On Traits, Attributes, and Incentives – Motivating Heterogeneous Agents

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

2017

vorgelegt von

Katharina Christine Schüßler

Referent: Prof. Dr. Florian Englmaier
Korreferent: Prof. Stephen Leider, PhD
Promotionsabschlussberatung: 31. Januar 2018

Tag der mündlichen Prüfung: 16.01.2018

Namen der Gutachter: Florian Englmaier, Stephen Leider, Simeon Schudy

Acknowledgments

At this point, I would like to thank all the people who have made this dissertation possible. First and foremost, I would like to thank my supervisor, Florian Englmaier, who not only convinced me to pursue a PhD in economics and provided invaluable assistance along the way, but also made this journey much more fun. Without his constant enthusiasm and unrestricted support I wouldn't be where I am today. I am also grateful to my second advisor Stephen Leider, who not only made it possible for me to visit the University of Michigan, but also provided invaluable feedback on early-stage ideas, presentations, papers, and academic life in general. I also thank Simeon Schudy for his advice and for being the third member of my thesis committee.

In addition, I would like to thank all the other people who also provided invaluable support in writing this dissertation, my co-author Matthias Fahn, my colleagues at the MGSE, especially Max Breu and Raphael Guber, and the whole Micro research group at LMU – thank you all for discussing research ideas, problems, and papers with me. To Anne, Maria, Max, Michi, Nadine, Raphael, Sarah: thanks for being my support and my friends! You made the PhD journey worthwhile. Also, I owe a big thank you to Manu Beckstein and Julia Zimmermann for helping me wrestle with all administrative issues. In addition, I feel very grateful to have been part of the EBE graduate program, to be part of a highly motivated cohort and towards our program coordinators, Joachim Winter and Florian Englmaier. Funding through the Elitenetwork of Bavaria and the CRC TR 190 "Rationality and Competition" are gratefully acknowledged.

Last but not least: I am greatly indebted to my family and friends who stayed by my side during my academic journey and supported me with encouragement, faith, and love. You know who you are.

Contents

List of Figures

LIST OF FIGURES

List of Tables

Preface

Traditionally, the focus in economic research has been on explaining changes in behavior as reactions to changes in constraints like prices, information, taxes, costs, and incentives. Situational factors have been put center stage, and identification of treatment differences has been the predominant goal, which is why economics provides powerful methodological tools to achieve this aim. However, this point of view builds on the assumption that preferences are stable and homogenous, which implies that behavior changes can be fully explained by changes in the environment. Thus, for a long time, differences in preferences were discarded as non-explanations; a position that is neatly summarized in the infamous paper "de gustibus non est distputandum" by Stigler and Becker (1977), who argue that it is sensible to treat tastes as stable and similar among individuals. Their conclusion is that then, "one does not argue over tastes for the same reason that one does not argue over the Rocky Mountains – both are there, will be there next year, too, and are the same to all men." (Stigler and Becker, 1977, p.76).

While challenging the neoclassical notion of rationality by introducing non-standard preferences to economic models, the focus of behavioral economists largely remained on situational constraints. In fact, some of the most influential behavioral economists take the rather extreme standpoint that economic preference parameters or personal traits neither possess predictive validity, stability, nor causality. Instead, they believe behavior to be almost entirely determined by constraints and incentives (see the discussion in Almlund et al., 2011). This position can be traced back to the beginnings of behavioral economics as a movement: To enhance the arguably narrow view of neoclassical economics on human motivation (Fehr and Falk, 2002; Camerer, 2004), behavioral economists explored the neighboring field of psychology, thereby being influenced by

the predominant spirit of the time in psychological research. While the exploration of individual differences in personality had been an important domain in psychology from the start, early accounts focused on the reductive trait-based approach. The fundamental assumption of this approach – that behavior is stable over time and situations – was only challenged with the influential monograph by Walter Mischel (1968), who stressed the importance of situational determinants of behavior. In light of popular behavioristic reasoning at that time, the ensuing person-situation debate left the field in limbo for the following decades. As a consequence, the subsequently predominant view in psychological research was that of experimental social psychological approaches focusing on the importance of the situation. Thus, the situational focus of behavioral economics reflects the development in psychology at the time which has influenced how economics view individual heterogeneity. As Richard Thaler puts it, "[t]he great contribution to psychology by Walter Mischel [...] is to show that there is no such thing as a stable personality trait." (Thaler, 2008).[2]

By now, however, personality psychology has emerged again as a thriving research field, and economic scholars from different fields have also started to argue that integrating individual heterogeneity with respect to preferences and personality is meaningful to further economists' understanding of incentives as it enlarges the way economists describe the world (Borghans et al., 2008; Rustichini, 2009; Almlund et al., 2011). This significance is established by the existence of pronounced heterogeneity and its predictive power for economic outcomes. More specifically, compelling evidence for pronounced heterogeneity has been brought forward, not least by behavioral economics scholars themselves, for instance in research on social preferences (Fehr and Schmidt, 2006) or risk preferences (Bruhin et al., 2010). While personality traits have been linked to important outcomes like life and job satisfaction, occupational choice and performance (Barrick and Mount, 1991; Ozer and Benet-Martînez, 2006; Mueller and Plug, 2006; Heineck and Anger, 2010; Moffitt et al., 2011; Cobb-Clark and Tan, 2011; Caliendo et al., 2014), economic preferences like risk and social preferences have pre-

[2] This is a rather extreme conclusion to draw from the person-situation debate, and most psychologists would not agree (Bowers, 1973). It stresses the rigorous position early behavioral economists took.

dictive power for these outcomes as well (Dohmen et al., 2008, 2009, 2011). In addition, economic preferences and personality traits have been argued to be complementary (Becker et al., 2012).

This dissertation consists of four essays that all incorporate this general notion while exploring different aspects and utilizing different methods. In Chapter 1, the general point is made that individual heterogeneity in traits and preferences might give rise to complementarities for human-resource management (HRM) practices, which in turn constitute an explanation for persistent productivity differences between seemingly similar firms. In contrast, Chapters 2 to 4 focus on different facets of individual heterogeneity and their impact on various economic outcomes: In Chapter 2, the role of intrinsic reciprocity for dynamic contracting is investigated. Chapter 3 focuses on the role of individual heterogeneity in personality traits on contribution behavior in a public goods context, and Chapter 4 explores the role of heterogeneity in beliefs on competitiveness. To provide a comprehensive picture, the chapters make use of various methods including theoretical modeling, econometric analysis of survey data, and laboratory experiments.

Chapter 1, which is joint work with Florian Englmaier, is an overview article in which we stress the relevance of complementarities of HRM practices for explaining persistent productivity differences across organizations. We do so by drawing on the literature on strategic human-resource management and incorporate an additional aspect into the debate: by introducing individual-level heterogeneity in preferences among employees, insights from behavioral economics point to an additional dimension of potential complementarities. To illustrate these central points, we first give an overview of the concepts of complementarities and HRM practices before focusing on empirical methods to study both. Here, we concentrate on the World Management Survey (Bloom et al., 2014) on the one hand and insider-econometric studies (Ichniowski and Shaw, 2003) on the other hand, which constitute two broad approaches to investigate the complementarity of HRM practices. Based on this, we then proceed to show how insights from behavioral economics can inform the analysis. To this end, we develop a simple agency model illustrating how social preferences influence the design and effects

of incentive schemes. More specifically, in a hidden-action principal-agent framework with limited liability, we show that the principal is able to reduce the incentive intensity and the wage bill when contracting with a more reciprocal agent and illustrate that the composition of the workforce and the effect of HRM practices are interdependent. In addition, we investigate how auxiliary HRM practices can strengthen this interaction. We broaden our central argument by discussing other behavioral sub-fields that are also suited to inform research on complementarities. Finally, we give an outlook on further research.

Chapter 2, which is joint work with Matthias Fahn and Anne Schade, focuses on the role of reciprocal behavior in employment relationships. In this paper, we address the question if and how two predominant explanations for reciprocal behavior, inherent preferences for reciprocity and repeated interaction, interact. Developing a theoretical model of a long-term employment relationship, we first show that reciprocal preferences are more important when an employee is close to retirement. At earlier stages, repeated interaction is more important because relatively more future rents are available to provide incentives. Intrinsic reciprocity is only formally used in contracts in the last stages of the employment relationships when future rents are too small to make repeated interaction feasible. Preferences for reciprocity still affect the structure of an employment relationship early on, though, because of two reasons: First, preferences for reciprocity effectively reduce the employee's effort costs. Second, they allow to relax the enforceability constraint that determines the principal's commitment in the repeated interaction. Therefore, reciprocity-based and repeated-game incentives are dynamic *substitutes*, but *complements* at any given point in time. We test our main predictions using data from the German Socio-Economic Panel (SOEP) and find evidence for a stronger positive effect of positive reciprocity on effort for older workers and for workers that are close to retirement.

In Chapter 3, which is joint work with Michael Schüßler and Daniel Mühlbauer, our aim is to investigate factors that influence sequential contribution decisions in a public goods context. Following the notion that organizations often face public good dilemmas, we use a real-time provision-point mechanism to explore the process of achieving

cooperative equilibria in the laboratory. We develop and test two related sets of hypotheses. First, we discern how institutional changes (i.e., step return, availability of real-time information) influence contribution behavior in the lab. Second, by drawing on models of collective behavior, the theory of mixed-motive social dilemmas, and personality psychology, we test the link between individual heterogeneity (i.e., personality traits, distributional preferences) and the probability to contribute early or late. We find that while step return influences group success, availability of information about the behavior of fellow players only has an influence in combination with the former. In addition, individuals' distributional preferences as well as conscientiousness and extraversion help to explain the observed contribution sequences.

In Chapter 4, I investigate whether two mechanisms leading to biased beliefs about success, overconfidence and competition neglect, influence decisions to enter competitive environments. I use a controlled laboratory setting that allows to elicit belief distributions related to absolute as well as relative overconfidence to study it comprehensively. In addition, my setting allows me to introduce systematic variation in the possibility for competition neglect. Studying these two mechanisms simultaneously is especially relevant as both overconfidence and competition neglect potentially lead to the same behavioral pattern, namely excess entry. However, for de-biasing, it is essential to know whether individuals overestimate their chances of success because they overestimate their performance in absolute terms or relative to others, or because they do not realize that they face a selected set of competitors. With this in mind, I let individuals decide whether they want to enter a competition in a real-effort experiment while eliciting their beliefs with respect to their own and others' performance as well as their winning probability. I introduce two treatment variations: First, some participants receive detailed performance feedback addressing absolute and relative overconfidence before making their decision. Comparing entry decisions of individuals who have received feedback to those who have not enables me to investigate whether overconfidence plays a role for selection into competition – if there is a difference between those groups, it does. Second, I vary whether the competition group consists of all potential competitors or only of individuals who also chose to compete. Here, finding no difference in entry between the

groups implies that decision makers fail to adjust their decisions to the selected sample, thus exhibiting competition neglect. I find that there is systematic heterogeneity in perception biases, with low-performing individuals overestimating their own performance and their chances of success while underestimating performance in the competition group, while the opposite is true for high-performing individuals. While these biases in performance beliefs are ameliorated by feedback, individuals persistently disregard the composition of the competition group they face. Investigating determinants of entry decisions to tackle the key question of this paper whether competition neglect and overconfidence influence entry into competition, I find that both influence individuals' decisions. However, choices are closely tied to previous performance and assessments, and there are no gender differences.

All four chapters in this dissertation are self-contained: they can be read autonomously as they all contain an independent introduction and discussion. The appendices and a joint bibliography are provided after the main text.

Chapter 1

Complementarities of HRM Practices: A Case for a Behavioral Economics Perspective*

1.1 Motivation

One of the most important developments in the study of firms in economics and management has been the increasing availability of firm-level micro-data and the ensuing emphasis on firm heterogeneity. Newly available large and detailed data sets have provided strong evidence for enormous and persistent heterogeneity of firms (and workers) over a range of dimensions, even in narrowly defined industries. These observations are starkly at odds with theoretical predictions and are commonly referred to as "persistent productivity differences" (PPDs) across firms (cf. Syverson, 2011). Theory predicts that competitive forces will induce firms to quickly adopt innovations, only leaving room for short-term heterogeneity, but not for the persistent patterns that emerge in the data. Indeed, PPDs are evident even in seemingly uniform industries without room for differentiation, with high competition and with minor frictions. While these differences are quite impressive for developed countries like the U.S., where a firm at the 10th percentile of the productivity distribution generates only half of the output that a firm at the 90th percentile is able to generate with the same input (Syverson, 2004), they seem

* This chapter is based on joint work with Florian Englmaier which has already been published under the same title in the Journal of Institutional and Theoretical Economics (Englmaier and Schüßler, 2016). Reprinted with permission from Mohr Siebeck Tübingen.

to be even more pronounced for countries like India and China, where the average total factor productivity differentials between the 10th and 90th percentile are larger than 1:5 (Hsieh and Klenow, 2009).

In fact, researchers have come up with several explanations that include, for instance, differences in input quality (e.g., Abowd et al., 2005), learning-by-doing (e.g., Benkard, 2000), or differences in management practices (e.g., Bloom et al., 2014); however, even when controlling for these facts, there still remains substantial variation (see Syverson, 2011). From the perspective of firms, increasing productivity and thus ensuring to be (and stay) on the upper tail of the productivity distribution can be seen as a central goal, which gives rise to the question of how that can be achieved. By drawing on the literature on strategic human resource management (SHRM), we argue that complementary human resource management (HRM) practices are an important part of the answer and thus also important for understanding PPDs. Specifically, strategic management in general has been traditionally focused on how to achieve a sustained competitive advantage, which can be depicted as the antecedent of PPDs (Baron and Kreps, 1999; Lockett and Thompson, 2001). Besides that, the notion of complementarities is essentially the same as the notion of fit, which is a prominent theme in SHRM: while internal fit points to the fact that the implementation of matching practices can yield convex returns, external fit makes the same claim for taking into account external contingencies (Kepes and Delery, 2007).

