we are able to examine the perceived acceptability of common individual decision making patterns under risk. To the extent that responsibility for others acts as a cognitive motivator for a more careful consideration of the decision, we can draw conclusions about the perceived acceptability of a type of behavior by observing if

[1] This means that saying that choices under condition A are more risk averse than under condition B is taken as equivalent to saying that they are less risk seeking under A than under B, regardless of the absolute level of risky or safe choices (i.e, regardless of whether safe choices are more or less than 50% in both cases, or whether they cross the 50% mark).

and how people move from the individual baseline when responsible for others.

## 4.3 Experiment 1: Responsibility for Gains, Losses and Mixed Prospects

### 4.3.1 Experimental Design

We designed a laboratory experiment in which we asked subjects to take binary decisions between two alternatives, that are presented to them on a computer screen. Payoffs always affect the decision maker and the recipient in a perfectly parallel manner in the responsibility treatment, so as to avoid issues of payoff inequality (Bolton and Ockenfels, 2010; Rohde and Rohde, 2010).

**Subjects.** Overall, 144 subjects were recruited from a subject pool of the experimental laboratory MELESSA at Ludwig-Maximilian's University in Munich, Germany, via ORSEE (Greiner, 2004). The experiment took roughly 1.5 hours, and average earnings were 22.5 €. The experiments were run on computers using zTree (Fischbacher, 2007). 46% of subjects were female, and the average age was 24.07 years.

**Task.** Subjects were asked to choose between a safe prospect and a risky prospect. The safe prospect usually consisted in a sure amount of money, and sometimes in a prospect with lower volatility compared to the risky prospect. The risky prospect always gave a 50-50 chance to obtain one of two outcomes. The prospects could comprise only positive amounts, only negative amounts, or both positive and negative amounts (see below). Overall, subjects had to make 40 choices, with the order of presentation as well as the position of the two prospects randomized for each subject. Subjects took decisions sequentially and had no opportunity to return to an earlier decision to revise it. All of the above was explained in the instructions.

**Prospects.** The 40 choices to be made by all subjects in the experiment were constructed systematically in the following way: We chose five different stake levels that we denote henceforth by **b** where $b \in \{2, 4, 6, 8, 10\}$. For every stake level, we had subjects chose between the following eight different prospect pairs:

- *Base Case*: These prospect pairs offered a choice between the safe payment $b$ and a prospect providing a 50% chance to win twice the safe amount $b$ or zero otherwise.
- *Sensitivity up*: Compared to the basic choice pair, the safe payment is increased by 25% to assess subjects' the degree of risk aversion. The risky option is unchanged.
- *Sensitivity down*: Similar to "Sensitivity up", but the safe payment is reduced by 25%, again in order to measure risk aversion. The risky option is unchanged.
- *Positive shift*: Every amount is increased by 50% of the safe payment in the base category. These choices were included to see how choices changed when shifting away from the 0 € outcome.
- *Lottery choice*: The risky prospect now remains identical to the basic case, but the safe payment is replaced by a prospect with a lower variance ($0.5b$ and $1.5b$) than the risky prospect (0 and $2b$).
- *Mixed prospects*: To obtain these prospects, the safe amount in the base case was subtracted from all outcomes, thus obtaining a prospect with an expected value of 0 €. The safe amount was always 0, the prospect a lottery between $-b$ and $b$.
- *Mean-preserving spread*: The two risky outcomes of the base case were respectively increased and decreased by 50% of the sure amount. The expected value of the prospect thus remains the same; however, the variance of the prospect increases, and a loss equal to 50% of the sure amount is introduced into the prospect.
- *Loss Shift*: The mirror image of the base case where every amount was negative instead of positive. These prospects were inserted to directly compare risk taking behavior for gains and losses.[2]

[2]Additional prospects in the gain domain were not mirrored for ethical reasons. Indeed, replicating all gain prospects for losses would have resulted in a high chance of overall losses during the experiment.

The following table gives an overview of the eight different prospect pairs as a function of the stake level $b$.

Table 4.1: Overview of Lotteries

| Choice Type | Option A ("Safe") | | Option B ("Risky") | |
|---|---|---|---|---|
| Base | b | | 0 | 2 b |
| Sensitivity Up | 1.25 b | | 0 | 2 b |
| Sensitivity Down | 0.75 b | | 0 | 2 b |
| Positive Shift | 1.5 b | | 0.5 b | 2.5 b |
| Lottery Choice | 0.5 b | 1.5 b | 0 | 2b |
| Mixed Prospect | 0 | | -b | b |
| MPS | b | | -0.5 b | 2.5 b |
| Loss Shift | -b | | -2 b | 0 |

For a complete overview of all prospect pairs, see Table A4.2 in appendix A4.2.

**Treatments.** Subjects were randomly assigned to one of two treatments. In the individual treatment, subjects took their decisions only for themselves. In the responsibility treatment, half of the subjects were randomly assigned the role of decision maker and the other half to the role of passive recipient. The decision maker was told that she had to take the decision on behalf of herself and another subject sitting in the laboratory, whose identity was not disclosed. All other subjects were told that they were in a passive role and that somebody else in the laboratory would take the decisions on their behalf. With a lag of one period, recipients were shown the decision problem and the choice of their corresponding decision maker. They could then indicate whether they were "satisfied" or "not satisfied" with the decision, but this did not affect payoffs nor was it shown to the decision maker.

**Incentives.** 3 out of the 40 decisions were randomly drawn for every subject to be payoff relevant once the experiment was over. Subjects did not learn about any payoffs or extractions before the very end of the experiment. The random incentive system was chosen in order to avoid possible income effects, and because it is the standard procedure used in this kind of tasks. We extracted 3 out of the 40 choices in order to reduce the probability that subjects would actually lose money in the experiment. To make the random mechanism behind lotteries as transparent as possible, we had one participant throw a dice for every lottery that determined what outcome of the lottery is obtained. In the responsibility treatment, we implemented the payout procedure such that always three identical decisions were randomly chosen for the two paired subjects - a decision maker and her passive recipient would thus always obtain the same payoff from a choice. Subjects were told that it was possible - though unlikely - that they would lose money in the experiment. They could either pay such losses directly or work them off in the lab for a wage of 5 € per half hour.

### 4.3.2 Results

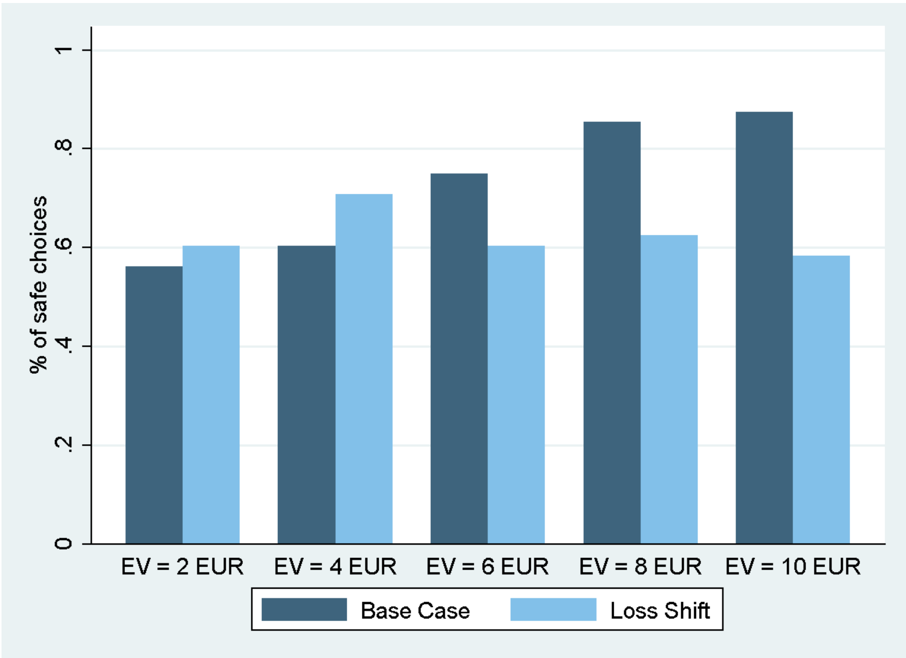
**Prospect Choices: Overview** Before discussing treatment effects, it seems desirable to discuss general risk attitudes and how they change for the different types of prospects employed when we look at the individual treatment only. In the base case we find a considerable degree of risk aversion across all stake levels, with about 73% of subjects choosing the sure amount over the prospect with equal expected value ($p < 0.001$)[3]. As one would expect, choices of the sure amount further increase when the sure amount is higher than the expected value of the prospect (Sensitivity Up), and decrease when the sure amount is lower (Sensitivity Down) in which case we observe a majority of choices for the prospect ($p < 0.01$). When compared to the base case all outcomes are moved upward by 50% of the sure amount (Positive Shift), we observe increased choices of the prospect, although choices still display significant risk aversion ($p < 0.01$). This can be explained by aspiration level theory, whereby subjects aspire to win at least some money, thus making a prospect with a non-zero minimal outcome more attractive (Payne et al., 1980, 1981).

When the choice is between two non-degenerate prospects (Lottery Choice), choice frequencies of the safe prospect are further increased relative to the base case, indicating a similar heuristic, since the safe choice now provides a combination between a safe minimum amount and a potentially higher outcome. For mixed prospects, the choice

[3]p-values reported are two-sided and refer to binomial tests for intermediate stakes, with a safe amount of $b = 6$, unless specified otherwise.

frequency of safe choices is only slightly increased compared to the base case (this however underestimates the effect given the lowering of the stake levels: see below as well as appendix A4.1 for a more nuanced discussion). For the mean-preserving spread, choices of the risky prospect increase, but risk aversion remains the dominant pattern ($p < 0.001$). This may indicate that the increase in the good outcome more than makes up for the slight loss that has been introduced in the bad outcome. Finally, for pure loss choices, subjects are considerably more risk seeking than for gains, and in absolute terms risk neutrality cannot be rejected ($p = 0.19$). It is also commonly found in the literature

Figure 4.1: Stake Effects for Gains and Losses

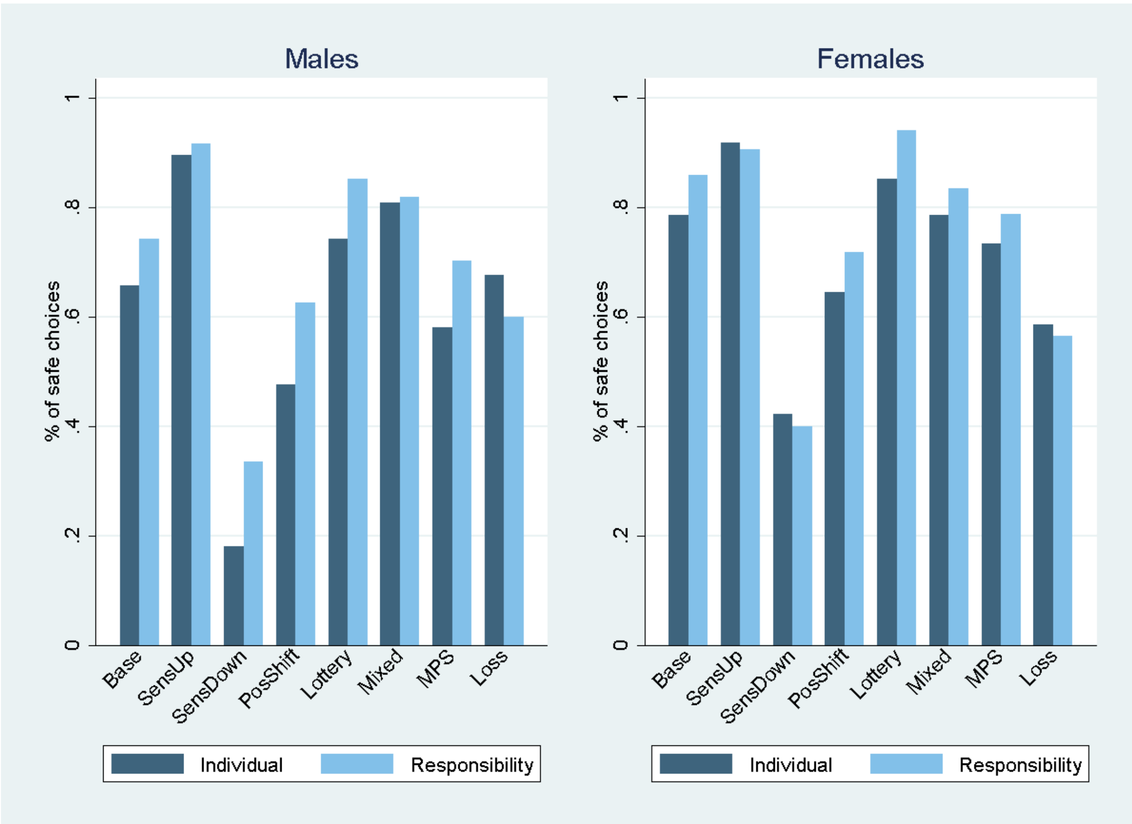



Choice frequency of the safe prospect for basic prospect and loss shift across stake levels

that risk attitudes are influenced by stake levels (Abdellaoui et al., 2011; Binswanger, 1980; Holt and Laury, 2002; Kachelmeier and Shehata, 1992). We thus take a look at the influence of the different stake levels on decisions. Figure 4.1 shows choices for the safe alternative separately for the basic prospect pairs and the pure loss pairs. The stake effect is clearly visible for the basic gain prospects, with increasing expected values resulting in increased levels of risk aversion. Indeed, we cannot reject risk neutrality for the lowest stakes ($p = 0.47$), with risk aversion increasing with stake levels and being highly significant for the highest stake level ($p < 0.001$). For losses, on the other hand, there is no clear trend and risk aversion has only a very slight (and non-significant: $p = 0.31$ for the highest stake level) tendency to increase with absolute stake values.[4] A parametric analysis of these descriptive results can be found in appendix A4.1. We next turn to the differences between the individual and the responsibility treatment.

**Individual Decisions versus Responsibility** Figure 4.2 shows choice frequencies for the safe prospect by treatment, for males and females respectively.[5] One can clearly see how for the base case subjects are more risk averse under responsibility than in the individual decisions - this holds both for males and females.

Figure 4.2: Choice Frequency of the Safe Prospect for Different Prospect Pairs




Frequency of choosing the safe prospect for different prospect pairs, by treatment and gender

[4]The Spearman correlation coefficient between the stake size b and choice for the safe option in the individual treatment is indeed significantly positive for the base lotteries ($p < 0.001$), but not different from zero for losses ($p = 0.57$).

[5]We display the effects by gender because of the large gender effects in risk taking typically found in the literature (Donkers et al., 2001; Eckel and Grossman, 2008), which are also present in our data.

The same tendency is visible in almost all other positive prospect pairs, except for the upward sensitivity prospect pair, in which there is no difference. There is only a very slight indication of responsibility inducing more risk aversion in the mixed prospect pair, while this tendency is again more pronounced for the mean-preserving spread (MPS) pair. For pure loss choices, however, the tendency is inverted, with responsibility decreasing risk aversion. Table 4.2 presents a random effects Probit model regressing choices for the safe prospect on a variety of explanatory variables. Regression I regresses choices on the treatment dummy, a dummy variable indicating the pure loss prospects, a dummy indicating mixed prospects, and two interaction terms between the latter two and the treatment dummy.

Table 4.2: Experiment 1: Choice of Safe Prospect

| **Dep. Var.:** choice of safe prospect | **I** | **II** | **III** |
|---|---|---|---|
| responsibility | 0.070** | 0.080** | 0.099** |
| | (0.036) | (0.037) | (0.049) |
| pure loss | -0.043** | -0.067** | -0.067** |
| | (0.033) | (0.033) | (0.033) |
| responsibility × pure loss | -0.098** | -0.106** | -0.106** |
| | (0.049) | (0.049) | (0.049) |
| mixed prospect | 0.131*** | 0.112*** | 0.112*** |
| | (0.027) | (0.027) | (0.027) |
| responsibility × mixed prospect | -0.024 | -0.032 | -0.032 |
| | (0.051) | (0.051) | (0.051) |
| EV difference | | 0.196*** | 0.197*** |
| | | (0.012) | (0.012) |
| SD difference | | 0.022*** | 0.022*** |
| | | (0.002) | (0.002) |
| female | 0.083** | 0.086** | 0.107** |
| | (0.033) | (0.036) | (0.049) |
| responsibility × female | | | -0.046 |
| | | | (0.076) |
| Constant | ✓ | ✓ | ✓ |
| Obs | 3840 | 3840 | 3840 |
| Subjects | 96 | 96 | 96 |
| Wald $\chi^2$ | 64.43 | 417.03 | 417.27 |

Random effects Probit Regression. Coefficients show marginal effects relative to choices in the individual treatment. *** represents significance at p=0.01, ** at p=0.05, * at p=0.10

Being responsible for somebody else's payoffs as well as one's own increases risk aversion relative to affecting only one's own payoffs; the latter is a *simple main effect*, indicating the effect of responsibility for all prospects except the pure loss prospects (i.e., with the pure loss dummy held *constant at zero*) and the mixed prospect (i.e., with the mixed dummy held *constant at zero*). The effect of the pure loss dummy indicates that for pure loss prospects subjects are more risk seeking compared to all other gain prospects. The interaction between the treatment dummy and the one identifying pure loss prospects indicates that for pure loss prospects the effect of responsibility goes in the opposite direction compared to pure gain prospects, and thus shows that subjects in the responsibility treatment are *more risk seeking* (or less risk averse) for losses compared to subjects in the individual treatment. The significant effect of the mixed-prospect dummy shows that subjects choose the safe option significantly more often for the mixed prospect than for pure gain prospects. The insignificant interaction between the treatment dummy and the mixed dummy on the other hand indicates that there is no significant treatment effects for mixed prospects, with the effect thus going in the same direction as for gains. Finally, we also find that females are significantly more risk averse than males. Such an effect is commonly found for decision making under risk (Donkers et al., 2001; Eckel and Grossman, 2008).

Regression II keeps the same independent variables as regression I, and adds the difference in expected value (defined as the expected value of the safe prospect minus the expected value of the risky prospect) and the difference in standard deviations (defined as the standard deviation of the risky prospect minus the standard deviation of the safe prospect, which is thus always positive). The higher the difference between the safe prospect and the risky prospect in terms of expected value, the more likely subjects will choose the safe prospect. Also, the larger the difference in terms of standard deviation, the more likely subjects are to choose the safer alternative. The main treatment effects discussed above are stable, indicating increased risk aversion under responsibility in the gain domain, increased risk seeking in the loss domain, and no treatment effect in the mixed domain.

Regression III further adds an interaction term between the gender dummy and the treatment dummy. The effect is not significant, which goes to show that being responsible for somebody else does affect males and females in the same way. Once again, all the effects previously discussed remain stable. We next turn to the analysis of the satisfaction ratings of recipients in the responsibility treatment.

**Choice Satisfaction of Recipients** In the responsibility treatment, recipients saw the decision maker's choice with one period lag and indicated whether they were satisfied with the decision or not. Although this rating was not incentivized, it may nevertheless give an indication of the extent to which decision makers adapted their decision to the commonly acceptable one, or correctly intuited which decision would be deemed more acceptable while doing so. Since satisfaction ratings were not communicated to the decision maker and had no influence on payoffs whatsoever, recipients had indeed no reasons to systematically misrepresent their preferences. Also, the fact that providing such ratings was the only occupation of recipients during the experiments leads us to suspect that they took this task seriously. Table 4.3 shows a random effects Probit model regressing

Table 4.3: Experiment 1: Satisfaction Ratings

| **Dep. Var.:** Satisfied with Decision | I | II |
|---|---|---|
| Safe prospect chosen | 0.346*** | 0.208*** |
| | (0.028) | (0.038) |
| Pure loss | 0.046 | -0.002 |
| | (0.036) | (0.040) |
| Pure loss × safe prospect chosen | -0.268*** | -0.174** |
| | (0.075) | (0.072) |
| Mixed prospect | | -0.227*** |
| | | (0.080) |
| Mixed prospect × safe prospect chosen | | 0.187*** |
| | | (0.286) |
| EV difference | | 0.068*** |
| | | (0.012) |
| female | 0.034 | -0.016 |
| | (0.032) | (0.039) |
| female × safe choice | | 0.078** |
| | | (0.037) |
| Constant | ✓ | ✓ |
| Obs | 1920 | 1920 |
| Subjects | 48 | 48 |
| Wald $\chi^2$ | 196.08 | 230.93 |

Random Effects Probit Regression. Coefficients indicate marginal changes in satisfaction levels relative to a choice of the risky prospect; *** represents significance at p=0.01, ** at p=0.05, * at p=0.10

the recipients' satisfaction with each choice on a number of independent variables. The highly significant effect of the safe prospect being chosen by the decision maker shows that safe choices are deemed more satisfactory in the gain domain (this being a simple main effect measuring the effect of safe choices with the pure-loss dummy held constant at zero). While the fact that a prospect offers only negative outcomes per se does not affect satisfaction ratings, choosing the safe amount in pure loss prospects is generally not perceived as satisfactory by recipients, as shown by the highly significant interaction effects of the pure loss and safe choice dummies. This finding confirms that risk seeking is deemed more acceptable than safe choices in the loss domain. There is no main gender effect for satisfaction ratings.

Regression II confirms the stability of the findings we have just discussed, and adds some more variables. The significantly negative main effect of the mixed prospect dummy indicates that choices of the prospect are considered even less satisfactory in the mixed domain as compared to the gain domain. In a parallel fashion, satisfaction increases relative to the pure gain domain when a safe amount is chosen, giving again an indication of loss aversion on the side of recipients. Choices are deemed more satisfactory the higher the difference in expected value, providing an indication that higher differences in expected value increase the agreement between decision makers and recipients on which choice is the best one. Finally, in keeping with previous findings on gender effects, women generally deem choices of the safe prospect as more satisfactory than choices of the risky prospect.

At the end of the experiment we asked subjects to rate their degree of risk aversion on a scale from being very risk seeking (1) to being very risk averse (6). This self-declared risk aversion correlates strongly with the number of safe choices taken in non-negative prospect pairs during the experiment itself on the basis of the Spearman correlation coefficient ($p = 0.01$) across both treatments. Self-declared risk attitudes are not significantly different between the two treatments ($p = 0.26$; Mann-Whitney-U-test, two-sided), nor is there a significant difference between decision makers and recipients in the responsibility treatment ($p = 0.72$; Mann-Whitney-U-test, two-sided). Finally, we also asked subjects to rate themselves according to their risk aversion relative to other participants in the experiment. The rating went from 1 (indicating that a subject considered herself to be amongst the four most risk-loving participants in the session of 24) to 6 (indicating that a subject considered herself to be amongst the four most risk averse participants in the session). On average, decision makers in the responsibility treatment had a rating of 4.17, indicating that they considered themselves more risk averse than the median participant in the experiment, and thus ruling out that they may have considered recipients on average to be more risk averse than they are themselves. This finding corresponds to existing evidence according to which subjects generally consider others as more risk loving than themselves (Hsee and Weber, 1997).