In this paper, we want to build on this idea and incorporate an additional aspect into the debate on complementary effects of HRM practices as a possible mechanism for bringing about PPDs. Behavioral economics highlights additional dimensions of potential complementarities by introducing individual level heterogeneity in preferences among employees. Assuming that firms can recruit from a heterogeneous labor force has stark consequences for complementary effects of HRM practices and can result in fundamentally different systems of practices; e.g., recruiting workers with strong social preferences, much authority, little monitoring, and relatively mute explicit incentives vs. recruiting workers with no social preferences, very formalized processes with little authority and relatively strong explicit incentives. These starkly different systems of

practices may yield comparably high returns as long as they are tailored to the composition of the workforce. Thus, considerations based on behavioral economics give rise to substantially richer interactions.

We give a brief summary over different views in SHRM in Section 1.2 before focusing on two macro-level approaches to measuring the impact of HRM practices used in organizational economics: the world management survey (WMS) and “insider econometrics" studies (cf. Ichniowski and Shaw, 2003). In Section 1.3, we then demonstrate what behavioral economics is able to contribute by making a case for focusing more on micro-foundations. We do this by developing a simple model introducing agents with social preferences to the standard principal-agent framework, relating our insights to empirical findings, and posing new questions. Building on these, in Section 1.4, we describe what we identify as an empirical agenda and conclude.

1.2 Review of the Literature

This section comprises of two parts: first, we aim at a brief, concise explanation of the general view on complementarities and HRM practices; then, we provide an overview of empirical methods to study complementarities of HRM practices.[1]

1.2.1 Concepts

Complementarities

As Brynjolfsson and Milgrom (2013) point out, complementarities can be thought of as a set of interrelated decisions a firm has to take in order to function properly. Assume for example a firm follows a low-cost strategy; then, subsequent decisions concerning

[1] It is important to note that this article is not meant to be an exhaustive overview of the extensive literature on the World Management Survey, insider econometric studies, or on complementarities of HRM practices. We rather provide a brief summary of these literatures to act as a background for developing our main argument – that behavioral economics insights can contribute to the study of complementarities of HRM practices – and based on that sketch a research agenda. Almost necessarily, in doing so we paint a subjective picture and brazenly over-represent own work throughout the paper. For excellent reviews see, e.g., Ichniowski and Shaw (2013), Bloom et al. (2014), or Jackson et al. (2014).

the acquisition of and the investment in human capital depend on this strategy. It can be assumed that the decisions to keep hiring and training costs low are indeed complementary to the strategy decision. Hölmstrom and Milgrom (1994) describe the general pattern of practices being complementary when "using one more intensely increases the marginal benefit of using others more intensely" (Hölmstrom and Milgrom, 1994, p.973).[2]

Obviously, this way of thinking can be applied to a variety of situations. As shown in the example above, in organizational economics, thinking about complementarities between organizational decisions or practices has proven useful for explaining predominant practice patterns as systems of complements that then appear together (see Brynjolfsson and Milgrom, 2013, for a concise treatment of complementarity in organizations).

Strategic Human Resource Management and The Impact of HRM Practices

In a similar vein but largely unnoticed by economists, scholars in SHRM have been investigating the impact of HRM practices on organizational-level outcomes in general for more than 20 years (Huselid, 1995; Becker and Gerhart, 1996; Becker and Huselid, 2006). In this still ongoing discourse, increasingly sophisticated theoretical formulations have been proposed (e.g., Delery, 1998; Gerhart, 2007; Jiang et al., 2012).[3] In general, three different theoretical approaches addressing the link between HRM and firm performance have been proposed: the universalistic, the contingency, and the configurational approach.

In early work, an universalistic perspective has been taken which follows the proposition that there exists a relationship between the adoption of particular HRM practices and increased organizational performance (Delery and Doty, 1996). Assuming homogeneous organizations, an adoption of those "best practices" is expected to increase firm performance independent of any contingencies (Huselid, 1995). In principle, the litera-

[2] Aside from this insightful and straightforward intuition, economists have defined complementarities with mathematical precision using the concept of supermodularity (Milgrom and Roberts, 1995).

[3] See Jackson et al. (2014) for a discussion of this literature.

ture building on the World Management Survey also adopts an universalistic approach.

In contrast, the contingency approach suggests that HRM practices should be aligned with important contingency factors such as labor markets, competition, or firm strategies. Organizational performance is assumed to be fostered by the interaction between the external fit to contingencies and the set of HRM practices in place (Delery and Doty, 1996). Accordingly, the use of HRM practices lacking this external fit may result in ambiguity, which in turn reduces organizational performance (Schuler and Jackson, 1987; Schuler, 1989).

Lastly, the configurational approach assumes that HRM needs complex alignment with both external and internal contingency factors such that complex and idiosyncratic sets of practices at different levels originate (Delery and Doty, 1996). The underlying assumption of this perspective is that the use of a coherent system of mutually reinforcing HRM practices has greater effects on organizational performance than the sum of each individual practice effect (cf. the literature following the insider econometrics approach, e.g., Ichniowski et al., 1997). Note that this notion of complementary practices and synergies is almost identical to the idea of complements in organizational economics and in line with the formal definition of supermodularity. Furthermore, the notion of complementary HRM practices has also been raised and discussed by economists (Pfeffer, 1994; Baron and Kreps, 1999; Lazear and Shaw, 2007).

1.2.2 Empirics

After having discussed the general view on complementarities and HRM practices, we now focus on giving a brief overview of empirical approaches to identify complementarities and their role in explaining firm productivity.

As already described in the preceding paragraphs, there is a plethora of theories in SHRM on complementary HRM practices. However, although theorists keep on emphasizing the importance of an (internal) fit of different practices, direct empirical tests remain scarce (Cappelli and Neumark, 2001; Gerhart, 2007; Kepes and Delery, 2007). Throughout this section, our main focus is on economic approaches.

Traditionally, case studies on single firms have been used to build a firm intuition about underlying mechanisms in the complementarity-productivity relation – prominent examples include cases like Lincoln Electric's business methods and incentive design (Berg and Fast, 1975) as well as specific changes like the introduction of digital imaging technologies (Autor et al., 2002) or of an enterprise resource planning (ERP) system (McAfee, 2002).

However, as this approach can be misleading because of lacking generalizability (Brynjolfsson and Milgrom, 2013), several other methods have been used to systematically study these effects. In the following, we focus on two broad approaches, namely the WMS and insider econometrics studies.

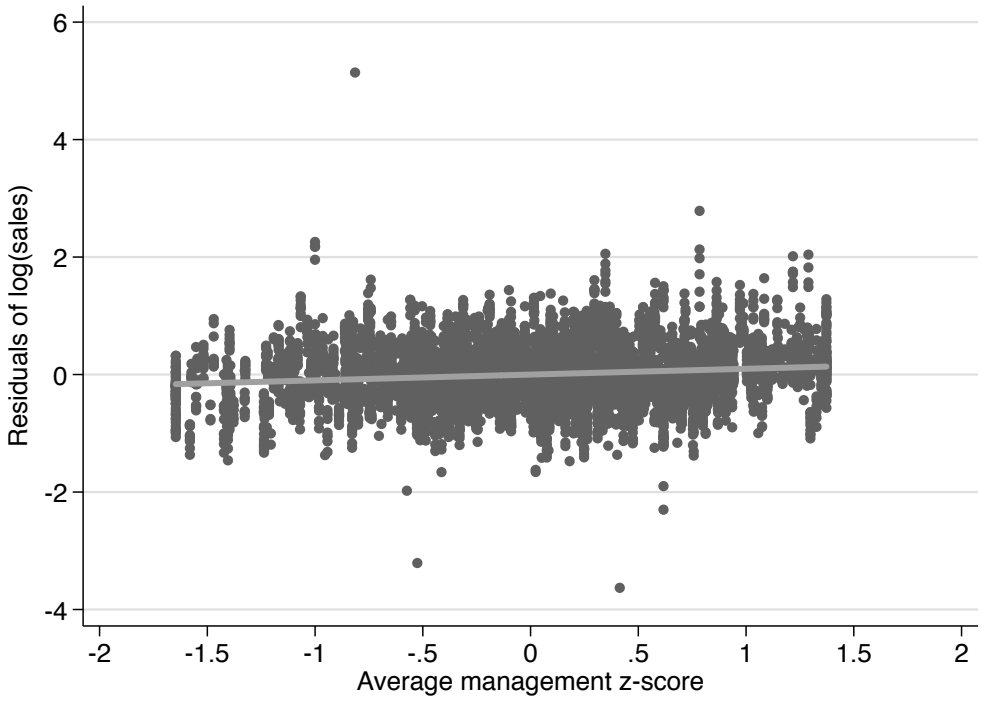
World Management Survey

Over the last decade, the WMS, a survey run by a team of researchers around Nicholas Bloom and John Van Reenen, has led to numerous important insights to explain PPDs across firms. In particular, the WMS explores how "good" management practices can explain firm heterogeneity and focuses on monitoring, targets, and incentives, to explore the impact of management practices on firm productivity in different sectors and countries; for an overview, see Bloom et al. (2014).[4] In closely related work, these authors have documented complementarities between (what they argue constitute) "good" management practices and more general firm investment behavior, namely in IT (Bloom et al., 2012). They document that good management in the above sense makes IT investments more profitable and show that a significant share of the productivity advantage of US firms over European firms can be explained by IT usage together with "better" management.

Although some of the heterogeneity across firms can be explained by the insights provided by WMS data, a lot of unexplained variance is left. This drawback is illustrated in

[4] While the WMS proper is a compilation of a large number of semi-structured telephone interviews, starting out as cross-sectional but recently also building up a panel component, the correlational evidence from the WMS is recently corroborated by smaller randomized control trials (RCTs); see, e.g., Bloom et al. (2013).

Figure 1.1: *Residual plot of log(sales) on average management scores*



Notes: The figure plots the residuals of a regression of the log number of sales on the log number of employees, a key performance measure used by Bloom and Van Reenen (2007), and the average z-standardized management scores of the surveyed firms. Data are generously provided at: www.http://worldmanagementsurvey.org/. The figure is based on own calculations.

Figure 1.1, which is based on the data used in Bloom and Van Reenen (2007), and plots the residuals of a regression of the log number of sales on the log number of employees, one of the performance measures used by Bloom and Van Reenen (2007), against the average z-standardized management scores of the surveyed firms. As the observations scatter widely around the regression line, the average management score still seems to be a noisy measure, even when controlling for firm size. Hence, to generalize from this picture, a lot of unexplained variation is left to be explained.

Concerning HRM practices, an important drawback is that the WMS focuses on a specific set of HRM practices rather than depicting the whole system of HRM practices of a firm. Furthermore, only mere correlative patterns are observed which are not able to depict complementarities or internal fit, but only show which practices tend to be used together. Hence, even if the WMS data are helpful in explaining some of the variation

in productivity, substantial amounts of PPDs remain unexplained. A particular aspect of HRM practices that is not at the center of the WMS but that might matter a lot, is their complementarity. The study of these complementarities has been at the center of insider econometrics studies, covered in the next section.

Insider Econometrics

Insider econometrics studies aim at identifying the performance contribution of bundles of HRM practices more closely (Ichniowski and Shaw, 2003). This within-firm work has suggested that the specific combination of complementary HRM practices enhances productivity. Aimed at producing empirical estimates of the value of alternative HRM practices, numerous studies in this field indicate that various innovative practices are complements. The defining characteristic of a wide range of studies that can be summarized as belonging to the insider econometrics literature – see, for example, Ichniowski et al. (1997); Lazear (2000); Hamilton et al. (2003); Shearer (2004); Bandiera et al. (2007, 2009); Wolf and Zwick (2008) – is that they use highly detailed, production-line specific data to tie HRM practices like pay-for-performance schemes, work teams, cross-training, and routinized labor-management communication to productivity growth. In sum, these insider econometrics studies show that factors other than incentive pay are also important determinants of firm productivity.

1.2.3 Interim Conclusion

While both WMS as well as Insider Econometrics have been concerned with the impact of management practices on organizational performance, their focus is quite different: the former concentrates on measuring (aspects of) management quality, showing cross-sectional correlations with productivity (Bloom and Van Reenen, 2007), while the latter focuses on investigating within-firm data suggesting that human resources components of management (HRM) can affect productivity in a complementary fashion (Ichniowski et al., 1997). Despite a common interest, it is important to note that the underlying model of measurement as well as the theoretical rationale differ to some extent: whereas the WMS assumes additive index building with different practices causing a

higher order factor termed management quality, insider econometrics claim that those practices exhibit interaction effects. As Jiang et al. (2012) argue, those effects can be either negative (deadly combinations) or positive (fruitful combinations), which corresponds to the economic notion of complementarities. We argue that albeit WMS has been an influential and important step, from our perspective it is highly unlikely that those practices act virtually independent from one another as assumed by additive indexing. Moreover, we support the claim put forward by SHRM scholars as well as by the researchers advocating the importance of complementarities in organizational economics that those firms which manage to adapt practices that are complementary to (or fit their) external and internal environment are able to establish a sustained competitive advantage and, in turn, PPDs arise.

Still, specific knowledge is fairly limited, last but not least because the traditional case study approach as well as the correlational approach based on WMS data and insider econometrics studies are challenged by identification problems, unobserved heterogeneity and reversed causality being prominent ones (see Becker and Huselid, 2006; Brynjolfsson and Milgrom, 2013, for a detailed discussion). To date, only few studies use approaches like panel data, natural and designed experiments that address this problem (e.g., Athey and Stern, 2002).

In addition, there is still little understanding of the microfoundations of the effect of complementary HRM practices. Hence, for the remainder of this paper, our approach is to focus on two functional areas of HRM and the practices within those that have traditionally been of great interest to economists: incentive design (compensation and benefits) and hiring (recruiting). Using these, we develop a simple model and use it as proof of concept to illustrate that enriching theory with behavioral concepts and sound microfoundations of individual behavior helps to better understand complementarities of HRM practices. In addition, we emphasize the idea of individual heterogeneity between workers and argue that those differences play an important role for how complementarities between practices come to life and persist.

1.3 What Can Behavioral Economics Research Contribute?

After having defined our topic and briefly summarized some evidence in the previous section, we want to develop our argument that behavioral economics research allows us to gain additional insights by highlighting an additional dimension of potential complementarities: employees' heterogeneous preferences interact in non-trivial ways with the design of incentive schemes and the choice of HRM practices more generally. This indicates a case for jointly analyzing recruiting and the organizational choices of firms.

Generally, most of the empirical literature in organizational economics so far does not focus on how recruiting and as a consequence the matching of types might interact with other HRM practices. This omission is reasonable from a neoclassical viewpoint as these factors can to a large degree be separated from optimal incentive design. To illustrate this line of reasoning, think of worker heterogeneity: ability has been a standard dimension of heterogeneity that has been considered in the literature. Every firm wants, ceteris paribus, to hire more able workers. The firm then solves its contracting problem conditional on the characteristics of its workforce. Adding heterogeneity in risk attitudes, however, leads to a somewhat more advanced problem, as the optimal intensity of incentives is now affected by the risk attitudes of the given employee pool (cf. Bandiera et al., 2015b). In both of these cases, ability and risk attitude heterogeneity lead to differences in incentive heterogeneity, but do not affect the structure and nature of organizational choices in a broader sense – while there might result changes in slopes, the general structure of incentives is not changed. Thus, recruiting of workers and incentive or organizational design can be (and have been) treated largely separately.

However, in contrast to that, much research has been conducted in behavioral economics in the course of the last two decades that indicates that this clean separation of effects might not be a correct representation of reality, but that recruiting and matching of types interact with other HRM practices like incentive schemes. For instance, Delfgaauw and Dur (2007) develop a theoretical model assuming workers who are heterogeneous with respect to their intrinsic motivation and show how incentive schemes

can help to attract motivated workers in this setting, and Jones et al. (2014) conduct an experiment and a field study to point out that employer signaling of corporate social performance might be a cause for selection and sorting effects of employees in firms.

Based on these insights, our goal in the remainder of this paper is to highlight how complementary effects of HRM practices and interactions with heterogeneous social preferences matter and how these can be studied. We focus on social preferences within an organizational context, and show how these preferences influence the optimal use of HRM practices by developing a simple, illustrative model introducing reciprocal preferences into the principal-agent framework based on Englmaier and Leider (2012a) in Section 1.3.1. We see this model as a proof of concept that behavioral economic insights can be helpful for explaining facts and patterns in the study of HRM practices and PPDs. Of course, this is only a small piece of what behavioral economics research has to offer and there are numerous other areas that can be fruitfully included in the study of organizations. We provide a discussion on the general mechanism behind it as well as a variety of other particularly fruitful topics in Section 1.3.2.

1.3.1 Social Preferences: an Illustrative Model

Social preferences are a good choice to offer a proof of concept for the role behavioral economic insights can play in studying heterogeneous preferences of agents and how they influence the optimal configuration of HRM practices. Some evidence pointing to this conclusion has been made by scholars studying these preferences.[5] First, there are numerous laboratory experiments indicating that social preferences matter in organizational settings; recently, various researchers have begun to take the question of how reciprocity matters in workplace interactions to the laboratory. Cabrales et al. (2010) find that heterogeneous social preferences, measured in standard lab tasks, are a significant predictor of choices for firms and workers in a multi-stage labor market experiment, and Bartling et al. (2012) directly relate to the idea of complementary sys-

[5] Together with other kinds of non-standard preferences like time preferences (present bias) and risk preferences (reference-dependence), social preferences have been researched most intensely in the field of behavioral economics (DellaVigna, 2009). Fehr and Schmidt (2006) or Camerer and Weber (2013) provide extensive reviews of this line of research.

tems of HRM practices and show that both work attitude and labor market competition are causal determinants of the viability of high-performance work systems in an experimental labor market. Englmaier et al. (2014) investigate one-shot labor relations in the lab and are able to show that employers take into account heterogeneity of workers in the productivity as well as in the social preference dimension. In their setting, firm profits are highest when employing reciprocal workers. Two complementary studies conducted by Englmaier and Leider (2012b,c) explore determinants for the existence of the gift-exchange mechanism both in the laboratory and in the field and show that context matters for the viability of gift exchange: in both settings, the reciprocal reaction of workers is dependent on how much the manager benefits from extra effort.

Second, there are also a few attempts to investigate this relationship in real-world settings using administrative and firm level data, which usually use survey questions to measure reciprocal inclinations. For instance, Leuven et al. (2005) propose reciprocity as explanation for why firms overinvest in general and specific training compared to standard theory predictions and show empirically that training investments are greater when the workforce is reciprocal using the NIPO Post-initial Schooling Survey, a cross-sectional survey with a representative sample of the Dutch population. Dohmen et al. (2009) explore the link between reciprocal preferences and behavioral labor market outcomes and show with data from the German Socioeconomic Panel (G-SOEP) that positive reciprocity is related to high wages, high effort, and life success in general. Huang and Cappelli (2010) demonstrate that screening for "work ethic" is related to the prevalence of less monitoring, more teamwork, less employee turnover, higher wages, and higher firm productivity in US census data. Englmaier et al. (2016) use personality tests in recruitment as a proxy for a reciprocal workforce and show that this is linked to wage generosity, provision of non-pecuniary incentives, team work, and higher firm productivity in general using data from the UK Workplace Employment Relations Study (WERS). However, although all of these studies provide excellent starting points for future research, they all build on cross-sectional data and cannot identify causal effects.

Summing up, there is ample evidence from the lab as well as the field indicating that social preferences matter in organizational settings. Hence, in the remainder of the

section, we develop a simple agency model incorporating reciprocal inclinations on the side of the agent to illustrate how behavioral economics can inform the analysis.

Base Model

Following Englmaier and Leider (2012a), we consider a simple binary principal-agent framework where both the principal and the agent are risk-neutral. There are two states of the world that are characterized by two possible outcomes, q_1 and q_2, with $q_1 < q_2$ and $q_i > 0 \quad \forall i \in \{1, 2\}$. The agent can choose between two actions, a_L and a_H, with related costs $c(a_L) < c(a_H)$. The probabilities of the principal's outcomes are conditional on the action taken by the agent: $Pr(q_2|a_j) = \pi_j \quad \forall j \in \{L, H\}$, where $\pi_H > \pi_L$, which implies that the higher outcome is a better signal for high effort. Hence, the principal's expected return is given by $ER(a_j) = (1 - \pi_j)q_1 + \pi_j q_2 \quad \forall j \in \{L, H\}$.

We assume that effort is not contractible; hence, wages can only be paid conditional on outcomes. Thus, the principal offers the following contract to the agent: $(w, b, \hat{a})$, where w denotes a secure wage payment (a salary) for the agent in every state of the world, b an additional bonus that is paid in case that outcome q_2 is realized and $\hat{a}$ represents a non-binding request for an action.[6] This request reflects the performance level expected by the firm that is known to the worker (e.g., through job descriptions or a code of conduct) and serves to fix the agent's beliefs about the principal's intended generosity; see the discussion below and in Englmaier and Leider (2012a).

The agent has an inherent concern for reciprocity η with $\eta \in [0, +\infty)$. To focus the analysis, we assume here that this type η is publicly observable.[7] His utility from taking

[6] When developing the model in the following subsections, we always discuss optimal wages, as is common in agency models. As the model is formulated in utility terms, however, note that these are to be interpreted broadly and might include multiple ways of transferring utility to employees. A discussion of this is provided in Section 1.3.2.

[7] If we assume instead that firms have access to a noisy signal on η upon hiring, there are two types of errors that might arise: either an in fact suitable, say reciprocal, applicant is mistaken to be non-reciprocal and is not hired, or a non-reciprocal applicant is mistaken to be a reciprocal type and is hired by a firm that has its HR systems geared towards employing reciprocal types. If, because the employment relationship is long-term, the worker continues to mimic a reciprocal type, our conclusions are not affected. If not and the worker starts to take advantage of the low-powered incentive environment, additional measures like having an intense probation period seem particularly important in such reciprocity-based settings. Even if these do not work, as long as the initial signal upon

action a_j when $\hat{a}$ is requested is given by

$$\begin{aligned} U(a_j, \hat{a}) = \; & (1-\pi_j)w + \pi_j(w+b) - c(a_j) \\ & + \; \eta\left[(1-\hat{\pi})\, w + \hat{\pi}(w+b) - c(\hat{a}) - \bar{u}\right] ER(a_j). \end{aligned}$$

As is evident from the utility function, a reciprocal agent ($\eta > 0$) does not only derive utility from the wage payment as is common in agency models (first part of utility function), but also from internalizing the principal's welfare (third part). This part of the function represents the case that reciprocal utility is "triggered" when the proposed wage scheme provides the agent with a rent in excess of his outside option $\bar{u}$. In the simple case of our model, $\bar{u}$ is assumed to be fixed exogenously. For simplification, we also assume that $\bar{u} = 0$, $c(a_L) = 0$, and define $c(a_H) = c$.

Benchmark Case Trivially, when assuming that effort is contractible, the first best solution is implemented by any wage scheme compensating the agent for his effort costs c. Furthermore, without further restrictions, it can easily be shown that the principal can induce a risk-neutral agent to exert high effort for first-best costs, even when effort is not contractible.[8]

The properties of the optimal contract are summarized in the following proposition.

Proposition 1.1. *Under hidden action, the risk-neutral principal proposes the following contract to the risk-neutral agent:*

$$w = -\frac{\pi_L}{\Delta\pi} c < 0, \qquad b = \frac{c}{\Delta\pi}, \qquad \textit{and} \qquad \hat{a} = a_H.$$

In doing so, she can induce the agent to exert a_H at first-best implementation cost $B^{FB}(a_H) = c$, leaving the agent with no rent. This holds for every $\eta \in [0, +\infty)$.

The optimal contract punishes the agent if the low outcome is realized and rewards him if the good outcome is realized. The intuitive explanation for this result is straightfor-

recruiting is precise, the firm presumably has to be able to live with few "rotten apples".

[8] See Appendix A.1 for the exposition of the problem and a derivation of the optimal contract.

ward: since effort is not contractible, the principal has to align the agent's interests to hers. Recall that the agent is risk-neutral; thus, the principal can just structure incentives steeply and transfer risk to the agent without having to compensate him for that. Note that just as in Englmaier and Leider (2012a), reciprocity does not have an effect on the optimal contract in the first best; the "standard contract" prevails.

Limited Liability Now, consider the case where the agent is wealth constrained and thus has to receive a minimal wage in every state of the world that cannot be negative; more formally, consider $w \geq \underline{w} \geq 0$. With an additional limited liability constraint (Innes, 1990) like this, the principal then faces a trade-off between rent extraction and incentive provision when designing the optimal contract.

The properties of the solution to this problem are summarized in the following proposition.[9]

Proposition 1.2. *Under hidden action with limited liability ($w \geq \underline{w} \geq 0$), the risk-neutral principal proposes the following contract to the risk-neutral agent:*

$$w_1 = \underline{w}, \qquad b = \frac{c}{\Delta\pi} - \frac{\eta \Delta ER}{\Delta\pi + \eta\Delta ER\pi_H}\left(\underline{w} + \frac{\pi_L}{\Delta\pi}c\right), \qquad \textit{and} \qquad \hat{a} = a_H.$$

In doing so, she can induce the agent to exert a_H, but only at implementation cost

$$B_\eta^{SB}(a_H) = c + \frac{\Delta\pi}{\Delta\pi + \eta\Delta ER\pi_H}\left(\underline{w} + \frac{\pi_L}{\Delta\pi}c\right).$$

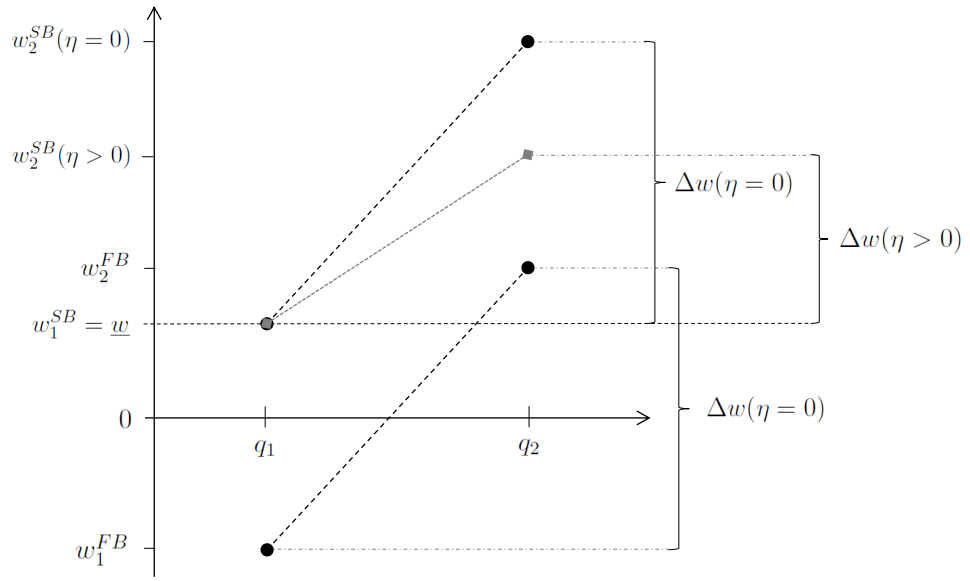
The agent's utility is then

$$U(a_H) = (1 + \eta ER_H)\left[\frac{\Delta\pi}{\Delta\pi + \eta\Delta ER\pi_H}\left(\underline{w} + \frac{\pi_L}{\Delta\pi}c\right)\right].$$

First of all, note that this collapses to the standard case if $\eta = 0$. Hence, the standard case is nested in our model. Comparing the optimal contract for selfish and reciprocal agents reveals that $B_{\eta>0}^{SB}(a_H) < B_{\eta=0}^{SB}(a_H)$; i.e., the principal can implement a_H at lower cost with a reciprocal agent. In both cases, the principal has to pay a rent due to wealth

[9] See Appendix A.1 for the exposition of the problem and a derivation of the optimal contract.