### 4.3.3 Discussion

For gain prospects, we find that responsibility increases risk aversion. An account based on the assumption that decision makers consider others to be more risk averse than they are themselves seems to be ruled out by the answers to the relative risk attitude ranking questions discussed above. Also, Hsee and Weber (1997) found that in a series of different experimental designs subjects systematically predicted others to be less risk averse than themselves. We can thus conclude that subjects do not simply try to adapt their decisions to what they think may be others' risk attitudes.

A different possibility is that subjects comply to an implicit social rule dictating increased caution when responsible for somebody else as well as oneself, thus increasing their risk aversion when responsible for somebody else. This explanation is distinct from the argument discussed in the last paragraph, inasmuch as such a social norm may push subjects to be more risk averse when deciding for others even in cases where they expect that others would be more risk loving than themselves if left to decide for themselves. Such a cautious shift explanation however cannot explain our increased risk seeking for loss prospects. Arguably, different social rules dictating a cautious shift for gains and a 'risky shift' for losses could well exist, but such a hypothesis does have a distinctly ad hoc flavor. Given that individual risk attitudes have been established to be much richer than the simple risk-aversion/risk-seeking dichotomy implicit in such explanations (Abdellaoui, 2000; Abdellaoui et al., 2010; Bleichrodt and Pinto, 2000), we rather hypothesize that risk attitudes typically found in individual decision making are accentuated under conditions of responsibility.

Prospect theory would predict risk aversion to prevail both for medium to large probabilities, so that a theory based on the amplification of the fourfold pattern of risk attitudes predicted by prospect theory cannot be separated from an account based on a social rule favoring increased risk aversion under responsibility based only on the evidence collected for gain prospects. Risk seeking, however, seems to appear more acceptable than risk aversion in the loss domain for the medium probabilities used in our experiment. Evidence in this direction comes both from the behavior of decision makers, who under conditions of responsibility in the loss domain are induced to become more risk seeking rather than more risk averse; and from recipients, who are much more likely to be dissatisfied with a decision in the loss domain when the decision maker chose the sure loss rather than the prospect. This, in turn, cannot be explained by a uniform social norm dictating increased caution under conditions of responsibility.

We thus propose as an alternative hypothesis that the fourfold pattern of risk attitudes predicted by prospect theory-risk aversion for medium to large probability gains and small probability losses, risk seeking for medium to large probability losses and small probability gains-is amplified by responsibility. At this point, the hypothesis that responsibility accentuates the fourfold pattern may be no more plausible than the already discussed hypothesis of different social norms for decisions under gains and under losses. Luckily however, there is a possibility to disentangle such different explanations. The hypothesis of an accentuated fourfold pattern of risk attitudes as found in prospect theory and the social norm argument make very different predictions for different probability levels in the gain domain, which makes it easy to test them against each other. For large probabilities, both prospect theory and the social norm argument predict an increase in risk aversion under conditions of responsibility. For small probabilities, on the other hand, the social norm hypothesis still predicts an increase in risk aversion; quite to the contrary, however, prospect theory and the argument of an amplification of the fourfold pattern laid out above now predict an increase in risk seeking under conditions of responsibility.

The same test can also be adopted to rule out yet another alternative explanation that we cannot rule out on the basis of the results from above. When deciding for others as well as themselves-so the objection goes-decision makers effectively decide over twice the amount of money. Given the common finding that risk aversion increases in stake levels, the increased amounts over which decisions are taken may thus well be the factor underlying the finding of increased risk aversion in the responsibility treatment, rather than the responsibility effect itself. This explanation is indeed plausible for the medium probability gains used in experiment 1 (although it cannot account for the findings for loss prospects). Notice, however, how this explanation would again predict increased risk aversion for small probability gains under higher stakes, which has been found repeatedly (Kachelmeier and Shehata, 1992; Lefebvre et al., 2010). We thus now proceed to testing the effect of responsibility on decisions for different probability levels in the gain domain.

## 4.4 Experiment 2: Disentangling Social Norm and Amplification Accounts

### 4.4.1 Experimental Design

**Subjects.** 180 subjects were recruited from a subject pool of the experimental laboratory MELESSA at Ludwig-Maximilian's University in Munich, Germany, using ORSEE (Greiner, 2004). The experiment was run together with another, unrelated, experiment. 59% of subjects were female, and the average age was 23.88 years.

**Task.** This task was run after another, unrelated experiment.[6] Subjects were asked to choose between a safe option and a risky option in a fashion similar to experiment 1. However, we now only looked at choices in the gain domain. The safe option always consisted in a sure amount of money, while the prospect providing a chance of either 10% or 90% to win 10 €. Overall, subjects had to make 10 choices where the order of presentation was randomized for every subject. Subjects took decisions sequentially and had no opportunity to return to an earlier decision to revise it.

**Prospects.** The choice was always between a sure amount of money and a prospect. There were two prospects, one providing a 10% chance to win 10 € and zero otherwise; and one providing a 90% chance to win 10 € and zero otherwise. The sure amount could take one of five different amounts for each prospect: 0.8, 1, 1.2, 1.5 and 2 € for the 10%

[6] Although the preceding experiment was unrelated, care was taken to distribute the treatments of this experiment orthogonally to the treatments in the other experiment.

prospect, and 7, 8, 8.5, 9, and 9.5 € for the 90% prospect.

**Treatments.** Subjects were randomly assigned to one of two treatments that exactly replicated those of experiment 1: an individual treatment in which subjects took their decisions only for themselves; or a responsibility treatment, in which half of the subjects were randomly assigned the role of decision maker and half the subjects were assigned the role of passive recipient.

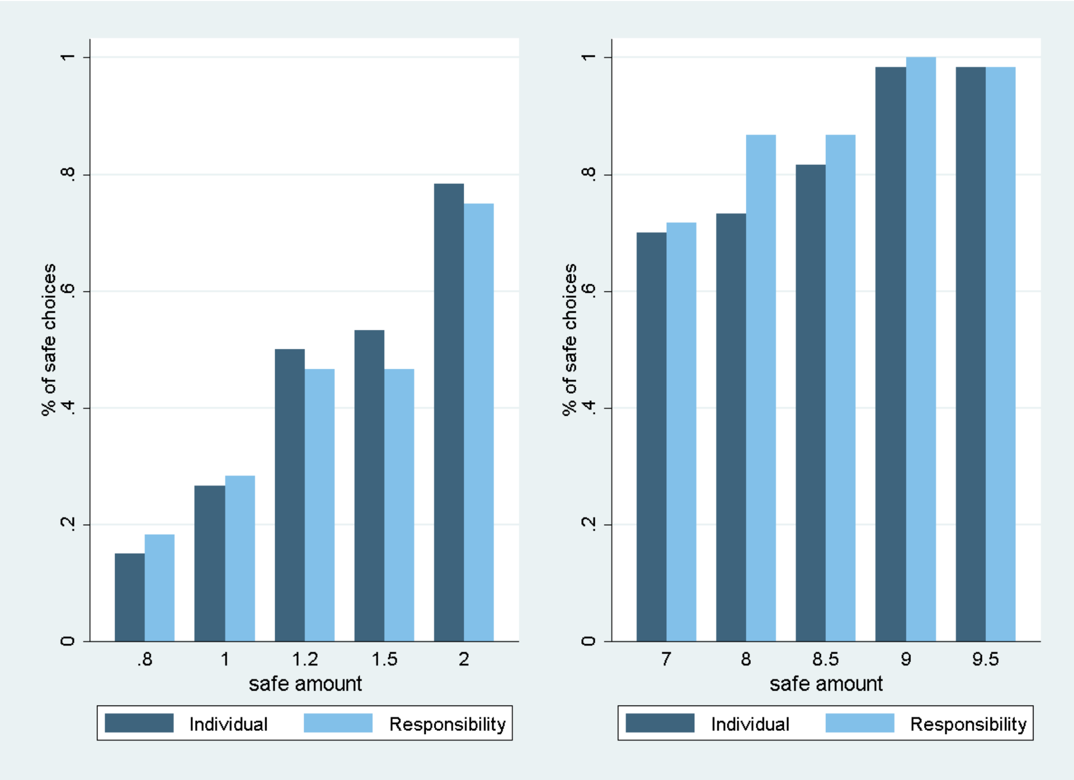
**Incentives.** One decision was randomly extracted to be played for real pay. Since in the unrelated experiment subjects could obtain at least an approximate knowledge about their payoffs, we decided to fully reveal earnings from the experiment in order to be able to control for the exact income effect in a regression (rather than having unknown perceptions of earnings).

### 4.4.2 Results

Figure 4.3 displays the choice frequencies by treatment separately for small and large probabilities. On average we find the typical pattern of risk seeking for small probabilities and risk aversion for large probabilities. Indeed, when the subjects face a choice between a prospect and a sure amount of equal expected value, only about 27% of subjects choose the sure amount for the 10% probability ($p < 0.001$, binomial test), while 99% of subjects do so for the 90% probability ($p < 0.001$, binomial test). For the 10% probability, subjects who are responsible for somebody else choose the sure amount less often for all but the smallest two certain amounts, where choices of the safe amount are generally low. For the 90% probability, responsible subjects always choose the sure amount at least as often as subjects who only decide for themselves.

Table 4.4 presents a random effects Probit model regressing choices of the safe alternative on a variety of explanatory variables. The effect of the responsibility treatment dummy now indicates the simple main effect of being responsible when probabilities are large (Jaccard and Turrisi, 2003). Subjects are thus more likely to choose the sure amount for a 90% probability of winning when responsible compared to the individual treatment.

Figure 4.3: Individual Decisions versus Decisions under Responsibility




Choices of Safe Amount by Treatment for $p = 0.1$ (left) and for $p = 0.9$ (right)

Under small probabilities, subjects are significantly more risk seeking than under large probabilities, as indicated by the highly significant effect of the small probability dummy. More importantly, the interaction of the small-probability dummy with the treatment dummy indicates that this risk-seeking tendency is further enhanced relative to the individual treatment when subjects are responsible for somebody else. As may be expected, the difference in expected value between the sure amount and the prospect (defined as in experiment 1) is also highly significant. Finally, we find a significant, if small, income effect, which goes as expected in the direction of increased risk seeking by subjects who have realized higher earnings from the previous experiment.