Figure 1.2: *Illustration of the wage schemes*



Notes: The figure depicts the exemplary first best contract (lower dashed line), the second best contract for a selfish agent (parallel dashed line above the first best line) and the second best contract for a reciprocal agent (dotted line).

constraints of the agent; however, while this rent is "lost" on a selfish agent, it serves as an additional incentive for a reciprocal agent. This is also reflected in the bonus payment b, which is lower for reciprocal agents as the (non-zero term comprising of the) limited liability rent, weighted by a term reflecting reciprocity η and effectiveness of reciprocal behavior ΔER, can be subtracted. The intuition for the smaller expected wage bill in this case is straightforward: the wage differential for the reciprocal agent can be smaller because the positive rent provides an additional motivation to the reciprocal agent. More formally, $b_{\eta>0} < b_{\eta=0}$.

The three cases of the exemplary first best contract, the second best contract for a selfish agent and the second best contract for a reciprocal agent in comparison are illustrated in Figure 1.2.

Comparative statics. To see what happens to the wage differential if the principal is faced with a more reciprocal agent, we take the derivative of the optimal b with respect

to η:

$$\frac{\partial b}{\partial \eta} = -\frac{\Delta\pi\Delta ER}{\left(\Delta\pi + \eta\Delta ER\pi_H\right)^2}\left(\underline{w} + \frac{\pi_L}{\Delta\pi}c\right) < 0.$$

Hence, the principal can offer a lower b to a more reciprocal agent, which is a rather intuitive result.[10] Subsequently, the expected wage bill for employing the agent also decreases in η.[11] This is summarized in the following proposition.

Proposition 1.3. *b is decreasing in η and ΔER. As a consequence, the expected wage bill $B_\eta^{SB}(a_H)$ is also decreasing in these factors.*

Interim conclusions. To sum up, in a hidden action principal-agent framework with limited liability, we can show that the principal is able to reduce the incentive intensity and the wage bill when contracting with a more reciprocal agent.

Note that even in this simple model, it is possible to immediately see that there are different strategies for firms to combine complementary HRM practices in hiring and incentive design: if a firm manages to attract a reciprocal workforce, it can implement several reciprocity-based practices (e.g., little explicit incentives, generous treatment, job security, little monitoring) while if it attracts non-reciprocal workers, the system optimally geared to this looks starkly different (e.g., internal competition, steep incentives, close performance monitoring). While these two firms, due to the use of complementary practices, look starkly different (some might say one appears well managed and the other badly managed), they might be relatively comparable in terms of profitability.

[10] Also, since the increase in the principal's expected revenue due to choosing a_H rather than a_L, ΔER, is always multiplied with η in b, the same holds true if the value of effort increases for the principal.

[11] Considering the limit case $\eta \to \infty$, we can show that for an infinitely reciprocal agent, the principal could even induce high effort a_H for first best cost:

$$\lim_{\eta\to\infty} B_\eta^{SB}(a_H) = \lim_{\eta\to\infty}\left[c + \frac{\Delta\pi}{\Delta\pi + \eta\Delta ER\pi_H}\left(\underline{w} + \frac{\pi_L}{\Delta\pi}c\right)\right] = c$$

which equals $B^{FB}(a_H)$.

Other HRM Practices in the Model Framework

In their seminal textbook, Milgrom and Roberts state that "[...] important features of many organizations can be best understood in terms of deliberate attempts to change the preferences of individual participants [...]. As a result, organizationally desired behavior becomes more likely." They continue to argue that this "is clearly an element of leadership [...] and it has much to do with practices of organizing semi-permanent groups of workers and encouraging them to interact socially as well as at work." (Milgrom and Roberts, 1992, p.42). While Milgrom and Roberts then continue down the neo-classical route, positing a selfish, rational actor framework, with the progress made by behavioral economic research we can begin to more carefully study the content of their statement. In this spirit, we include the possibility for HRM practices to influence reciprocal behavior in our model by modifying the reciprocity parameter η.

Without tapping into the vast nature versus nurture debate in detail, it can be stated that the assumption of stability of preferences in economics and of personality traits in psychology has been heavily challenged in recent years (Woods et al., 2013). Albeit the trait approach of personality has been the prevalent paradigm for the last eighty years, it is being criticized by scholars putting characteristics of the situation at center stage (e.g., Mischel and Shoda, 1995). In economics, the assumption of stable preferences has been scrutinized as well (see, e.g., Bowles, 1998). In light of all those issues, we argue that preferences that matter most in the work context may at least to a degree be subject to change by HRM practices in a firm. For instance, specific aspects of work design, like fostering intense team work, may affect preferences, in particular, strengthen pre-existing social preferences.

When conceptualizing how one could incorporate HRM practices in the relatively general model framework outlined above, one can think of several different ways to do so, each with different theoretical and empirical implications. Here, we focus on ways of influencing the employee's reciprocal motivation, η, by HRM practices like team building, empowerment, or delegation of decision rights. In a similar vein, Pfeffer (2007) stresses that there are certain types of high-commitment HRM practices that serve to

activate reciprocal motivation. These include investments in training, information sharing, or decentralized decision making as signals of trust. As a consequence, we have to think of η as consisting of two parts – the agent's inherent concern for reciprocity, η_1, as well as a second malleable part that can be influenced by HRM, η_2. Depending on how these two factors interact, different predictions can be derived from the model.

To fix ideas, consider a multiplicative model: $\eta = \eta_1 \times \eta_2$.[12] For simplification, we assume that the agents' preference types are known. We substitute η in Proposition 1.2 and derive comparative statics by taking the derivative of the optimal bonus payment:

$$\frac{\partial b}{\partial \eta_2} = -\frac{\eta_1 \Delta\pi \Delta ER}{\left(\Delta\pi + \eta_1\eta_2 \Delta ER \pi_H\right)^2} \left(\underline{w} + \frac{\pi_L}{\Delta\pi} c\right) < 0.$$

Here, the interplay of the two reciprocal factors is complex, which can be seen when considering the second and cross-derivative:

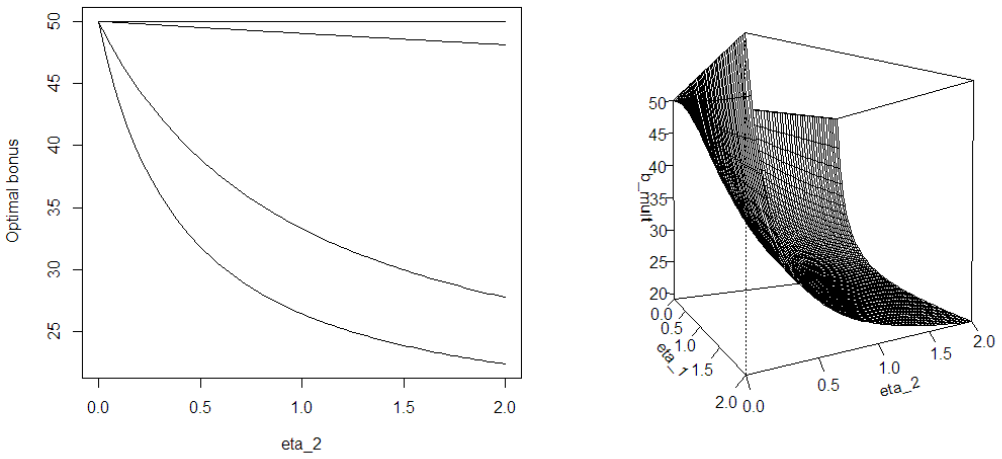
$$\frac{\partial^2 b}{\partial \eta_2^2} = \frac{2\eta_1^2 \Delta\pi \left(\Delta ER\right)^2 \pi_H}{\left(\Delta\pi + \eta_1\eta_2 \Delta ER \pi_H\right)^3} \left(\underline{w} + \frac{\pi_L}{\Delta\pi} c\right) > 0,$$

$$\frac{\partial^2 b}{\partial \eta_1 \eta_2} = -\frac{\Delta\pi \Delta ER \left(\Delta\pi - \eta_1\eta_2 \Delta ER \pi_H\right)}{\left(\Delta\pi + \eta_1\eta_2 \Delta ER \pi_H\right)^3} \left(\underline{w} + \frac{\pi_L}{\Delta\pi} c\right) \lesseqgtr 0.$$

While the former is globally positive, the latter is only positive for a threshold value of $\eta_1\eta_2 > \Delta\pi / \Delta ER \pi_H$, i.e., for a sufficient amount of reciprocal motivation already present. This hints at a separation of cases – below the threshold, investments in η_2 will hardly reduce the bonus payment necessary and are thus not profitable for the firm, while investment for levels above the threshold pay off. To illustrate this relationship, consider the situation that the firm has the opportunity to implement a training program that is known to increase η_2 from 1 to 1.5, but that has implementation costs of $ 1,000. Note that the investment decision cannot be made without considering the level of η_1: when it is small, the optimal bonus payment will only decrease by a small amount that does not exceed $ 1,000. For a large enough η_1, however, the increase of η_2 by

[12] See Appendix A.2 for an additive modeling, $\eta = \eta_1 + \eta_2$, and the discussion of its implications.

Figure 1.3: *Optimal bonus dependent on η in multiplicative model*



Notes: The figure depicts the optimal bonus as function of η_1 and η_2, the parameters capturing the innate reciprocal inclination of a worker and the HRM policies targeted at increasing this reciprocal inclination towards the firm, for a multiplicative formulation $\eta = \eta_1 \times \eta_2$.

0.5 will result in a bonus decrease that exceeds $ 1,000 – in the former case, the firm optimally does not invest, while in the latter case, it does, although the effectiveness of the training itself has not changed. In this way, a higher level of η_1 increases marginal returns from investing in η_2, which reflects the complementarity definition given above. This relationship is also depicted graphically in Figure 1.3.

It is evident that the firm can foster reciprocal reactions of their employees by investing into practices that increase η_2 as long as the employee has an inherent concern for reciprocity at all ($\eta_1 > 0$). Again, it is important to note that the strength of this influence and hence the decrease of the necessary bonus payment depend on the level of the agent's inherent concern for reciprocity. In other words, η_1 and η_2 are complements and thus interdependent.

Implications. Modeling employees' preferences and preference enhancing HRM policies as complements makes thinking about these investments more complicated (and interesting). When the ex ante level of overall reciprocal motivation is low, either because the workforce selected has a low concern for reciprocity or because there are no other investments in the form of HRM practices, additional investment does not lead

to significant gains in terms of lower bonus payments for the firm and is hence hardly profitable. In contrast, when ex ante reciprocal motivation is already above a certain threshold, gains from increasing investment in either dimension are larger. Phrased differently, as η_1 is related to hiring practices used in personnel selection, these practices are complementary to practices that directly increase η_2. As an application, one could think of using a selection process focused on screening for reciprocal agents and utilizing work design to influence reciprocity, for instance, by granting high discretion (cf. Pfeffer, 2007). This is then optimally complemented by a flat, generous incentive scheme.

Conclusions from the Model

To conclude and sum up the insights from modeling a simple contracting situation with reciprocally motivated agents, we have seen that the composition of the workforce and the effect of HRM practices are interdependent. In this respect, the design of incentive schemes and the effects from selection and matching cannot be treated separately. This in turn suggests that those practices are indeed complements yielding potential convex returns in terms of organizational performance and ultimately explanatory power for disaggregating PPDs.

Before going on by broadening our focus, we now discuss in what way this model (and behavioral contracting models in general) can yield testable predictions and thus help exploring complementarities of HRM practices. Of course, by incorporating individual preferences, the need to observe both these and individual-level outcomes is evident, something which is hard to observe in field settings.[13] However, in our opinion, this is an excellent example for a case where there is a clear role for laboratory and field experiments that make it possible to test specific mechanisms in a clean manner and to measure individual preferences using behavioral outcomes, which matters especially when individual heterogeneity comes into play (see Cohn et al., 2015, for an argument

[13] However, economic researchers are beginning to incorporate measures for individual preferences like reciprocity, risk attitudes or patience in large administrative data sets like the G-SOEP (see, e.g., Dohmen et al., 2009, 2011), which makes it possible to investigate related questions using field data.

along these lines). This then, in the spirit of a feedback loop, can inform larger-scale empirical approaches using administrative data.

1.3.2 Other Aspects

As already stated right at the beginning of this section and hinted at in the last paragraphs, we use social preferences of workers as an illustrative example to support our case that behavioral economics has much to contribute when it comes to complementarities of HRM practices. However, of course social preferences are not the only behavioral aspect of potential relevance in this field. Possible topics for future research incorporating complementarities of HRM practices will involve further areas of behavioral economics research. Besides social preferences, which have probably been studied most extensively, there are also various other concepts yielding insights that are likely to matter in this setting. We aim at demonstrating this point by first highlighting that the general mechanisms we demonstrated cannot only be found for social preferences, but rather appear to be more general, at least when utilizing a broader interpretation of the model, and then giving a brief overview of research in other areas in behavioral economics that we perceive as relevant.

Broader Interpretations of the Model

In the following, we consider some ways in which our framework can be thought of in a more general way, indicating that we are hinting at a more general pattern. There are two aspects we want to stress: first, our results can be compared to arguments from identity economics (Akerlof and Kranton, 2000, 2005). Second, as already noted above, despite the fact that the standard agency model incorporates monetary benefits and we also frame our model in this terms, incentives do not have to be defined as being purely monetary.

Identity. Although we define the non-standard motivation of agents in our model as the agent being reciprocally motivated, many of our arguments are parallel to arguments of the identity literature. For instance, Akerlof and Kranton (2005) also explore the role of

identity as supplement to monetary compensation by incorporating identity utility that depends on the deviation from norms and ideals. They derive a result that is similar to ours: being an insider leads to a reduced wage differential. When relating their model to the workplace, the authors explicitly define that the dichotomy of insiders and outsiders can be thought of as the dichotomy of extrinsic and intrinsic motivation in management. Here, modern approaches like investments in corporate culture aim at aligning the goals of workers with those of the firm, essentially trying to shift workers from being outsiders to being insiders.

In a more general sense, our formulation – based on Englmaier and Leider (2012a) – like the identity formulation of Akerlof and Kranton (2005), the model incorporating trust by Sliwka (2007), or the theoretical formulations incorporating motivated agents (e.g., Besley and Ghatak, 2005; Cassar, 2014), can all be thought of as the utility representation of workers internalizing the firm's goals. In this respect, the aspect of motivation that is influenced by HRM practices, η_2 in our model, can be reinterpreted as investment into category affiliation, or, exactly like in the identity framework, as investments into goal alignment. The only differences between the models are then the way of thinking about the underlying mechanism and the underlying inherent motivation of the agents. Again, differentiating between the explanations and carving out the underlying mechanism is ultimately an empirical task with a clear-cut role for lab and field experimental investigation.

Benefits. As already briefly noted at the beginning of Section 1.3, narrowly interpreted, the model (as well as the standard agency model) describes a situation in which a monetary payment, the wage, is exchanged for effort exertion by the agent. However, interpreting the model in a broader sense is possible. All the formulations are essentially set up in utility terms; hence, one could also think about the bonus as any benefit for the agent generating a utility rent and about effort exertion as everything that is beneficial for the company, which includes, for instance, retention of the employee.

Again, this way of generalizing the model has numerous empirical implications. Naturally, thinking about the problem in this generalized way means that it loses structure,

but also again highlights the role of experiments to identify which benefits can substitute for money and which actions of the agent can substitute for effort – and, ultimately, to clarify the intricate relationship between different types of incentives and benefits (cf. Oyer, 2008).

Other Areas of Study from Behavioral Economics of Particular Relevance

Besides the aspects of social preferences discussed above, there are also other interesting areas within the field of behavioral economics that might prove useful for thinking about HRM practices and possible complementarities.

First, there is an increasing number of papers focusing on different types of non-monetary incentives in work relations. This includes research on a variety of different aspects, for example, on the broad area of "respect" (Ellingsen and Johannesson, 2007). Here, some authors have focused on exploring the role of spontaneous recognition for employee performance (e.g., Chen et al., 2010; Bradler et al., 2013; Hoogveld and Zubanov, 2014), which is proposed to work through conformity to a performance norm (Bernheim, 1994). Similarly, status incentives in general, i.e., non-pecuniary rewards for good performance like medals, prizes, awards, or job titles, have gained some interest in recent research and have been examined both theoretically (e.g., Besley and Ghatak, 2008; Auriol and Renault, 2008) and empirically (e.g., Ashraf et al., 2014; Charness et al., 2014; Kosfeld and Neckermann, 2011; Gubler et al., 2013). Proposed mechanisms for these kinds of incentives to work are that their value is based on scarcity and the human tendency to crave for social status and recognition (Besley and Ghatak, 2008). Related to that and the literature on tournaments, several researchers have also been exploring the role of relative rank as a motivator (e.g., Azmat and Iriberri, 2010; Bandiera et al., 2015a; Barankay, 2011a,b; Blanes i Vidal and Nossol, 2011). Other topics include job characteristics like perceived meaningfulness of tasks (e.g., Ariely et al., 2008; Grant, 2008) or (lack of) discretion (Falk and Kosfeld, 2006; Dickinson and Villeval, 2008).

Summing up, all of these studies find that there are several other ways for principals to motivate their workers than pure incentive pay. These "new" HRM tools are of course

very interesting in themselves, but so far, this area of research is mostly comprised of studies that look at one tool in isolation without incorporating a whole system of practices. As with evidence on social preferences, the scope for exploring complementarities between those practices then is very limited. From our perspective, more comprehensive studies incorporating whole systems of HRM practices are needed to address possible complementarities.[14]

In addition, some researchers have also begun to look into personality traits – mostly the "Big Five" (agreeableness, conscientiousness, neuroticism, extroversion, and openness) – and find that these also interact with outcome dimensions like inherent motivation and productivity (e.g., Segal, 2012). As with heterogeneous social preferences, firms' optimal systems of HRM practices are likely to depend crucially on workforce composition with respect to these traits.

Furthermore, within behavioral economics in general, there are a few other topics that we think of as promising. There is an important body of research on time inconsistent preferences of agents that lead to self-control problems and procrastination (O'Donoghue and Rabin, 1999a,b). Here, one major insight is that sophisticates, i.e., those agents who know about their time inconsistency and seek to avoid it, might actively search for commitment devices, and firms that know about the problem of time inconsistency might in turn desire to employ sophisticates. As a consequence, offering commitment devices might be one strategy for achieving this. Examples include up or out schemes, conditional training or rotation schemes, or gym memberships[15], which can all be interpreted in this way (see Englmaier et al., 2015, for an attempt to capture this idea theoretically within a long-term employment contracting framework). In addition, non-standard beliefs like overoptimism and overconfidence might also matter when thinking about interrelations between HRM practices (see, e.g., Larkin and Leider,

[14] Recent exceptions include Bartling et al. (2012) and Kosfeld et al. (2014).

[15] This might not seem comparable to the aforementioned practices at first glance. However, gym membership and attendance have been in the focus of studying time inconsistent behavior in the field. At the same time, gym memberships are a prime example for non-wage benefits employers are offering with increasing frequency and attract specific types of workers (cf. Lazear and Shaw, 2007). See Oyer (2008) for a discussion of factors that influence firms' benefit choices.

2012; Sautmann, 2013). However, note that in this areas, as well as in other sub-fields of behavioral economics research, insights almost exclusively stem from experiments conducted in laboratory settings, with few firm data available so far. Here, perhaps even more than in the area of social preferences, it is crucial to push for comprehensive multi-level field data to understand the prevalence and patterns of practices in the field.

1.3.3 Interim Conclusion

Throughout this section, we have made the point that enriching theoretical models with behavioral aspects has several important implications for the study of complementary HRM practices.

Note that, in general, adding agents with heterogeneous, non-standard preferences to the problem will make the contracting environment become more complicated because all of these preferences are relatively difficult to measure, while firms' optimal strategies crucially depend on the composition of the workforce. Hence, HRM practices that decrease those information asymmetries and in turn increase matching quality are also a potential source of complementary HRM practices. To achieve this, firms can utilize various methods – they can screen workers by offering menus of contracts and let them self-select into occupations, or carefully test and screen applicants. Another method would be to just ask current employees to suggest fitting applicants, or, in other words, request referrals. Recently, this mechanism has been receiving increased attention by several field experiments focusing either on different types of incentives for providing referrals (Beaman and Magruder, 2012), performance differences between referred vs. unreferred employees (Pallais and Sands, 2016) or other organizational relevant outcomes such as turnover (Burks et al., 2015). Summing up, these studies emphasize that efficient matching of types matters and in turn may be a reason for the use of referrals. However, the aforementioned studies focus solely on referrals and therefore are silent about other HRM practices that firms employ. In contrast to that, our model predictions rely on the interactions between practices, and complementarities seem to be important for productivity outcomes. Consequentially, more research that explicitly refers to the complementarities between HRM practices and the role of screening in general and

referral hiring in particular is needed.

1.4 Outlook

For answering the more general question how the potential explanatory power of HRM practices can best be studied, we suggest combining different methods and angles. Within this general agenda, WMS-style cross-sectional studies are important as they map correlations and identify effective clusters of practices; to do so, in light of the above suggested role for behavioral insights, the scope of these cross-sectional studies has to be broadened to capture a wide set of practices employing a measurement model to assess potential complementarities. In parallel, insider-econometric style studies help to provide in-depth insight into what matters in specific firms. In addition to these approaches, we argue that, in a first step, it is important to develop solid theoretical models yielding testable predictions, in particular including behavioral constructs. In a second step, these predictions are tested using both laboratory and field experiments to establish causality for theoretically suggested mechanisms.

One example for how we envision research on organizations, management practices and interactions to progress is the Organizations Research Group (ORG)[16], an initiative at the University of Munich with the goal to explicitly study these topics by combining theory, lab and field experiments, and large-scale survey data. Hence, an important feature of ORG is the Organizational Observatory (O^2), a large-scale survey that aims at collecting rich primary data on organizations, their structure and design, with an emphasis on management practices that have not been studied extensively in the WMS, while striving for comparable data quality by utilizing the same survey methodology. Already started and being continuously improved, O^2 is ultimately planned to have a panel structure and to combine firm-level data on practices with employee-level data.[17] Beyond generating survey data, ORG also has the aim to serve as hub for field exper-

[16] http://www.organizations-research-group.uni-muenchen.de/index.html

[17] Related to this, we want to stress that there are some other groups of researchers that try to investigate the same kinds of questions by producing high-quality, detailed data; one example for this is the Linked Personnel Panel (LPP) that combines firm- and employee-level data (see Kampkötter and Sliwka, 2016, in this volume).

iments with and in firms. In combination with method-based training for graduate students through structured graduate programs like the “International Graduate Program Evidence-Based Economics"[18], ORG provides expertise for interested firms and organizations to conduct randomized interventions. Thus, we hope for ORG to be an integral part of the integrated research on organizations that we envision in the future.

Summing up, by investigating the role of HRM practices for explaining persistent productivity differences and exploring the concept of complementarities between them, we have shown that integrating different research strands and employing different methods helps to gain new insights and to understand the field more thoroughly. Especially for the role of complementarities of HRM practices, we have demonstrated how linking seemingly unrelated results from behavioral economics research on social preferences can be fruitful for further research in this area. However, by giving an overview of different issues and upcoming topics like determinants of productivity and referral hiring, we have also highlighted that more research, and especially more comprehensive firm data, are needed.

To conclude, we briefly sketch the virtuous feedback cycle we envision to evolve between differing research methodologies. Consider the following example: cross-sectional studies like the WMS, the LPP, or O^2 provide the possibility to relate within-firm patterns to market-level characteristics like market structure or the intensity of competition. So far, models on behavioral aspects have been largely silent on this topic that is likely to have important effects on internal organization choices, which is why we hope that empirical findings from these surveys will stipulate theoretical research. Again, this theorizing will generate new predictions that have to be tested empirically. Here, experimental studies have the role to causally establish hypothesized mechanisms, while insider econometrics studies and new modules in surveys can again help to identify practices and complementarities among them.

[18] http://evidence-based-economics.de/home.html

Chapter 2

What Drives Reciprocal Behavior? The Optimal Provision of Incentives over the Course of Careers*

2.1 Motivation

Humans reciprocate. They repay kindness with kindness and hostility with hostility. Several possible explanations exist for why individuals display reciprocal behavior, the most prominent ones being inherent preferences for reciprocity and repeated interaction (see Sobel, 2005, for an overview). Whereas inherent preferences for reciprocity reflect the idea that an individual can enjoy additional utility when returning favors he or she has received (based on the gift exchange concept, Akerlof, 1982), repeated interaction can give rise to reciprocal behavior even if individuals only care about their own material payoffs. A vast amount of evidence supports both drivers of reciprocal behavior, however mostly trying to isolate one from the other.

In this paper, we address the questions whether inherent preferences for reciprocity are also relevant in long-term employment relationships, and if and how they affect relational contracts, where repeated interaction generates incentives. We show that both kinds of incentives do interact with each other in an optimal incentive system, and that

* This chapter is based on joint work with Matthias Fahn and Anne Schade.

their relative importance depends on the career phase: At early stages, incentives generated by repeated interaction are more important because more future rents can be used to provide incentives. At later stages, reciprocity-based incentives become more and more important and gradually replace repeated-game incentives. However, reciprocal preferences are still important early on. First, they reduce an employee's effective effort costs. Second, they relax the employer's enforceability constraint which determines its commitment in the relational contract. Therefore, reciprocity-based and repeated-game incentives are dynamic *substitutes*, but *complements* at any given point in time.

After deriving these – and other – results within a theoretical model, we test its implications using data from the German Socio-Economic Panel (SOEP). As predicted, we find evidence for a positive effect of reciprocity on effort, and that this effect is stronger for older employees. These results indicate that reciprocity-based as well as repeated-game incentives interact in real-world incentive systems.

More precisely, we develop a dynamic principal-agent model with a finite time horizon. Effort is observable but not verifiable, and yields a verifiable output measure. Standard spot contracts based on output are feasible, but necessarily associated with a rent going to the agent. Furthermore, the agent reacts reciprocally towards any *voluntary* rent, i.e., any unconditional wage payment. We first show that static spot contracts either take the form of a standard "bonus contract" (with a wage of zero) or – if the agent's preferences for reciprocity are sufficiently strong – of a "reciprocity contract" (with no bonuses). Next, we take into account that repeated-game incentives based on effort can also be provided by using a so-called relational contract. Here, the principal promises a bonus based on exertion of the desired effort level. Because effort is not verifiable, the principal's promise must be credible, which is the case if paying the promised effort-based bonus triggers sufficiently higher continuation profits than refusing to do so. In our case, this can be achieved despite a finite time horizon, because we assume that once promises made in the relational contract are reneged on, the agent's preferences for reciprocity towards the principal disappear. Therefore, the principal can be punished for reneging if a reciprocity contract is optimal in a spot relationship – because upon reneging, she only has the option to use (less profitable) bonus spot contracts.

Generally, the enforceability of effort in the relational contract is determined by a so-called dynamic enforcement constraint, which states that the effort-based bonus must not exceed the difference between future discounted profits on and off the equilibrium path. This yields a first source of complementarity between relational and reciprocity contracts because the principal has more to lose from reneging if the difference between profits generated by a reciprocity contract and the profits generated by a bonus contract in the last period is larger. Therefore, the relational contract can implement higher effort if the agent's preferences for reciprocity are more pronounced. Moreover, there exist two additional channels how the agent's reciprocal inclinations amplify the performance of the relational contract. First, receiving an extra rent effectively reduces the agent's effort costs. Therefore, it is optimal for the principal to always pay a fixed wage. Second, a binding dynamic enforcement constraint is relaxed and more effort can consequently be implemented with a higher fixed wage. All this implies that incentives triggered by reciprocal preferences and relational contracts are *complements* at any given point in time. However, they are dynamic *substitutes* in a sense that – as time proceeds – repeated-game incentives which are utilized by the relational contract are gradually replaced by reciprocal incentives. This is because the dynamic enforcement constraint is automatically tightened as time moves on (less remaining periods reduce the difference between the principal's on- and off-path continuation profits), and a tighter constraint amplifies the benefits of reciprocity-based incentives.