Regression II adds two further interaction terms. Almost all effects can be seen to be stable. The gender effect, which had not been significant in regression I, is now also significant: since this is a simple effect, the positive effect of the female dummy now indicates increased risk aversion by females relative to males for large probability prospects. This effect is qualified by the interaction of the female dummy with the

Table 4.4: Experiment 2: Choice of Safe Prospect

| **Dep. Var.:** Choice of Safe Prospect | **I** | **II** |
|---|---|---|
| Responsibility | 0.107* | 0.148* |
| | (0.058) | (0.086) |
| Small probability | -0.666*** | -0.589*** |
| | (0.036) | (0.047) |
| Small probability × responsibility | -0.135** | -0.133** |
| | (0.068) | (0.068) |
| EV difference | 0.295*** | 0.295*** |
| | (0.025) | (0.025) |
| Female | 0.071 | 0.197*** |
| | (0.044) | (0.061) |
| Past profit | -0.008*** | -0.007** |
| | (0.002) | (0.003) |
| Female × small probability | | -0.209** |
| | | (0.068) |
| Past profit × responsibility | | -0.003 |
| | | (0.005) |
| Constant | ✓ | ✓ |
| Obs | 1200 | 1200 |
| Subjects | 120 | 120 |
| Wald $\chi^2$ | 264.55 | 263.57 |

Random Effects Probit Regression. Coefficients show marginal effects relative to choices in the individual treatment; *** represents significance at p=0.01, ** at p=0.05, * at p=0.10

small-probability dummy. The negative effect of that interaction shows that females are significantly *more risk seeking* relative to males for small probabilities. Past profits remain significant, though less so than in regression I. Most importantly, however, there is no interaction effect between past profits from the preceding experiment and our treatment manipulation, showing that this is not interfering with our results.

**Satisfaction ratings** Exactly as in experiment 1, recipients in experiment 2 saw the decisions of their assigned decision maker with a lag of one period, and had to indicate whether they were satisfied with the decision or not. Table 4.5 reports the results of a random effects Probit model regressing the satisfaction dummy on a number of explanatory variables. The first dummy shows the simple main effect of choosing the safe amounts over the large probability prospect: choosing the safe amount for large probability prospects is deemed much more satisfactory in general than choosing the prospect. The dummy indicating the simple main effect of a small probability choice is also positive, indicating

Table 4.5: Experiment 2: Satisfaction Ratings

| **Dep. Var.:** Satisfied with Choice | **I** | **II** |
|---|---|---|
| Safe Choice | 0.530*** | 0.458*** |
| | (0.119) | (0.129) |
| Small probability (10%) | 0.226** | 0.238** |
| | (0.113) | (0.113) |
| Safe choice $\times$ small probability | -0.658*** | -0.680*** |
| | (0.130) | (0.126) |
| EV difference | -0.228*** | -0.233*** |
| | (0.064) | (0.064) |
| EV difference $\times$ safe choice | 0.341*** | 0.345*** |
| | (0.072) | (0.071) |
| Female | 0.124*** | 0.023 |
| | (0.044) | (0.060) |
| Female $\times$ safe choice | | 0.146** |
| | | (0.064) |
| Past profit | | -0.002 |
| | | (0.003) |
| Constant | ✓ | ✓ |
| Obs | 600 | 600 |
| Subjects | 60 | 60 |
| Wald $\chi^2$ | 61.7 | 64.59 |

Random Effects Probit Regression. Coefficients indicate marginal changes in satisfaction levels relative to a choice of the risky prospect. *** represents significance at p=0.01, ** at p=0.05, * at p=0.10

considerable agreement with choices of the prospect in this instance.[7] Choosing however the safe amount for small probability prospects is considered to be very dissatisfying, as shown by the large negative coefficient of the interaction effect. Recipients are in general less satisfied with choices of the prospect the closer the safe amount is to the expected value of the prospect, which is indicated by the simple effect of the relative dummy. They are however more satisfied with a choice of the safe alternative for relatively small deviations in expected value than for small probabilities. Females tend to be much more satisfied when the safe amount is chosen for the large probability prospects, while past profits of the recipients have no influence on satisfaction ratings.

[7] Indeed, the dummy indicates the satisfaction levels for small probabilities with all interactions that include that dummy held constant at zero (Jaccard and Turrisi, 2003). This in turn means that the safe-choice dummy must be zero, thus resulting in the interpretation that the effect indicates satisfaction with choices of the prospect; this satisfaction in turn is measured relative to the (much fewer) choices of the prospect for the large probability prospect.

### 4.4.3 Discussion

The social norm hypothesis and the amplification of fourfold pattern hypothesis make very different predictions on behavior for small probabilities in the gain domain. While for large probabilities both theories predict an increase in risk aversion under responsibility, for small probabilities the social norm argument predicts a cautious shift towards increased risk aversion (or reduced risk seeking), whereas the amplification argument predicts increased risk seeking. Having directly tested these contradictory predictions in experiment 2, we conclude that the social norm dictating a cautious shift under conditions of responsibility has been discredited as an explanation of the results: an increased fourfold pattern of risk attitudes explains our results well. At the same time, this finding also excludes explanations based on which our initial effects could have been due to stake effects rather than responsibility.

While an accentuation of the fourfold pattern of risk attitudes is a good fit for our results, we have not fully proven that such an accentuation takes place. Indeed, we miss results for small probability losses. While such an additional result may seem desirable, our experiment was designed with the explicit purpose of testing two different predictions in the gain domain against each-other. While the fourfold-pattern hypothesis finds strong evidence in our data, it is not impossible that a different explanation could exist for our results. Indeed, even if the interpretation of an increase in the fourfold pattern of risk attitudes prevails-or at least an increase in typical risk attitudes found at the individual level-such an interpretation is merely descriptive in nature. The more fundamental question remains why we observe such a shift in risk attitudes under responsibility.

We can only speculate about the answer at this point. One possibility is to examine the finding in the light of Wegener and Petty (1995) flexible self-correction model. The model postulates that people may shift away from their 'natural' or spontaneous behavior when motivated to do so. The extent to which they correct their behavior, however, as well as the direction in which they correct it, will fundamentally depend on their naïve theory of the bias. This explanation appears however highly unsatisfying, given that there is no way of determining what such unconsciously determined naïve theories of bias may be-with the consequence that such an account could be used to ex-post justify any kind of behavior that one may find. The fact that typical individual risk attitudes are accentuated under conditions of responsibility provides an indication that increased responsibility does by no means push decisions closer to expected utility maximization-generally held to be normative-but rather farther away from it. There seems however to be general agreement on this tendency, as indicated by our satisfaction rating patterns. Indeed in experiment 1, we found recipients to be generally satisfied with safe choices in the gain domain, but dissatisfied with such choices in the loss domain. Given that safe choices have already been found to decrease under conditions of responsibility in the loss domain, this is indeed a strong indication for the perceived social acceptance (or at least desirability) of such choices. A similar pattern can be seen in experiment 2, where safe choices were deemed satisfactory for the large probability prospect, but very unsatisfactory for the small probability prospect.

Whatever the psychological reasons behind our findings may be, the mere economic fact of more extreme patterns under responsibility remains. Such factors may have important consequences for economic predictions and for policy design. Probability weighting-from which the fourfold pattern is thought to derive to a large extent-has been used to explain the simultaneous take-up of insurance and lottery play (Wakker, 2010). The fourfold pattern of risk attitudes has also been used to explain reference point effects that have been observed in financial markets (Baucells et al., 2011; Wiseman and Gomez-Mejia, 1998) and for investment behavior by firms (Fiegenbaum, 1990; Fiegenbaum and Thomas, 1988). Our results provide a further indication that typical risk attitudes found for individuals may not only generalize to professional agents or firms, but even be reinforced to some extent. Given that these patterns seem very resilient to debiasing, explicit rules may be needed to rein in excessive risk taking in certain conditions, or special training programs for managers may seem desirable in order for them not to fall prey to automatic decision making patterns that may be suboptimal from the point of view of the company that they manage.

## 4.5 Conclusion

We systematically explored decision situations in which a decision maker bears responsibility for somebody else's outcomes as well as for her own. In the gain domain, and for medium to large probabilities, we confirmed the intuition that being responsible for somebody else's payoffs increases risk aversion. Looking at risk attitudes in the loss domain, however, we found an increase in risk seeking under conditions of responsibility.

This raises issues about the extent to which changed behavior under responsibility may depend on a social norm of caution in situations of responsibility, or to what extent pre-existing risk attitudes found at the individual level may simply be enhanced under responsibility. To further explore this issue, we designed a second experiment to explore risk-taking behavior for gain prospects offering very small or very large probabilities of winning. For large probabilities, we found increased risk aversion, thus confirming our earlier finding. For small probabilities, on the other hand, we found an increase of risk seeking under conditions of responsibility. The latter finding thus discredits hypotheses of a social rule dictating caution under responsibility, and points towards an amplification in the fourfold pattern of risk attitudes found for individual decisions.

At the present point we can only speculate on what may underlie such an amplification of individual risk attitudes. Additional evidence - possibly from neighboring disciplines such as neuroscience - will probably be needed to fully understand the underlying dynamics. Nevertheless, our findings point out how important and resilient to debiasing these risk attitudes are, and hence the importance of considering them in policy design or for the training and supervising of decision makers.

## 4.6 Appendix A4

### Appendix A4.1: Prospect Type Regression

Table A4.1: Regressions on Prospect Types

| **Dep. Var:** Choice of Safe Prospect | **I** | **II** | **III** |
|---|---|---|---|
| Sensitivity up | 0.185*** | 0.185*** | 0.185*** |
| | (0.023) | (0.023) | (0.023) |
| Sensitivity down | -0.465*** | -0.465*** | -0.485*** |
| | (0.032) | (0.032) | (0.033) |
| Positive shift | -0.161*** | -0.161*** | -0.170*** |
| | (0.034) | (0.034) | (0.035) |
| Lottery choice | 0.101*** | 0.101*** | 0.098*** |
| | (0.026) | (0.026) | (0.026) |
| Mixed Lottery | 0.059** | 0.059** | 0.060** |
| | (0.028) | (0.028) | (0.028) |
| Mean-preserving | -0.065** | -0.65** | -0.069** |
| | (0.032) | (0.032) | (0.033) |
| Loss | -0.171*** | -0.172*** | -0.183*** |
| | (0.034) | (0.032) | (0.035) |
| Female | | 0.080** | 0.080** |
| | | (0.036) | (0.036) |
| Age | | 0.010** | 0.010** |
| | | (0.005) | (0.005) |
| Stake size | | | 0.064*** |
| | | | (0.006) |
| Obs | 3840 | 3840 | 3840 |
| Subjects | 96 | 96 | 96 |
| Wald Chi2 | 510.8 | 515.91 | 608.33 |

Random Effects Probit Regression. Coefficients show marginal effects relative to choices in the basic prospect pair; *** represents significance at p=0.01, ** at p=0.05, * at p=0.10.

Table A4.1 shows a random effects Probit model, with coefficients indicating the deviation of choices with respect to the basic prospect pair. In addition to the effects already discussed in the main text, it shows that females are on average significantly more risk averse than males. Also, risk aversion increases with age. Both findings are commonly found in decision making under risk (Donkers et al., 2001; Eckel and Grossman, 2008). More interestingly, we find an effect of stake size, represented by the expected value of the prospect (taken in absolute terms for the pure loss prospect). The higher the stakes of the decision, the more risk averse subjects become on average. This is in agreement with general findings in the literature (Binswanger, 1980; Kachelmeier and Shehata, 1992;

Lefebvre et al., 2010). Controlling for stake effects also makes the effect of the mixed prospects much more significant. This increased effect derives from the fact that the mixed prospects are obtained by adjusting the expected value of the prospect downward from the basic prospect pair. Since subjects tend to be less risk averse for lower stakes, the increased risk aversion found for mixed prospects appears more relevant once one controls for the decreased stakes in those choice pairs.