The optimal incentive scheme has implemented effort at its highest level in early stages of the employment relationship, where it remains until the dynamic enforcement constraint starts to bind. Then, the principal's reduced credibility effectively constrains her ability to pay a sufficiently high effort-based bonus. This decreases implementable effort, which in turn lets the principal respond with an increase of the fixed wage in order to mitigate the necessary effort reduction. Overall, however, a binding dynamic enforcement constraint reduces equilibrium effort because implementing an additional unit is then more expensive with reciprocity-based incentives than with relational incentives. Therefore, although these two are dynamic substitutes, the substitution is incomplete. Concluding, once the dynamic enforcement constraint starts binding, effort gradually

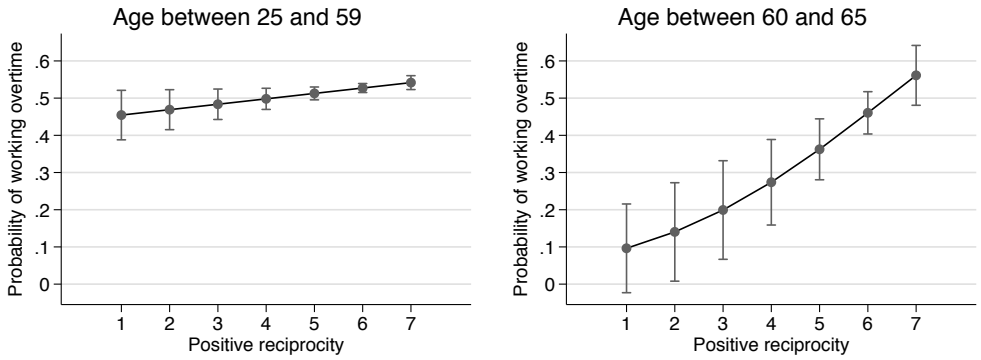
decreases and reaches its lowest level in the last period of the game. The effort reduction goes hand in hand with a gradual increase of the fixed wage.

We explore the empirical validity of our theoretical results using representative survey data on ∼8,000 employees from the German Socio-Economic Panel (SOEP). We utilize the fact that in the 2005 wave of the survey, the SOEP included measures of intrinsic reciprocity. Our approach follows Dohmen et al. (2009), who use overtime as a proxy for non-contractible effort and show that reciprocal inclinations are linked to high effort, high wages, and general life success. Their results hence support the notion that reciprocal preferences help to enforce effort. However, because the SOEP does not contain information on actual incentives systems, Dohmen et al. (2009) do not explore the extent to which reciprocal preferences are optimally utilized in a firm's incentive system – in particular in interaction with repeated-game incentives. Our model, though, allows to draw conclusions from observable outcomes on actually used incentive systems, because we derive that reciprocal inclinations and repeated interaction take on different roles in different stages of a career. More precisely, we develop the following predictions in Section 2.3: First, we predict that an individual with stronger reciprocal preferences is expected to exert more effort, and, second, that effort is decreasing over time. Third, we predict that the positive effect of reciprocal preferences on effort becomes stronger over the course of an employment relationship. This prediction directly follows from reciprocal and relational incentives being dynamic substitutes: At later stages of a career, the incentive system puts more weight on reciprocal incentives, in particular for individuals with more pronounced preferences for intrinsic reciprocity. Therefore, equilibrium effort responds more strongly to reciprocal preferences later on.

Furthermore, we use the agent's realized utility levels to derive empirical predictions. The SOEP contains a measure on an individual's job satisfaction, which we argue is a good proxy for utility experienced in the employment relationship. Our model predicts that more reciprocal individuals enjoy higher levels of job satisfaction, that it increases over time, and in particular that the positive effect of reciprocal preferences on satisfaction becomes stronger over the course of an employment relationship.

Our empirical analysis, conducted in Section 2.4, largely confirms our model's predic-

Figure 2.1: *Predictive marginal effects of positive reciprocity*



Notes: The figure plots predicted marginal changes of the propensity to work overtime at different levels of positive reciprocity depending on age group (left panel) and retirement propensity (right panel), holding all other factors constant. Error bars indicate 95% confidence intervals.

tions. More specifically, we are able to show that while positive reciprocity generally has a positive effect on the propensity to work overtime – a result that corresponds to those reported by Dohmen et al. (2009) –, this effect is much more pronounced for older workers and workers who indicate that they are close to retirement. For example, when including an interaction term between reciprocity and a dummy indicating that an employee is at least 60 years old, we find that the positive effect of positive reciprocity is much more pronounced for workers above the age cutoff. This is illustrated in Figure 2.1, which depicts the predicted marginal effects of positive reciprocity on the propensity to work overtime as proxy for effort in the subsamples.

Furthermore, when estimating the effect of reciprocity on job satisfaction, our predictions are largely confirmed; in particular, the effect of positive reciprocity on job satisfaction is larger for individuals who are close to retirement.

In Section 2.4.2, we explore alternative specifications. First, we show that our results are robust to different specifications of the propensity to work overtime, in particular if only unpaid overtime is considered (our main specifications follows Dohmen et al. (2009) and includes all forms of overtime). Second, we explore the intensive margin of effort by using overtime hours instead of only a binary question whether individuals have worked overtime or not. Finally, we also use data from 2010 and 2015, where

reciprocity was again included in the SOEP. There, we still observe the expected patterns and can thus exclude that those are mainly driven by cohort effects.

Finally, note that a specific age cutoff (we use 60 in our main specification) is not crucial for our results. We present results for a large number of specifications (for different age cutoffs, overtime measures, or included survey waves). Whereas significance levels differ among the specifications, all of them indicate that preferences for reciprocity assume a larger role in later stages of a career.

Before introducing our theoretical model, we now give an overview of literature relevant for our research question.

Related Literature

The deviation from the assumptions of self-interest and greed is one of the most robust, thoroughly researched, fundamentals in the field of behavioral economics (DellaVigna, 2009). There, inference on intrinsic reciprocity is based on Akerlof's conceptual idea of gift exchange, i.e., that employees exert voluntary effort if they feel well treated by firms (Akerlof, 1982). Seminal work by Fehr et al. (1993, 1998) attempts at testing the gift-exchange paradigm experimentally and has inspired a plethora of research that establishes the prevalence of the norm of reciprocity (see, e.g., Camerer and Weber, 2013, for an overview of existing experimental research). This is important for organizations because the existence of reciprocal individuals has the potential to influence the employment relationship in fundamental ways. However, employment relationships are inherently dynamic, and most of the approaches identifying reciprocal preferences have been careful in muting all incentives potentially stemming from repeated interaction. Some recent experimental studies have started to address this issue by disentangling strategic (i.e., generated by repeated interaction) and intrinsic motives for cooperation. Reuben and Suetens (2012) use an infinitely-repeated prisoner's dilemma to assess the relative importance of strategic motives and intrinsic reciprocity and find that cooperation is mostly driven by strategic concerns. Similarly, Dreber et al. (2014) find that strategic motives seem to be more important than social preferences in an infinitely repeated prisoner's dilemma. Cabral et al. (2014) conduct an infinitely repeated veto

game to distinguish between different explanations for generous behavior. They find strategic motives to be the predominant motivation, but present evidence for the importance of intrinsic reciprocity as well.

Hence, experimental evidence suggests that repeated-game incentives are not only relevant in situations of repeated interaction, but rather seem to be the dominant mode to support cooperation. However, to understand how cooperation is achieved in long-term employment relationships, and in particular if and how incentive systems respond to the existence of reciprocal preferences, real-world evidence is needed. As described above, Dohmen et al. (2009) use data on individual-level survey measures for reciprocity from the German Socio-Economic Panel (SOEP), and show that reciprocal inclinations are linked to high effort, high wages, and general life success. Moreover, based on a double moral-hazard problem that can be overcome with promotion incentives for reciprocal agents, Dur et al. (2010) follow Dohmen et al. (2009) and use data from the German Socio-Economic Panel (SOEP) to show that reciprocal preferences are linked to performance appraisals, which serve as a proxy for promotion incentives. Furthermore, existing papers have linked firm-level proxies for reciprocity, like screening for work ethic or personality, to management practices and outcomes such as monitoring, teamwork, wage levels, and firm productivity. These papers provide at least suggestive evidence for the importance of reciprocity in employment relationships (Huang and Cappelli, 2010; Englmaier et al., 2016). Conducting field experiments, Bellemare and Shearer (2009, 2011) show that monetary gifts increase effort in a real-world working environment.

The theoretical literature on intrinsic reciprocity can be arranged along the lines whether reciprocal behavior is triggered by intentions or by outcomes, i.e., whether one counterpart's preferences for reciprocity can be used strategically. In the already-mentioned work by Akerlof (1982) – probably the first to formally model the idea of intrinsic reciprocity – employees are willing to exert additional effort if they are paid more than the market-clearing wage. Hence, firms can strategically raise wages in order to induce their employees to work harder. Applying this idea to a moral hazard framework, Englmaier and Leider (2012a) show that generous compensation can be a substitute for performance-based pay. On the other hand, Rabin (1993) argues that the perceived

kindness of an action should be the driving force to induce reciprocal behavior. He develops the techniques for incorporating intentions into game theory. Dufwenberg and Kirchsteiger (2004) apply these techniques to extensive games and explicitly account for the sequential structure of the respective games. Netzer and Schmutzler (2014) show that if only intentions matter, a self-interested firm cannot benefit from its employees' reciprocal preferences.

Whereas these two approaches assume that either only outcomes or only intentions are relevant, there is vast evidence that both aspects matter. Gneezy et al. (2000), Charness and Rabin (2002), or Falk et al. (2008) (among many others) present results that can be explained only if both aspects, fairness intentions as well as preferences for the distribution of outcomes, are taken into account. Falk and Fischbacher (2006a) develop a theory incorporating both aspects. They assume that an action is perceived as kind if the opponent has the option to treat someone less kind. Hence, intentions matter, however reciprocity can also be used strategically.

We build upon these ideas and assume that reciprocity is triggered by a generous wage and hence can be used strategically. However, intentions matter as well because only non-contingent payments matter, and because the agent's inclination to reciprocate disappears once the principal has broken any implicit promise made in the past.

We also contribute to the literature on relational contracts – self-enforcing, dynamic agreements based on non-verifiable information. Bull (1987) and MacLeod and Malcomson (1989) derive relational contracts with observable effort, whereas Levin (2003) shows that those also take a rather simple form in the presence of asymmetric information with respect to effort and the agent's characteristics. Malcomson (2013) delivers an extensive overview of relational contracts. Within this broader area, few papers have started to investigate how relational contracts and social preferences interact: Dur and Tichem (2015) incorporate social preferences into a model of relational contracts. They show that altruism undermines the credibility of termination threats which may reduce productivity and utilities. Contreras and Zanarone (2017) assume that employees suffer when their formal wage is below that of their colleagues. They show that these "social comparison costs" can be managed by having a homogeneous formal gover-

nance structure, while achieving necessary customizations through relational contracts. To the best of our best knowledge, we are the first to incorporate intrinsic preferences for reciprocity into a relational contracting framework. This allows us to derive specific predictions with respect to the interaction and relative importance of repeated-game incentives and reciprocity in an optimal incentive scheme. In light of the somewhat conflicting evidence on the interaction of the two mechanisms, we show that both are dynamic substitutes, but complements at any given point in time.

On a more general note, various papers have been investigating how agents with standard preferences respond to the (potential) existence of reciprocal agents. Kreps et al. (1982) form the basis for the notion that repeated-game incentives amplify intrinsic reciprocity (for an overview, see Mailath and Samuelson, 2006). The authors show that uncertainty about the presence of reciprocal types is enough for selfish types to rationally imitate reciprocal behavior in a finitely repeated prisoner's dilemma game. Andreoni and Miller (1993) and, utilizing gift-exchange games, Gächter and Falk (2002) present experimental evidence that is in line with this conception. Fehr et al. (2009a) also make a case for the complementary effects of reciprocal preferences and reputation. They state that cooperation is usually way more pronounced in repeated than in one-shot interaction and claim that this is due to selfish types imitating fair types.

We complement these arguments by showing that the positive effect of a long-run interaction on cooperation does not have to rely on signaling, but can also be generated by the optimal incentive scheme designed for individuals with reciprocal preferences.

2.2 Theoretical Model

2.2.1 Model Setup

Environment and Technology

There is one risk-neutral principal ("she") and one risk-neutral agent ("he"). At the beginning of every period $t \in \{1, ..., T\}$, with $1 < T < \infty$, the principal makes an employment offer to the agent. If the agent accepts the offer, he chooses an effort level

$e_t \geq 0$, which is associated with effort costs $c(e) = e^3/3$.[1] Furthermore, effort determines the probability with which a positive output – that is subsequently consumed by the principal – is realized. More precisely, the output is $y_t \in \{0, \theta\}$, with Prob$(y_t = \theta) = e_t$. Below, we will impose further assumptions to always guarantee an interior solution. If the agent rejects the offer, both players consume their exogenous outside options which are set to zero for simplicity.

Payments, Information & Contracts

The employment offer includes a prospective compensation package. It consists of a fixed wage w_t and discretionary bonus payments. An *output-based* bonus b_t is supposed to be paid if $y_t = \theta$ (it is without loss of generality to assume that no output-based bonus is paid if $y_t = 0$), an *effort-based* bonus B_t is supposed to be paid if the principal's requested effort level is chosen by the agent.

The output realization y_t is verifiable, and formal spot contracts can be used to enforce payment of b_t. Effort can be observed by both parties, however is not verifiable. Therefore, payment of B_t can only be enforced within a self-enforcing dynamic arrangement, a so-called relational contract. The agent is protected by limited liability, hence $w_t,\, b_t \geq 0$ (this assumption is not needed for most of our results).

Note that the agent's compensation, consisting of w_t, b_t, and B_t, does not only have to contain monetary components. It is a rather common perception in the literature on relational contracts (and beyond) that money is not the only source of motivation inside firms. For instance, Gibbons and Henderson (2012) conceive an individual's payoffs to include "everything that might affect an individual's experience of his or her job, including factors such as job assignment, degree of autonomy, status with the firm or work group, and other intangibles such as feelings of belonging or that one is making a difference" (Gibbons and Henderson, 2012, p. 1353). In the following, though, we will stick to the terms wage and bonus payments when referring to the agent's compensation to keep descriptions straightforward.

[1] We assume this specific functional form for analytical tractability. Other (convex) cost functions would deliver similar results.

Preferences and Equilibrium

Provided the agent has accepted the principal's employment offer at the beginning of a period t, and denoting the on-path effort level e_t^*, the principal's per-period profits on the equilibrium path are

$$\pi_t = e_t^* (\theta - b_t) - B_t - w_t.$$

The agent is also risk-neutral and in addition has preferences for reciprocity which are activated by any *non-contingent* payment the agent receives and thus seemingly by fixed wages. However, in general, a relational contract can either use current payments (in the form of bonuses) or future rents to motivate current effort. We rule out that the agent's preferences for reciprocity are triggered by wages paid as a reward for past effort. It turns out that in our setting it is without loss of generality to assume that only current bonus payments are used to incentivize the agent. Taking this into account, the agent's preferences for reciprocity are indeed activated by all fixed wage payments. Then, upon accepting the principal's offer, the agent's per-period utility on the equilibrium path is

$$u_t = e_t^* b_t + B_t + w_t - \frac{e_t^3}{3} + \eta w_t e_t^* \theta.$$

The parameter $\eta \in [0, \infty)$ captures the agent's inherent preferences for reciprocity and lets the principal's output (potentially) enter his utility. Note that the agent's preferences for reciprocity in period t are only activated by wage payments received in period t, and not by received past or expected future payments. Furthermore, η remains constant across periods, with one exception. If the principal has promised to pay a bonus B_t but reneges on that promise even though the agent has exerted the desired effort level, η drops to zero in all subsequent periods. We discuss our assumptions concerning the agent's preferences for reciprocity in the following section.

Finally, principal and agent agent share the discount factor $\delta \leq 1$, and we can use the following recursive formulations for players' discounted payoff streams:

$$\begin{aligned}\Pi_t =& e_t^* (\theta - b_t) - B_t - w_t + \delta \Pi_{t+1} \\ U_t =& e_t^* b_t + B_t + w_t - c(e_t^*) + \eta w_t e_t^* \theta + \delta U_{t+1}\end{aligned}$$

In what follows, we are interested in a subgame-perfect equilibrium that maximizes the principal's profits at the beginning of the game, Π_1.

Discussion of Assumptions

Before deriving properties of a profit-maximizing subgame-perfect equilibrium, we discuss our assumptions regarding the agent's preferences for reciprocity.

Our approach yields a hybrid between outcome- and intention-based reciprocity. On the one hand, intentions are not formally considered. On the other hand, reciprocity is only triggered by non-contingent payments and disappears once the principal breaks a promise. A purely output-based formulation would not contain these two properties. Therefore, our approach can generally be compared to Falk and Fischbacher (2006b), where an individual's reciprocal inclinations depend on outcomes, but also on the available options one's counterpart has at hand. This takes into account empirical evidence that while individuals respond to outcomes, intentions often matter as well (cf. Falk et al., 2008; Fehr et al., 2009a; Camerer and Weber, 2013).

We also assume that reciprocity only enters the agent's stage-game payoffs. However, one might expect reciprocal inclinations to also depend on past events in the employment relationship. We capture this idea by letting η drop to zero after a deviation by the principal as a simple way to take a potential history-dependence into account.

Furthermore, reciprocal behavior is triggered by a positive fixed wage (or put differently, by a wage that is above a reservation wage which is set to zero), and not by the agent's actual or perceived rent. This assumption is driven by two aspects: First, there is evidence (in particular from the lab) that generous wages cause reciprocal behavior even in the absence of performance-based incentives (cf. Fehr et al., 2009b; Charness and Kuhn, 2011). Second, incentives often cannot be provided *without* granting the agent a rent (for example, if the agent is protected by limited liability as in our case). We do not want reciprocal behavior being caused by rents that the agent collects in any case, but only by extra rents that the principal *chooses* to pay. Again, this relates to the

idea that also intentions – and not only outcomes – are supposed to matter.[2]

The reciprocity term in the agent's utility function also contains θ, and hence the extent to which the principal benefits from the agent's effort. This follows evidence pointing out that an important factor for reciprocity is the agent's assessment of the value generated for the principal (Hennig-Schmidt et al., 2010; Englmaier and Leider, 2012b).

Finally, we focus on positive reciprocity. In Section 2.5, we briefly discuss negative reciprocity.

2.2.2 Formal Spot Contracts

We first derive a profit-maximizing spot contract and hence omit time subscripts. There, in order to always guarantee an interior solution, we impose the technical assumptions $\theta < 3$ and $\frac{\eta\theta^2}{2} < 1$. Besides serving as a benchmark, such a contract will also be offered in period T, the last period of the game. In a spot contract, it is not possible to enforce a bonus based on non-verifiable effort, hence $B = 0$. Therefore, the only means to provide *direct* incentives is an output-based bonus b. Indirectly, though, the agent will also be incentivized by a positive fixed wage w. Because of his inherent preferences for reciprocity, a positive wage lets the output value also enter the agent's utility. Then, the agent's and principal's interests become partially aligned. Taking a slightly different perspective, one might also regard positive values of w and η as triggering a reduction of the agent's effective effort to $(e^*)^3/3 - \eta w e^* \theta$.

Given b and w, and presuming he decides to work for the principal, the agent chooses effort e^* in order to maximize his per-period utility $u = eb + w - e^3/3 + \eta w e \theta$. The conditions for using the first order approach hold, hence the agent's incentive compatibility (IC) constraint gives

$$e^* = \sqrt{b + \eta w \theta}. \tag{IC}$$

[2] Note that we are not able to use an approach introduced by Englmaier and Leider (2012a). They assume that the principal requests an effort level from the agent, and that the associated rent triggers reciprocal behavior. Then, consistency between actual and requested effort is required in equilibrium. This consistency requirement could not be met in our setting because actual effort would always be higher than requested effort (it is feasible in Englmaier and Leider (2012a) since they assume that effort is binary).

The principal sets b and w to maximize her expected per-period profits $\pi = e^* (\theta - b) - w$. However, she has to take into account that accepting the contract must be optimal for the agent. This is captured by the agent's individual rationality (IR) constraint,

$$e^* b + w - \frac{(e^*)^3}{3} + \eta w e^* \theta \geq 0. \tag{IR}$$

Furthermore, because of limited liability, payments must be non-negative.

Concluding, the principal's problem is to

$$\max_{w,b} e^* (\theta - b) - w,$$

subject to (IR) and (IC) constraints, and $w,\, b \geq 0$.

As a first result, we show that either only wages or bonus payments are used, not a combination of both.

Lemma 2.1. *Either bonus or wage payments are used to give incentives in a profit-maximizing spot contract. More precisely, there exists a threshold $\overline{\eta} > 0$ such that $b > 0$ and $w = 0$ for $\eta < \overline{\eta}$, and $b = 0$ and $w > 0$ for $\eta \geq \overline{\eta}$.*

The proof, as well as all other omitted proofs, can be found in Appendix B.1.

Intuitively, bonus and wage payments are not used together because they are substitutes in the principal's profit function: For equilibrium effort $e^* = \sqrt{b + \eta w \theta}$, the cross-derivative of per-period profits π with respect to w and b is negative. Put differently, for any bonus level the marginal profitability of using a bonus is decreasing in the wage. Therefore, either a pure *bonus contract* ($b > 0$ and $w = 0$) or a pure *reciprocity contract* ($b = 0$ and $w > 0$) is implemented by the principal.[3] When a reciprocity contract is used, a higher value of η is associated with lower effective effort costs and consequently also larger profits. Since the profitability of a bonus contract with a zero-wage is naturally unaffected by the size of η, a reciprocity contract is optimal given a sufficiently high η.

[3] However, note that this result is subject to the specific functional form of the agent's effort cost function – with other functional forms, bonus and wage payments might very well be used together. Still, our further results would not be qualitatively affected in this case.

Finally, note that profits under the bonus contract ($\pi = \frac{2\theta}{3}\sqrt{\frac{\theta}{3}}$) are strictly positive. Therefore, the principal will make an employment offer to the agent in any case.

The positive relationship between intrinsic reciprocity η and outcomes in the reciprocity contract is summarized in the following corollary:

Corollary 2.1. *Given $\eta \geq \overline{\eta}$ and hence a reciprocity contract is used, $\frac{de^*}{d\eta} > 0$, $\frac{d\pi}{d\eta} > 0$, $\frac{dw}{d\eta} > 0$, and $\frac{dU}{d\eta} > 0$.*

Proof. Using a reciprocity contract, outcomes are $e^* = \eta\theta^2/2$, $\pi = \frac{\eta\theta^3}{4}$, $w = \eta\theta^3/4$ and $U = \frac{\eta\theta^3}{4} + \frac{\eta^3\theta^6}{12}$, which all are increasing in η. □

Finally, note that the agent always receives a rent, that is, $U > 0$ under both types of contracts. This is straightforward for the reciprocity, but also for the bonus contract because of the agent's limited liability constraint. However, note that even without this constraint (implying that using a bonus contract, the principal could extract the whole rent), a reciprocity contract would eventually be optimal because of the associated reduction of effective effort costs. In this case, only the threshold $\overline{\eta}$ would be larger.

2.2.3 Relational Contracts

Now, we analyze how self-enforcing relational contracts based on effort can be used to motivate the agent. The principal would generally prefer an effort-based over an output-based contract because – as derived in the previous section – limited liability requires to grant the agent a rent in the latter case. Two aspects are of particular interest, namely the enforceability of relational contracts and whether and how they are affected by the agent's preferences for reciprocity. We will explore these aspects in the next subsections and furthermore derive the properties of a profit-maximizing relational contract.

Preliminaries

Relational contracts are self-enforcing implicit arrangements between economic agents. They work if the future surplus of continuing a cooperative relationship is sufficiently large compared to the future surplus without cooperation. Informally speaking, in our

setting, a relational contract involves a request from the principal to the agent to exert an effort level e_t^* (recall that effort can be observed by the principal), combined with a promise to pay the reward B_t in return. However, it must be in the principal's interest to pay the bonus when supposed to do so, which is specified by a dynamic enforcement (DE) constraint for every period t,

$$-B_t + \delta\Pi_{t+1} \geq \delta\tilde{\Pi}_{t+1}. \tag{DE}$$

Π_{t+1} describes the principal's on-path and $\tilde{\Pi}_{t+1}$ her off-path continuation profits. The (DE) constraint captures the requirement that future on-path profits must be sufficiently large compared to future off-path profits so that they offset today's costs of paying the bonus. Note that since the period-t output has already been realized and consumed, it is not included in the (DE) constraint and hence considered as sunk by the principal when making the decision whether or not to pay B_t. (DE) indicates that a bonus payment is only feasible if $\Pi_{t+1} > \tilde{\Pi}_{t+1}$, i.e., if future equilibrium play can be made contingent on the principal's current behavior.

Generally, relational contracts require a (potentially) infinite time horizon because of a standard unraveling argument that can be applied once a predetermined last period exists. Then, the equilibrium outcome in the last period is unique, implying the same for all subsequent periods. In our case, however, the situation is different if the spot reciprocity contract is (strictly) more profitable than the spot bonus contract, i.e., if $\eta > \overline{\eta}$. In this case, the principal's behavior in a period $t < T$ affects her future profits because A) the optimal spot contract is implemented (at least) in period T, and B) refusing to pay a promised bonus B_t lets η drop to zero. Therefore, the spot reciprocity contract is not feasible anymore once the principal reneged on a promise, and reneging is costly if $\eta > \overline{\eta}$.

In addition, relational contracts are unfeasible once the principal refused to pay a promised bonus (e.g., Abreu, 1988, shows that an observable deviation should be punished by a reversion to a player's minmax-payoff). Hence, after a deviation by the prin-

cipal, spot bonus contracts are implemented in every subsequent period,[4] and off-path continuation profits are $\tilde{\Pi}_t = \sum_{\tau=t}^{T} \delta^{\tau-t} \frac{2\theta}{3}\sqrt{\frac{\theta}{3}} = \frac{1-\delta^{T-t+1}}{1-\delta}\frac{2\theta}{3}\sqrt{\frac{\theta}{3}}$.

For $\eta \leq \overline{\eta}$, equilibrium profits in period T are unique, hence relational contracts are unfeasible, and the profit-maximizing spot bonus contract is implemented in every period.[5] Hence, to keep the analysis interesting, we will from now on assume that

$$\eta > \overline{\eta}.$$

This assumption is backed by the data we use for our empirical analysis, where the variable measuring positive reciprocity is relatively high for most individuals. The assumption also implies that a relational contract does not involve an output-based bonus b_t because it is dominated by using a fixed wage. Therefore, incentives are potentially given by a non-discretionary fixed wage w_t and an effort-based bonus payment B_t.

As mentioned above, we also assume that in a given period t, the agent is only motivated by period-t payments. This assumption is without loss of generality, for the following reasons: Generally, incentives in relational contracts can be provided via contingent current or future payments. In a setting like ours, though, replacing contingent future payments with the equivalent and appropriately discounted current amount does neither affect today's profits nor any constraints. Furthermore, sticking to current discretionary payments simplifies our analysis because the agent's reciprocal preferences are triggered only by unconditional payments. If a future wage was paid as a compensation for previous effort, we would have to differentiate between wages that are paid as a compensation for past effort and those that are not (if *any* fixed-wage payment triggered reciprocal behavior, using wages would be effectively cheaper than bonus payments, making it optimal to back-load wages as much as possible).

[4] As those are profitable, subgame perfection implies no termination after a deviation.

[5] We assume that no formal long-term contracts based on output realizations are feasible. This can be endogenized by assuming that the principal is not able to commit to fire the agent. If she were able to do so, a long-term contract involving a positive termination probability following a number of low output realizations would yield higher profits than a series of spot contracts (Ohlendorf and Schmitz, 2012; Fong and Li, 2017). Still, the possibility to write such a long-term contract would have no qualitative effect on our results, in particular with respect to the impact of the agent's reciprocal preferences on a profit-maximizing agreement.