# Appendix A4.2: Tables

| | Option A (Safe) | | | | Option B (Risky) | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **Lottery Number** | **Prob Left** | **Amount Left** | **Prob Right** | **Amount Right** | **Prob Left** | **Amount Left** | **Prob Right** | **Amount Right** | **Category** |
| 1 | 1 | 2 | 0 | 0 | 0.5 | 4 | 0.5 | 0 | Base |
| 2 | 1 | 2.5 | 0 | 0 | 0.5 | 4 | 0.5 | 0 | Sensitivity Up |
| 3 | 1 | 1.5 | 0 | 0 | 0.5 | 4 | 0.5 | 0 | Sensitivity Down |
| 4 | 1 | 3 | 0 | 0 | 0.5 | 5 | 0.5 | 1 | Positive Shift |
| 5 | 0.5 | 3 | 0.5 | 1 | 0.5 | 4 | 0.5 | 0 | Lottery Choice |
| 6 | 1 | 0 | 0 | 0 | 0.5 | 2 | 0.5 | -2 | Mixed Prospect |
| 7 | 1 | 2 | 0 | 0 | 0.5 | 5 | 0.5 | -1 | MPS |
| 8 | 0 | 0 | 1 | -2 | 0.5 | 0 | 0.5 | -4 | Loss shift |
| 9 | 1 | 4 | 0 | 0 | 0.5 | 8 | 0.5 | 0 | Base |
| 10 | 1 | 5 | 0 | 0 | 0.5 | 8 | 0.5 | 0 | Sensitivity Up |
| 11 | 1 | 3 | 0 | 0 | 0.5 | 8 | 0.5 | 0 | Sensitivity Down |
| 12 | 1 | 6 | 0 | 0 | 0.5 | 10 | 0.5 | 2 | Positive Shift |
| 13 | 0.5 | 6 | 0.5 | 2 | 0.5 | 8 | 0.5 | 0 | Lottery Choice |
| 14 | 1 | 0 | 0 | 0 | 0.5 | 4 | 0.5 | -4 | Mixed Prospect |
| 15 | 1 | 4 | 0 | 0 | 0.5 | 10 | 0.5 | -2 | MPS |
| 16 | 0 | 0 | 1 | -4 | 0.5 | 0 | 0.5 | -8 | Loss shift |
| 17 | 1 | 6 | 0 | 0 | 0.5 | 12 | 0.5 | 0 | Base |
| 18 | 1 | 7.5 | 0 | 0 | 0.5 | 12 | 0.5 | 0 | Sensitivity Up |
| 19 | 1 | 4.5 | 0 | 0 | 0.5 | 12 | 0.5 | 0 | Sensitivity Down |
| 20 | 1 | 9 | 0 | 0 | 0.5 | 15 | 0.5 | 3 | Positive Shift |
| 21 | 0.5 | 9 | 0.5 | 3 | 0.5 | 12 | 0.5 | 0 | Lottery Choice |
| 22 | 1 | 0 | 0 | 0 | 0.5 | 6 | 0.5 | -6 | Mixed Prospect |
| 23 | 1 | 6 | 0 | 0 | 0.5 | 15 | 0.5 | -3 | MPS |
| 24 | 0 | 0 | 1 | -6 | 0.5 | 0 | 0.5 | -12 | Loss shift |
| 25 | 1 | 8 | 0 | 0 | 0.5 | 16 | 0.5 | 0 | Base |
| 26 | 1 | 10 | 0 | 0 | 0.5 | 16 | 0.5 | 0 | Sensitivity Up |
| 27 | 1 | 6 | 0 | 0 | 0.5 | 16 | 0.5 | 0 | Sensitivity Down |
| 28 | 1 | 12 | 0 | 0 | 0.5 | 20 | 0.5 | 4 | Positive Shift |
| 29 | 0.5 | 12 | 0.5 | 4 | 0.5 | 16 | 0.5 | 0 | Lottery Choice |
| 30 | 1 | 0 | 0 | 0 | 0.5 | 8 | 0.5 | -8 | Mixed Prospect |
| 31 | 1 | 8 | 0 | 0 | 0.5 | 20 | 0.5 | -4 | MPS |
| 32 | 0 | 0 | 1 | -8 | 0.5 | 0 | 0.5 | -16 | Loss shift |
| 33 | 1 | 10 | 0 | 0 | 0.5 | 20 | 0.5 | 0 | Base |
| 34 | 1 | 12.5 | 0 | 0 | 0.5 | 20 | 0.5 | 0 | Sensitivity Up |
| 35 | 1 | 7.5 | 0 | 0 | 0.5 | 20 | 0.5 | 0 | Sensitivity Down |
| 36 | 1 | 15 | 0 | 0 | 0.5 | 25 | 0.5 | 5 | Positive Shift |
| 37 | 0.5 | 15 | 0.5 | 5 | 0.5 | 20 | 0.5 | 0 | Lottery Choice |
| 38 | 1 | 0 | 0 | 0 | 0.5 | 10 | 0.5 | -10 | Mixed Prospect |
| 39 | 1 | 10 | 0 | 0 | 0.5 | 25 | 0.5 | -5 | MPS |
| 40 | 0 | 0 | 1 | -10 | 0.5 | 0 | 0.5 | -20 | Loss shift |

Table A4.2: List of all 40 used Lotteries

# Appendix A4.3: Figures

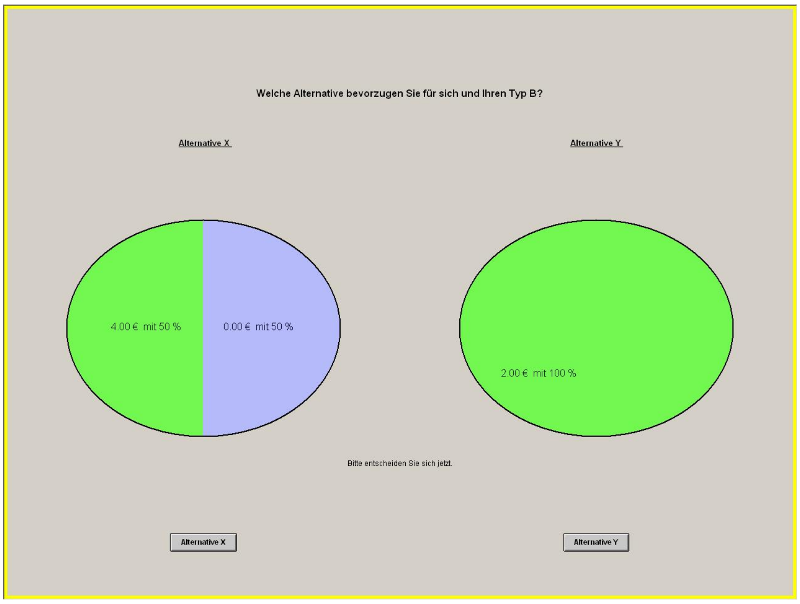



Figure A4.1: Gain Lottery

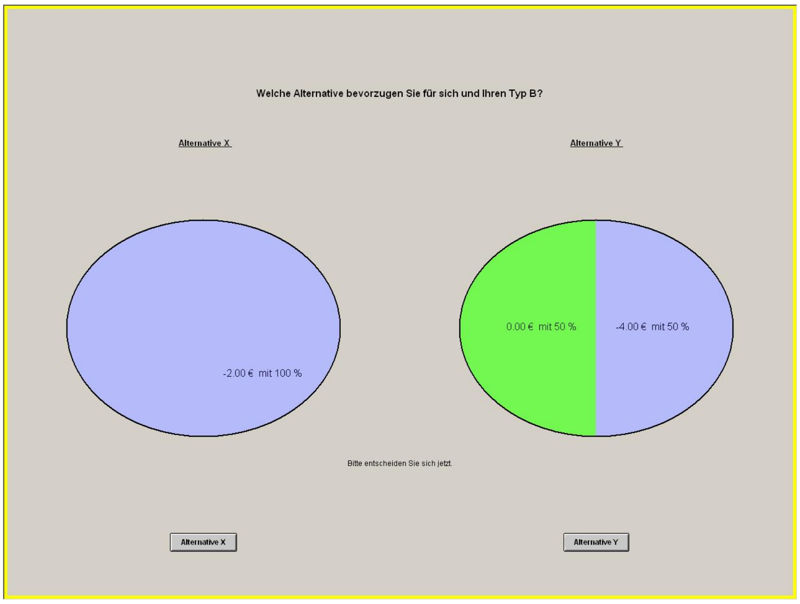



Figure A4.2: Loss Lottery

IND / RESP

# Appendix A4.4: Instructions (translated from German)

This experiment serves the investigation of economic decision making. You can earn money which will be paid to you in cash after the experiment. During the experiment you and all other participants will be asked to make decisions. In total, the experiment lasts for approximately **1 hour and 15 minutes**. Please raise your hand in case you have any questions during the experiment. One of the experimenters will then come to you and answer your questions in private. In the interest of clarity, we use male terms only in the instructions.

**Payment**

You receive 4 Euro for arriving in time in addition to your earnings from the experiment. There is a possibility that you suffer losses from specific decisions. Possible losses must be offset with your earnings from other decision situations and/or with your 4 Euro starting balance.
In (the very unlikely) case of an overall loss from the experiment, you may choose between paying it back in cash or by working as an assistant in the laboratory (5 Euro per half an hour).

**Support**

You are provided with a pen on your desk. Please type your decisions into the computer. While making your decisions, there is a clock counting down in the right upper corner of your computer screen. This clock serves as a guide for how much time it should take for you to make your decisions. Of course, you are allowed to exceed the time; particularly in the beginning, this may be happening quite frequently. Once time has run out, it is only the pure information screens, which do not ask you to make any decisions that will be dismissed.

---

## Lottery decision making

[IND: You do not interact with other participants of the experiment at any point during the experiment. Your final payment is determined exclusively by your own decisions and according to the rules explained in the following. Other participants do not find out about your decisions and about how much you have earned at any point during or after the experiment. In the same manner, you do not learn about other participants' decisions and their earnings at any point during or after the experiment.]

[RESP: You will be matched with another participant of the experiment. Your decisions <u>or</u> the decisions of the other participant determine your payment according to the rules explained in the following. At no point during or after the experiment other participants in the experiment learn your identity. In the same manner, you do not find out the identity of other participants at any point during or after the experiment.]

### Task

[RESP: There are two types of participants, **type A and type B**. The matching is such that a type A person is always matched with a type B person. At the beginning your computer screen will tell you which type you are. **The decision on which type you are is made randomly by the computer**. You will remain the same type throughout the experiment.

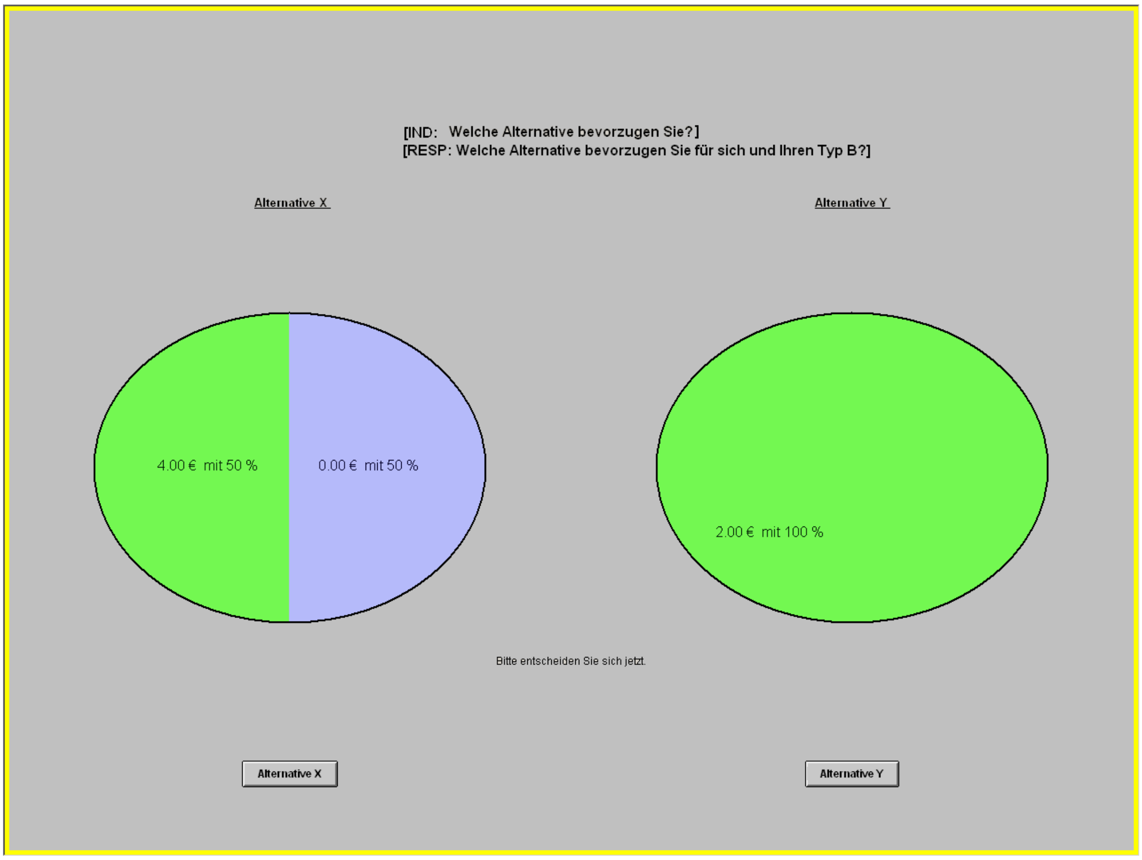
**Decisions are made by type A only**. Participants of type A make their decision **for themselves and at the same time for their matching partner of type B**. This means that **every decision that applies for type A applies to his matching partner of type B in exactly the same way**.]

**In total, there are 40 periods.** [IND: You] [RESP: Type A persons] have to make **one decision <u>per</u> period** for which [IND: you] [RESP: they] always have to choose between two alternatives.

Page 1 of 4

IND / RESP

A representative decision scenario may look like the following:



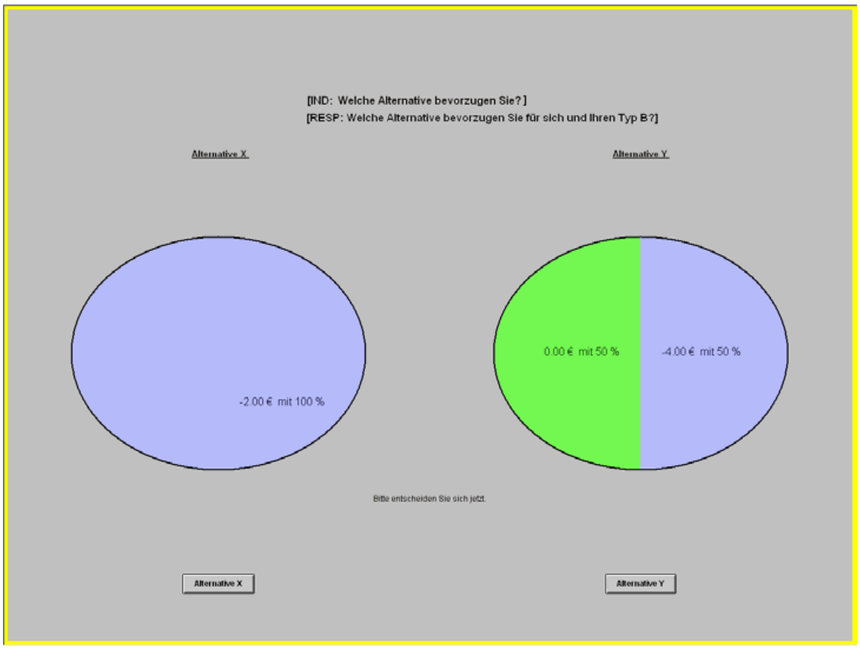

In the above example, [IND: you have] [RESP: type A player has] a choice between **alternative X**, that yields **4 Euro** with a probability of **50%** and **0 Euro** with the complementary probability of **50%** [IND: to you] [RESP: to him and to his matching partner of type B], and **alternative Y**, that yields **2 Euro** with a probability of **100%** [IND: to you] [RESP: to him and to his matching partner of type B]. [IND: You decide] [RESP: Type A player decides] on one of the two alternatives by clicking on either the button "Alternative X" or the button "Alternative Y" below the pie charts.

An alternative such as alternative Y from the above example is called a **"certain payment"** since it is paid out with a probability of 100%. An alternative such as alternative X is called **"lottery"** since one amount is paid out with a probability of 50 % and another amount is paid out with a probability of 50%.
The alternatives between which [IND: you have] [RESP: type A has] to choose in each period either represent a **choice between a certain payment and a lottery**, or a **choice between two different lotteries**. In both alternatives there may be **positive as well as negative amounts** involved.

Page 2 of 4

IND / RESP

A decision scenario involving negative amounts may look like the following:




In this example, [IND: you have] [RESP: type A player has] a choice between **alternative X**, that yields **-2 Euro** (a loss of 2 Euro) with a probability of **100%** [IND: to you] [RESP: to him and to his matching partner of type B], and **alternative Y**, that yields **0 Euro** with a probability of **50%** and **-4 Euro** (a loss of 4 Euro) with a complementary probability of **50%** [IND: to you] [RESP: to him and his matching partner of type B].

[RESP: **Type B players are provided with the information on the decisions of their type A partner with a lag of one period**. This means that type B players see the decision scenario on their screens with which their type A partner was confronted in the previous period and are told the alternative which their type A partner chose. Finally, type B players can indicate whether they were "content" with the decision or "not content". The statements of contentment do not influence type B's earnings or the earnings of his type A partner. **The statements of contentment do not get passed on to type A.]**

## Payment

**[IND: It is in your interest** to think thoroughly about each decision **because** each single decision may determine your payment at the end of the experiment.]

[RESP: If you are a type A player, it is in your interest **to think thoroughly about each decision** because **each single decision may determine your payment as well as the payment of your type B partner** at the end of the experiment.]
This happens as follows:

To determine final payments the computer randomly selects **three different periods that are relevant for the payment at the end of the experiment**. Each period is **equally likely to be selected** by the computer. The sum of the earnings from the three selected periods determines [IND: your final payment] [RESP: the final payment for type A as well as for this type B partner].
[IND: On your screen you get told which periods got selected at random and how you chose in these periods.]
[RESP: All participants are told on their screens which periods got selected at random and how type A chose in these periods.]

In case [IND: you] [RESP: type A] chose a **certain payment** in a selected decision period, [IND: you] [RESP: type A and his type B partner] receive the amount of the certain payment as [IND: your] [RESP: their] earning from this selected period.

IND / RESP

In case [IND: you] [RESP: type A] chose a **lottery**, the outcome of the lottery has to be determined first. To this end, lottery numbers from 1 to 6 get assigned to the possible earning amounts. As there are only lotteries involving probabilities of 50%, lottery numbers 1, 2 and 3 get assigned to one amount and lottery numbers 4, 5 and 6 get assigned to the other amount. The computer randomly determines which amount gets assigned to the low numbers and which amount gets assigned to the high numbers. Finally, a **randomly chosen participant is asked to roll a 6-sided die in public**. The amount corresponding to the lottery number that was rolled is then paid out for the selected period.

***Example 1:*** *The computer selects a period in which [IND: you] [RESP: type A] chose alternative X which yields **4 Euro** with a probability of 50% and **0 Euro** with a probability of 50%. Lottery numbers **1, 2 and 3 were assigned to the amount of 4 Euro** and numbers **4, 5 and 6 were assigned to the amount of 0 Euro** by the computer. [IND: You] [RESP: Type A and his type B partner] thus have a 50% chance to receive 4 Euro and a 50% chance to receive 0 Euro. If, for example, the lottery number 1 is rolled, the earnings from this period amount to **4 Euro** [IND: for you] [RESP: for type A and for his type B partner]. If, for example, the lottery number 5 is rolled, the earnings from this period amount to **0 Euro** [IND: for you] [RESP: for type A and for his type B partner].*

***Example 2:*** *The computer selects a period in which [IND: you] [RESP: type A] chose alternative Y which yields **-4 Euro** (a loss of 4 Euro) with a probability of 50% and **0 Euro** with a probability of 50%. Lottery numbers **1, 2 and 3 were assigned to the amount of -4 Euro** and numbers **4, 5 and 6 were assigned to the amount of 0 Euro** by the computer. [IND: You] [RESP: Type A and his type B partner] thus have a 50% chance to receive **0 Euro** and a 50% chance to receive **-4 Euro** (a loss of 4 Euro). If, for example, the lottery number 4 is rolled, the earnings from this period amount to **0 Euro** [IND: for you] [RESP: for type A and for his type B partner]. If, for example, the lottery number 3 is rolled, the earnings from this period amount to **-4 Euro** [IND: for you] [RESP: for type A and for his type B partner]. This loss must be offset with earnings from other decisions and/or with your starting balance of 4 Euro.*

Your payment is formed by the sum of your earnings in the three selected periods.

[RESP: **Two participants that are matched with each other** (type A and his type B partner) always have **identical earnings and thus final payments.]**

Please note that it is optimal [IND: for you] [RESP: for type A] to choose the alternative that [IND: you prefer for yourself] [RESP: he prefers for himself and for his type B partner].

There is **no possibility to increase the final payment by adopting a different behavior.**

# Bibliography

Abdellaoui, M. 2000. Parameter-free elicitation of utility and probability weighting functions. *Management Science* 46:1497–1512.

Abdellaoui, M., Baillon, A., Placido, L., and Wakker, P. 2011. The rich domain of uncertainty: Source functions and their experimental implementation. *American Economic Review* 101:699–727.

Abdellaoui, M., Bleichrodt, H., and Paraschiv, C. 2007. Loss aversion under prospect theory: A parameter-free measurement. *Management Science* 53:1659–1674.

Abdellaoui, M., L'Haridon, O., and Zank, H. 2010. Separating curvature and elevation: A parametric probability weighting function. *Journal of Risk and Uncertainty* 41:39–65.

Abeler, J., Altmann, S., Kube, S., and Wibral, M. 2010. Gift exchange and workers' fairness concerns: When equality is unfair. *Journal of the European Economic Association* 8:1299–1324.

Akerlof, G. 1982. Labor contracts as partial gift exchange. *Quarterly Journal of Economics* 97:543–569.

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

Akerlof, G. A. and Yellen, J. L. 1988. Fairness and unemployment. *American Economic Review* 78:44–49.

Altmann, S., Falk, A., and Huffman, D. 2009. Contract enforcement, fairness and involuntary unemployment. mimeo.

Ambrus, A. and Pathak, P. A. 2011. Cooperation over finite horizons: A theory and experiments. *Journal of Public Economics* 95:500–512.

Angelova, V., Güth, W., and Kocher, M. 2011. Co-employment of permanently and temporarily employed agents. *Labor Economics* forthcoming:.

Autor, D. H. and Scarborough, D. 2008. Does job testing harm minority workers? Evidence from retail establishments. *Quarterly Journal of Economics* 123:219–277.

Baker, G., Gibbons, R., and Murphy, K. 2002. Relational contracts and the theory of the firm. *Quarterly Journal of Economics* 117:39–84.

Bardsley, N. and Moffatt, P. 2007. The experimetrics of public goods: Inferring motivations from contributions. *Theory and Decision* 62:161–193.

Bartling, B., Fehr, E., and Schmidt, K. M. 2011. Screening, competition, and job design: Economic origins of good jobs. *American Economic Review* forthcoming:.

Baucells, M., Weber, M., and Welfens, F. 2011. Reference-point formation and updating. *Management Science* 57:506–519.

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

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

Binswanger, H. P. 1980. Attitudes toward risk: Experimental measurement in rural India. *American Journal of Agricultural Economics* 62:395–407.

Bleichrodt, H. and Pinto, J. L. 2000. A parameter-free elicitation of the probability weighting function in medical decision analysis. *Management Science* 46:1485–1496.

Blinder, C. and Choi, D. 1990. A shred of evidence on theories of wage stickiness. *Quarterly Journal of Economics* 105:1003–1015.

Bohnet, I., Greig, F., Herrmann, B., and Zeckhauser, R. 2008. Betrayal aversion: Evidence from Brazil, China, Oman, Switzerland, Turkey, and the United States. *American Economic Review* 98:294–310.

Bolton, G. E. and Ockenfels, A. 2000. ERC: A theory of equity, reciprocity, and competition. *American Economic Review* 90:166–193.

Bolton, G. E. and Ockenfels, A. 2010. Betrayal aversion: Evidence from Brazil, China, Oman, Switzerland, Turkey, and the United States: Comment. *American Economic Review* 100:628–33.

Booij, A., van Praag, B., and van de Kuilen, G. 2010. A parametric analysis of prospect theory's functionals for the general population. *Theory and Decision* 68:115–148.

Borghans, L., Duckworth, A. L., Heckman, J. J., and ter Weel, B. 2008a. The economics and psychology of personality traits. *Journal of Human Resources* 43:972–1059.

Borghans, L., Meijers, H., and ter Weel, B. 2008b. The role of noncognitive skills in explaining cognitive test scores. *Economic Inquiry* 46:2–12.

Bornstein, G. 1992. The free-rider problem in intergroup conflicts over step-level and continuous public goods. *Journal of Personality and Social Psychology* 62:597–606.

Bornstein, G. 2003. Intergroup conflict: Individual, group, and collective interests. *Personality and Social Psychology Review* 7:129–145.

Bornstein, G. and Ben-Yossef, M. 1994. Cooperation in intergroup and single-group social dilemmas. *Journal of Experimental Social Psychology* 30:52–67.

Bornstein, G., Erev, I., and Rosen, O. 1990. Intergroup competition as a structural solution for social dilemma. *Social Behavior* 5:247–260.

Bornstein, G., Gneezy, U., and Nagel, R. 2002. The effect of intergroup competition on group coordination: an experimental study. *Games and Economic Behavior* 41:1–25.

Bowles, S. 1985. The production process in a competitive economy: Walrasian, neo-hobbesian, and marxian models. *American Economic Review* 75:16–36.

Brandts, J. and Charness, G. 2004. Do labour market conditions affect gift exchange? Some experimental evidence. *Economic Journal* 114:684–708.

Brandts, J. and Charness, G. 2011. The strategy versus the direct-response method: A first survey of experimental comparisons. *Experimental Economics* 14:375–398.

Brandts, J. and Schram, A. 2001. Cooperation and noise in public goods experiments: Applying the contribution function approach. *Journal of Public Economics* 79:399–427.

Brown, M., Falk, A., and Fehr, E. 2004. Relational contracts in the nature of market interactions. *Econometrica* 72:747–780.

Brown, M., Falk, A., and Fehr, E. 2011. Competition and relational contracts: The role of unemployment as a disciplinary device. *Journal of the European Economic Association* forthcoming:.

Bull, C. 1987. The existence of self-enforcing implicit contracts. *The Quarterly Journal of Economics* 102:147–159.

Burlando, R. and Guala, F. 2005. Heterogeneous agents in public goods experiments. *Experimental Economics* 8:35–54.

Cabrales, A. and Charness, G. 2003. Optimal contracts, adverse selection and social preferences: An experiment. Working Paper, University of California at Santa Barbara.

Cabrales, A., Miniaci, R., Piovesan, M., and Ponti, G. 2010. Social preferences and strategic uncertainty: An experiment on markets and contracts. *American Economic Review* 100:2261–78.

Campbell III, C. and Kamlani, K. 1997. The reasons for wage rigidity: Evidence from a survey of firms. *Quarterly Journal of Economics* 112:759–789.

Charness, G. 2004. Attribution and reciprocity in an experimental labor market. *Journal of Labour Economics* 22:665–688.

Charness, G. and Jackson, M. O. 2009. The role of responsibility in strategic risk-taking. *Journal of Economic Behavior & Organization* 69:241–247.

Charness, G. and Kuhn, P. 2011. Lab labor: What can labor economists learn from the lab?, pp. 229–330. *In* O. Ashenfelter and D. Card (eds.), Handbook of Labor Economics, volume 4, Part A of *Handbook of Labor Economics*. Elsevier.

Charness, G. and Rabin, M. 2002. Understanding social preferences with simple tests. *Quarterly Journal of Economics* 117:817–869.

Chaudhuri, A. 2011. Sustaining cooperation in laboratory public goods experiments: A selective survey of the literature. *Experimental Economics* 14:47–83.

Cooper, D. and Kagel, J. 2009. Other regarding preferences: A selective survey of experimental results, pp. 1–77. *In* J. Kagel and A. Roth (eds.), Handbook of Experimental Economics Results, volume 2. Elsevier.

Dittrich, D. and Ziegelmayer, A. 2010. Laboratory bilateral gift exchange: Heterogeneous social preferences. mimeo.

Dohmen, T. and Falk, A. 2011. Performance pay and multidimensional sorting: Productivity, preferences, and gender. *American Economic Review* 101:556–90.

Dohmen, T., Falk, A., Huffman, D., and Sunde, U. 2009. Homo reciprocans: Survey evidence on behavioural outcomes. *Economic Journal* 119:592–612.

Donkers, B., Melenberg, B., and Van Soest, A. 2001. Estimating risk attitudes using lotteries: A large sample approach. *Journal of Risk and Uncertainty* 22:165–195.

Dufwenberg, M., Heidhues, P., Kirchsteiger, G., Riedel, F., and Sobel, J. 2011. Other-regarding preferences in general equilibrium. *The Review of Economic Studies* 78:613–639.

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

Eckel, C. C. and Grossman, P. J. 2008. Men, women and risk aversion: Experimental evidence, pp. 1061–1073. *In* C. R. Plott and V. L. Smith (eds.), Handbook of Experimental Economics Results, volume 1, chapter 113. Elsevier.

Eckel, C. C. and Petrie, R. 2011. Face value. *American Economic Review* 101:1497–1513.

Englmaier, F. and Leider, S. 2009. Managerial payoff and gift exchange in the field. mimeo.

Englmaier, F. and Leider, S. 2011. Contractual and organizational structure with reciprocal agents. *American Economic Journal: Microeconomics* forthcoming:.

Englmaier, F., Strasser, S., and Winter, J. 2011. Worker characteristics, contracting, and wage differentials: Evidence from a gift-exchange experiment. mimeo.

Englmaier, F. and Wambach, A. 2010. Optimal incentive contracts under inequity aversion. *Games and Economic Behavior* 69:312–328.

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.

Falk, A. 2007. Gift exchange in the field. *Econometrica* 75:1501–1511.

Falk, A., Fehr, E., and Zehnder, C. 2006. Fairness Perceptions and reservation wages - The behavioral effects of minimum wage laws. *Quarterly Journal of Economics* 121:1347–1381.

Falk, A., Huffman, D., and MacLeod, B. 2008. Institutions and contract enforcement. CESifo Working Paper No. 2284.

Fehr, E. and Falk, A. 1999. Wage rigidity in a competitive incomplete contract market. *Journal of Political Economy* 107:106–134.

Fehr, E. and Falk, A. 2008. Reciprocity in experimental markets, pp. 325–334. *In* C. R. Plott and V. L. Smith (eds.), Handbook of Experimental Economics Results, volume 1, chapter 38. Elsevier.

Fehr, E. and Gächter, S. 2000. Fairness and retaliation: The economics of reciprocity. *Journal of Economic Perspectives* 14:159–181.

FEHR, E., GÄCHTER, S., AND KIRCHSTEIGER, G. 1997. Reciprocity as a contract enforcement device: Experimental evidence. *Econometrica* 65:833–860.

FEHR, E., KIRCHLER, E., WEICHBOLD, A., AND GACHTER, S. 1998a. When social norms overpower competition: Gift exchange in experimental labor markets. *Journal of Labor Economics* 16:324–351.

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., KIRCHSTEIGER, G., AND RIEDL, A. 1998b. Gift exchange and reciprocity in competitive experimental markets. *European Economic Review* 42:1–34.

FEHR, E., KLEIN, A., AND SCHMIDT, K. M. 2007. Fairness and contract design. *Econometrica* 75:121–154.

FEHR, E. AND SCHMIDT, K. 1999. A theory of fairness, competitition and cooperation. *Quarterly Journal of Economics* 114:817–868.

FEHR, E. AND SCHMIDT, K. 2003. Theories of fairness and reciprocity - Evidence and economic ppplications, pp. 208–257. *In* M. Dewatripont, L. P. Hansen, and S. J. Turnovsky (eds.), Advances in Econonomics and Econometrics, volume 1, chapter 6. Cambridge University Press, 8th World Congress of the Econometric Society edition.

FIEGENBAUM, A. 1990. Prospect theory and the risk-return association: An empirical examination in 85 industries. *Journal of Economic Behavior & Organization* 14:187–203.

FIEGENBAUM, A. AND THOMAS, H. 1988. Attitudes toward risk and the risk-return paradox: Prospect theory explanations. *The Academy of Management Journal* 31:85–106.

FISCHBACHER, U. 2007. Z-tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics* 10:171–178.

Fischbacher, U. and Gächter, S. 2009. The behavioral validity of the strategy method in public good experiments. Discussion Papers 2009-25, The Centre for Decision Research and Experimental Economics, School of Economics, University of Nottingham.

Fischbacher, U. and Gächter, S. 2010. Social preferences, beliefs, and the dynamics of free riding in public goods experiments. *American Economic Review* 100:541–56.

Fischbacher, U., Gächter, S., and Fehr, E. 2001. Are people conditionally cooperative? Evidence from a public goods experiment. *Economics Letters* 71:397–404.

Fuchs, W. 2007. Contracting with repeated moral hazard and private evaluations. *American Economic Review* 97:1432–1448.

Gächter, S., Nosenzo, D., and Sefton, M. 2010. Peer effects in pro-social behavior: Social norms or social preferences? Discussion Papers 2010-23, The Centre for Decision Research and Experimental Economics, School of Economics, University of Nottingham.

Gächter, S., Nosenzo, D., and Sefton, M. 2011. The impact of social comparisons on reciprocity. *Scandinavian Journal of Economics* forthcoming:.

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

Gerhards, L. and Heinz, M. 2011. In good times and bad - reciprocal behavior in times of crisis under incomplete contracts. mimeo.

Gintis, H. 1976. The nature of labor exchange and the theory of capitalist production. *Review of Radical Political Economics* 8:36–54.

Gneezy, U. and List, J. 2006. Putting behavioural economics to work: Testing for gift exchange in labor markets using field experiments. *Econometrica* 74:1356–1384.

Goeree, J. and Yariv, L. 2007. Conformity in the lab. mimeo.

Greiner, B. 2004. The online recruitment system ORSEE 2.0 - A guide for the organization of experiments in economics. University of Cologne, Working Paper Series in Economics.

Gunnthorsdottir, A. and Rapoport, A. 2006. Embedding social dilemmas in intergroup competition reduces free-riding. *Organizational Behavior and Human Decision Processes* 101:184–199.

Hansen, G. 1985. Indivisible labor and the business cycle. *Journal of Monetary Economics* 16:309–327.

Heckman, J. J., Stixrud, J., and Urzua, S. 2006. The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics* 24:411–482.

Heineck, G. and Anger, S. 2010. The returns to cognitive abilities and personality traits in Germany. *Labour Economics* 17:535–546.

Herrmann, B. and Thöni, C. 2009. Measuring conditional cooperation: A replication study in Russia. *Experimental Economics* 12:87–92.

Holt, C. A. and Laury, S. K. 2002. Risk aversion and incentive effects. *American Economic Review* 92:1644–1655.

Hsee, C. K. and Weber, E. U. 1997. A fundamental prediction error: Self-others discrepancies in risk preference. *Journal of Experimental Psychology: General* 126:45–53.

Huang, F. and Cappelli, P. 2010. Applicant screening and performance-related outcomes. *American Economic Review* 100:214–218.

Ichniowski, C., Shaw, K., and Prennushi, G. 1997. The effects of human resource management practices on productivity: A study of steel finishing lines. *American Economic Review* 87:291–313.

Isaac, R. M. and Walker, J. M. 1988. Communication and free-riding behavior: The voluntary contribution mechanism. *Economic Inquiry* 26:585–608.

Isaac, R. M., Walker, J. M., and Thomas, S. H. 1984. Divergent evidence on free riding: An experimental examination of possible explanations. *Public Choice* 43:113–149.

Jaccard, J. and Turrisi, R. 2003. Interaction effects in multiple regression. Sage Publications.

Kachelmeier, S. J. and Shehata, M. 1992. Examining risk preferences under high monetary incentives: Experimental evidence from the People's Republic of China. *American Economic Review* 82:1120–1141.

Köbberling, V. and Wakker, P. P. 2005. An index of loss aversion. *Journal of Economic Theory* 122:119–131.

Kirchsteiger, G., Niederle, M., and Potters, J. 2001. Public versus private exchanges. *CentER Discussion Papers* 101:1–39.

Kocher, M., Luhan, W., and Sutter, M. 2010. Give every man his due? Relational contracts with individual and standardized wages. mimeo.

Kocher, M. G., Cherry, T., Kroll, S., Netzer, R. J., and Sutter, M. 2008. Conditional cooperation on three continents. *Economics Letters* 101:175–178.

Kosfeld, M. and von Siemens, F. A. 2011. Competition, cooperation, and corporate culture. *The RAND Journal of Economics* 42:23–43.

Kreps, D., Milgrom, P., Roberts, J., and Wilson, R. 1982. Rational cooperation in the finitely repeated prisoners' dilemma. *Journal of Economic Theory* 27:245–255.

Kube, S., Maréchal, M.-A., and Puppe, C. 2010. Do wage cuts damage work morale? Evidence from a natural field experiment. University of Zurich Working Paper No. 471.

Kurzban, R. and DeScioli, P. 2008. Reciprocity in groups: Information-seeking in a public goods game. *European Journal of Social Psychology* 38:139–158.

Kurzban, R. and Houser, D. 2005. Experiments investigating cooperative types in humans: A complement to evolutionary theory and simulations. *Proceedings of the National Academy of Sciences of the United States of America* 102:1803–1807.

Kydland, F. and Prescott, E. 1982. Time to build and aggregate fluctuations. *Econometrica* 50:1345–1370.

Lazear, E. and Rosen, S. 1981. Rank-order tournaments as optimum labor contracts. *The Journal of Political Economy* 89:841–864.

Ledyard, J. 1995. Public goods: A survey of experimental research. *In* J. Kagel and A. Roth (eds.), Handbook of Experimental Economics. Princeton University Press.

Lefebvre, M., Vieider, F. M., and Villeval, M. C. 2010. Incentive effects on risk attitude in small probability prospects. *Economics Letters* 109:115–120.

Levin, J. 2003. Relational incentive contracts. *American Economic Review* 93:835–857.

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

Liebrand, W. B. G. and McClintock, C. G. 1988. The ring measure of social values: A computerized procedure for assessing individual differences in information processing and social value orientation. *European Journal of Personality* 2:217–230.

Linardi, S. and Camerer, C. 2010. Can relational contracts survive stochastic interruptions? Caltech Social Science Working 1340.

Long, J. and Plosser, C. 1983. Real business cycles. *Journal of Political Economy* 91:39–69.

Lucas, R. 1977. Understanding business cycles. *Carnegie-Rochester Conference Series on Public Policy* 5:7–29.

MacLeod, B. 2003. Optimal contracting with subjective evaluation. *American Economic Review* 93:216–240.

MacLeod, B. and Malcomson, J. 1989. Implicit contracts, incentive compatibility, and involuntary unemployment. *Econometrica* 57:447–480.

MacLeod, B. and Malcomson, J. 1993. Wage premiums and profit maximization in efficiency wage models. *European Economic Review* 37:1223–1249.

MacLeod, B. and Malcomson, J. 1998. Motivation and markets. *American Economic Review* 88:388–411.

Marwell, G. and Ames, R. E. 1981. Economists free ride, does anyone else?: Experiments on the provision of public goods. *Journal of Public Economics* 15:295–310.

Maximiano, S., Sloof, R., and Sonnemans, J. 2007. Gift exchange in a multi-worker firm. *The Economic Journal* 117:1022–1050.

Mittone, L. and Ploner, M. 2009. Social effects in a multi-agent investment game: An experimental analysis. CEEL Working Paper 5.

Muller, L., Sefton, M., Steinberg, R., and Vesterlund, L. 2008. Strategic behavior and learning in repeated voluntary contribution experiments. *Journal of Economic Behavior & Organization* 67:782–793.

Murnane, R. J., Willett, J. B., and Levy, F. 1995. The growing importance of cognitive skills in wage determination. *The Review of Economics and Statistics* 77:251–66.

Offerman, T., Sonnemans, J., and Schram, A. 1996. Value orientations, expectations and voluntary contributions in public goods. *Economic Journal* 106:817–45.

Park, E.-S. 2000. Warm-glow versus cold-prickle: A further experimental study of framing effects on free-riding. *Journal of Economic Behavior & Organization* 43:405–421.

Payne, J. W., Laughhunn, D. J., and Crum, R. 1980. Translation of gambles and aspiration level effects in risky choice behavior. *Management Science* 26:1039–1060.

Payne, J. W., Laughhunn, D. J., and Crum, R. 1981. Further tests of aspiration level effects in risky choice behavior. *Management Science* 27:953–958.

Post, T., van den Assem, M. J., Baltussen, G., and Thaler, R. H. 2008. Deal or no deal? Decision making under risk in a large-payoff game show. *American Economic Review* 98:38–71.

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

Reuben, E. and Tyran, J.-R. 2010. Everyone is a winner: Promoting cooperation through all-can-win intergroup competition. *European Journal of Political Economy* 26:25–35.

Roe, B. and Wu, S. 2009. Do the selfish mimic cooperators? - Experimental evidence from finitely-repeated labor markets. IZA Discussion Paper No. 4084.

Rohde, I. and Rohde, K. 2010. Risk attitudes in a social context. mimeo.

Schmidt, K. M. 2010. Social preferences and competition. Discussion Papers in Economics 11313, University of Munich, Department of Economics.

Schmidt, U. and Zank, H. 2005. What is loss aversion? *Journal of Risk and Uncertainty* 30:157–167.

Schoemaker, P. J. H. 1990. Are risk-attitudes related across domains and response modes? *Management Science* 36:1451–1463.

Shapiro, C. and Stiglitz, J. 1984. Equilibrium unemployment as a worker discipline device. *American Economic Review* 74:433–444.

Starmer, C. 2000. Developments in non-expected utility theory: The hunt for a descriptive theory of choice under risk. *Journal of Economic Literature* 38:332–382.

Sutter, M. and Strassmair, C. 2009. Communication, cooperation and collusion in team tournaments - An experimental study. *Games and Economic Behavior* 66:506–525.

Tan, J. H. and Bolle, F. 2007. Team competition and the public goods game. *Economics Letters* 96:133–139.

Trautmann, S. and Vieider, F. 2010. Social influences on risk attitudes: Economic applications. mimeo.

Tversky, A. and Kahneman, D. 1992. Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty* 5:297–323.

van Dijk, F., Sonnemans, J., and van Winden, F. 2002. Social ties in a public good experiment. *Journal of Public Economics* 85:275–299.

Wakker, P. 2010. Prospect theory for risk and uncertainty. Cambridge University Press.

Wegener, D. T. and Petty, R. E. 1995. Flexible correction processes in social judgment: The role of naive theories in corrections for perceived bias. *Journal of Personality and Social Psychology* 68:36–51.

Weimann, J. 1994. Individual behaviour in a free riding experiment. *Journal of Public Economics* 54:185–200.

Wiseman, R. M. and Gomez-Mejia, L. R. 1998. A behavioral agency model of managerial risk taking. *The Academy of Management Review* 23:133–153.

Zelmer, J. 2003. Linear public goods experiments: A meta-analysis. *Experimental Economics* 6:299–310.

# Eidesstattliche Versicherung

Ich versichere hiermit eidesstattlich, dass ich die vorliegende Arbeit selbständig und ohne fremde Hilfe verfasst habe. Die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sowie mir gegebene Anregungen sind als solche kenntlich gemacht. Die Arbeit wurde bisher keiner anderen Prüfungsbehörde vorgelegt und auch noch nicht veröffentlicht.

23. September 2011

Sebastian Strasser

# Curriculum Vitae

| | |
|---|---|
| seit September 2007 - | Doktorandenstudium in Ökonomie<br>Munich Graduate School of Economics, LMU München |
| September 2006 -<br>August 2007 | Master of Science in Economics & Finance<br>University of Warwick, Großbritannien |
| September 2004 -<br>Juli 2006 | Diplom-Betriebswirt (FH)<br>European School of Business, Reutlingen |
| September 2002 -<br>Juli 2004 | Diplome d'Etudes Supérieures Européennes de Management<br>Ecole Superieure de Commerce, Reims, Frankreich |
| Juni 2001 | Abitur<br>Christoph-Jacob-Treu Gymnasium, Lauf a.d. Pegnitz |
| Mai 1982 | Geboren in Sulzbach-Rosenberg, Deutschland |