Incentive Compatibility

The relational contract specifies an effort level e_t^* the agent is supposed to exert on the equilibrium path. He will do so if his (IC) constraint is satisfied. Before stating this constraint, we have to specify what happens if he deviates in a period t. First, he does not receive the period-t bonus B_t. Second, we assume that after a deviation by the agent, the reciprocity parameter is *not* reduced but remains at η, and third, that continuation play is not affected by the agent's behavior. The second assumption is not crucial for our results, but it seems more realistic to presume that the degree of the agent's reciprocal preferences only depends on the principal's behavior. It implies that if the agent deviates, he does not necessarily deviate to an effort level of zero. The third assumption, however, is important because the agent's rent under a sequence of (spot) reciprocity contracts might be higher than under a relational contract (see below). Therefore, if the agent's behavior affected continuation play, and in particular if a deviation triggered a breakdown of the relational contract, the agent might be tempted to deviate in order to enjoy the higher rent of a sequence of reciprocity contracts in the future.

Concluding, for any off-path effort level $\tilde{e}_t$, the (IC) constraint equals

$$B_t + w_t - \frac{(e_t^*)^3}{3} + \eta w_t e_t^* \theta \geq w_t - \frac{(\tilde{e}_t)^3}{3} + \eta w_t \tilde{e}_t \theta.$$

Subgame perfection implies that if the agent deviates, he will select an effort level $\tilde{e}_t = \text{argmax}\,(-e^3/3 + \eta w_t e\theta)$, i.e., $\tilde{e}_t = \sqrt{\eta w_t \theta}$. Hence, the (IC) constraint becomes

$$B_t - \frac{(e_t^*)^3}{3} + \eta w_t e_t^* \theta \geq \frac{2\left(\sqrt{\eta w_t \theta}\right)^3}{3}. \tag{IC}$$

This implies that an (IR) constraint for the agent is automatically satisfied because his per-period rent, $B_t + w_t - (e_t^*)^3/3 + \eta w_t e_t^* \theta$, is non-negative given the (IC) constraint. Also note that $e_t^* \geq \tilde{e}_t$ (because $B_t \geq 0$).

The Complementarity of Relational and Reciprocity-Based Incentives

In this section, we derive some first results and show that reciprocity-based incentives also improve the performance of the relational contract.

To simplify the principal's problem, note that the (IC) constraint must bind in any profit-maximizing equilibrium. If it did not bind, the bonus B_t could be slightly reduced, which would increase profits and relax the (DE) constraint without violating the (IC) constraint. This allows us to plug $B_t = (e_t^*)^3/3 - \eta w_t e_t^* \theta + 2/3 \left(\sqrt{\eta w_t \theta}\right)^3$ into the (DE) constraint, which yields

$$\frac{(e_t^*)^3}{3} - \eta w_t \theta e_t^* \leq \delta \left(\Pi_{t+1} - \tilde{\Pi}_{t+1}\right) - \frac{2}{3}\left(\sqrt{\eta w_t \theta}\right)^3. \tag{DE}$$

The enforceability of relational contracts is generally determined by a comparison of today's effort costs with discounted future payoffs (compared to future deviation payoffs). Only if the latter are large enough, they are sufficient to cover today's costs of exerting effort. Here, two additional terms enter which are implied by the agent's preferences for reciprocity; first, reciprocal preferences reduce the necessary bonus payment to achieve a certain effort level e_t^*; second, if the agent deviates, he still selects a positive effort level given the wage is positive.

Concluding, for $\eta > \overline{\eta}$, the principal's problem is to maximize

$$\Pi_1 = \sum_{t=1}^{T} \delta^{t-1} \pi_t,$$

subject to a (DE) constraint for every period t, and subject to $w_t \geq 0 \ \ \forall\, t$.[6]

The equilibrium is sequentially efficient, hence the problem is equivalent to maximizing $\pi_t = e_t\theta - B_t - w_t = e_t\theta - \left((e_t^*)^3/3 - \eta w_t e_t^* \theta + 2/3 \left(\sqrt{\eta w_t \theta}\right)^3\right) - w_t$ in every period t, subject to the relevant constraints.

After generally addressing the enforceability of a relational contract, we will now ana-

[6] Note that in period T, the (DE) constraint equals $\frac{(e_T^*)^3}{3} - \eta w_T \theta e_T^* \leq -\frac{2}{3}\left(\sqrt{\eta w_T \theta}\right)^3$, which for $e_T^* = \sqrt{\eta w \theta}$ (the agent's effort in a spot reciprocity contract) is trivially satisfied.

lyze the relationship between reciprocal and effort-based incentives. To do so, we first abstract from issues of enforceability. Put differently, we assume that the (DE) constraint does not bind, i.e., is satisfied for the principal's preferred effort level and derive respective effort and wage levels. Note that this situation is equivalent to one where formal contracts based on effort would be feasible.

Lemma 2.2. *Assume the (DE) constraint does not bind in a period $t < T$. Then, setting a strictly positive wage is optimal.*

Lemma 2.2 implies that even if the principal is not restricted in setting her preferred effort-based bonus B_t, she still decides to pay a strictly positive fixed wage (which amounts to $w_t = \left(\eta^2\theta^3 - 1\right)^2 / 4\eta^3\theta^3$), and the agent receives a rent. This is because the agent's concern for reciprocity reduces his effective effort costs, *but only in combination with a strictly positive wage* w_t. The agent's effective effort costs are $(e^*)^3/3 - \eta w e^* \theta$, and implemented effort $e_t^* = \left(1 + \eta^2\theta^3\right)/2\eta\theta$ is also strictly larger than the "standard" first best without reciprocal preferences, $\sqrt{\theta}$.[7] In the following, we will refer to the implemented effort and wage levels for a non-binding (DE) constraint as *first-best* levels. At these first-best levels, the costs for the principal to implement one additional unit of effort are the same when using relational as when using reciprocity-based incentives, and those costs are equal to the principal's marginal benefits.

in a next step, we explicitly take the enforceability of relational contracts into account and assess how the agent's preferences for reciprocity affect outcomes with a *binding* (DE) constraint.

Lemma 2.3. *Assume the (DE) constraint binds in a period $t < T$. Compared to the situation with a non-binding (DE) constraint, the fixed wage is larger and implemented is effort smaller.*

Besides reducing effective effort costs, a fixed wage also relaxes the principal's (DE) constraint by reducing the bonus that must be paid to implement a given effort level.

[7] The condition $\left(1 + \eta^2\theta^3\right)/2\eta\theta > \sqrt{\theta}$ is equivalent to $\left(1 - \sqrt{\eta^2\theta^3}\right)^2 > 0$, where the strict inequality holds since $\eta > \overline{\eta}$ also implies $\eta^2\theta^3 > 1$.

Therefore, if the (DE) constraint binds (i.e., it does not hold for first-best effort), the fixed wage is larger than when it does not bind.

All this implies that relational and reciprocity-based incentives are complements *at any given point in time.* Reciprocity-based incentives reduce effective effort costs. Furthermore, they relax the (DE) constraint and therefore allow to enforce more effort within the relational contract. In the following, we will explore how this interaction evolves over time.

Relational and Reciprocity-Based Incentives as Dynamic Substitutes

In this section, we derive conditions for when the (DE) constraint actually binds, and in particular how this relates to the tenure of the employment relationship. This allows us to characterize how the optimal use of relational and reciprocity-based incentives evolves over the course of the employment relationship.

Generally, the (DE) constraint might or might not bind in any period $t < T$, depending on discount factor δ, reciprocity parameter η and productivity θ. Furthermore, the (DE) constraint becomes tighter in later periods.

Lemma 2.4. *The principal's dynamic enforcement constraint might or might not bind in period $T-1$. More precisely, for any discount factor δ, the (DE) constraint holds for first-best effort and wage levels if η is sufficiently large. For any values η and θ, the (DE) constraint does not hold for first-best effort and wage levels if the discount factor is sufficiently small.*

Furthermore, $\Pi_{t-1} - \tilde{\Pi}_{t-1} > \Pi_t - \tilde{\Pi}_t$ *for all* $t \leq T$.

The principal's commitment in a relational contracts is given by what she has to lose given she deviates. If the discount factor is small, she cares less about a potential reduction of future profits and is therefore less willing to compensate the agent for his effort. Furthermore, a larger reciprocity parameter η increases future profits on the equilibrium path (by more than future off-path profits), and furthermore reduces today's effective effort costs (by more than first-best effort goes up). The second part of

Lemma 2.4 states that the difference between on- and off-path continuation profits goes down over time. The intuition for this is driven by two aspects. First, the remaining time horizon and therefore the periods in which profits can be generated is reduced as time elapses. Second, this triggers a re-enforcing effect because implementable effort in a period is increasing in the difference between on- and off-path continuation profits. Since $\Pi_T - \tilde{\Pi}_T > 0$, the (DE) constraint allows to implement a larger effort level in period $T-1$ than in period T. Therefore, per-period on-path profits in period $T-1$ are larger than in period T (whereas per-period off-path profits are the same in every period), and implementable effort in period $T-2$ is even larger than in period $T-1$. Hence, the (DE) constraint in earlier periods is less tight than later on.

Lemma 2.4 implies that if the (DE) constraint *binds* in a given period $\tilde{t}$, it will also bind in all subsequent periods $t > \tilde{t}$. If it is *slack* in a given period $\hat{t}$, it will also be slack in all previous periods $t < \hat{t}$. This allows us to derive the following effort- and (fixed) wage-dynamics.

Proposition 2.1. *Equilibrium effort is weakly decreasing over time and equilibrium wage weakly increasing, i.e.,* $e_t^* \leq e_{t-1}^*$ *and* $w_t \geq w_{t-1}$. *Both inequalities hold strictly if and only if the (DE) constraint binds in period* t.

Furthermore, $e_t^* < e_{t-1}^*$ *and* $w_t > w_{t-1}$ *imply* $e_{t+1}^* < e_t^*$ *and* $w_{t+1} > w_t$, *whereas* $e_{t+1}^* = e_t^*$ *and* $w_{t+1} = w_t$ *imply* $e_t^* = e_{t-1}^*$ *and* $w_t = w_{t-1}$.

Proposition 2.1 states that the profit-maximizing equilibrium is characterized by a downward sloping effort and an upward sloping wage profile. As long as the future is sufficiently valuable for the (DE) constraint not to bind, both are time-invariant. Once the (DE) constraint binds, the principal cannot credibly promise her preferred bonus payment anymore. On the one hand, this reduces equilibrium effort. On the other hand, the principal responds with a wage increase which increases equilibrium effort – directly due to the agent's preferences for reciprocity, and indirectly because it relaxes the principal's (DE) constraint and allows her to request more effort from the agent. However, the effort increase caused by the higher wage does not fully compensate for the effort reduction caused by the binding (DE) constraint because the costs of implementing an

additional unit of effort are now larger with reciprocity-based than with relational incentives. As time proceeds, the (DE) constraint becomes tighter and tighter (Lemma 2.4). Hence, relational incentives are gradually substituted by reciprocity-based incentives (fixed wage ↑), with the substitution however being incomplete (effort ↓).

2.3 Predictions

In the previous sections, we derived the properties of a profit-maximizing long-term arrangement for an agent with given reciprocal preferences. Now, we will explore to what extent the agent's preferences for reciprocity affect his effort choices over the course of his career. This allows us to generate several comparative statics and consequently a number of testable predictions.

2.3.1 Effort

First, we derive a general result concerning the effect of an individual's reciprocity on effort. I our model , the effect of a larger η on effort is unambiguously positive irrespective of an individual's career stage.

Prediction 2.1. *More reciprocal individuals exert more effort.*

Proof. This relationship holds in any period: For periods $t < T$ and $\eta > \overline{\eta}$, $\frac{\partial e_t^*}{\partial \eta} = \frac{\eta^2\theta^3(1+\lambda_{DEt})-1}{2\eta^2\theta(1+\lambda_{DEt})} > 0$. For period T and $\eta > \overline{\eta}$, $\frac{\partial e_T^*}{\partial \eta} = \frac{\theta^2}{2} > 0$. Finally, for $\eta \leq \overline{\eta}$, $\frac{\partial e_t^*}{\partial \eta} = 0$ in all periods t. □

This prediction follows from our result that reciprocal and relational incentives are complements at any given point in time. There, providing incentives becomes cheaper if η goes up, hence more effort is implemented.

The next prediction picks up the results stated in Proposition 2.1 and refers to the dynamics with respect to effort.

Prediction 2.2. *Effort is lower in later stages of a career.*

Proof. This immediately follows from Lemma 2.1. □

Next, we explore how η affects incentive schemes and consequently outcomes at different stages of an individual's career. Prediction 2.3 yields the main result of this paper:

Prediction 2.3. *The positive effect of reciprocal preferences on effort becomes stronger over the course of the employment relationship.*

Proof. This follows from Lemma 2.4, where we show that λ_{DEt} increases over time, and $\frac{\partial e_t^*}{\partial \lambda_{DEt} \partial \eta} = \frac{1}{2\eta^2\theta(1+\lambda_{DEt})^2} > 0$. □

Prediction 2.3 follows from reciprocal and relational incentives being dynamic substitutes. When an agent approaches the end of his employment relationship, the incentive system puts more weight on reciprocal incentives, hence the role of η is intensified. Therefore, the reduction of incentive costs caused by a higher η is more pronounced and equilibrium effort reacts more strongly.

2.3.2 Utility

In the following, we derive predictions concerning the agent's utility. In the empirical section, we test these predictions using survey measures on an individual's job satisfaction. This section is also supposed to capture potential interactions between an individual's inherent preferences for reciprocity and his or her compensation. Although the SOEP contains measures such as "monthly wage" or "annual salary", those are incomplete if an individual's total compensation also contains non-monetary components, like career concerns, job assignment, status, feelings of belonging, that one is making a difference (see Gibbons and Henderson, 2012, p. 1353), or discretion over decisions[8]. However, all these aspects are supposed to increase an individual's satisfaction with his or her job, therefore the respective measure – and the agent's utility as the theoretical counterpart – seems better suited.[9]

[8] This has been observed by Cyert and March (1963) and later taken up by Li, Matouschek, and Powell, who state that "payments within organizations often take the form of promises about future decisions and decision making rather than monetary transfers" (Li et al., 2017, p. 218).

[9] Note that Dohmen et al. (2009) also explore the interaction between reciprocity and satisfaction. They compare individuals who are satisfied with their jobs with individuals who are not, and find that the positive effect of reciprocity on effort is only observed for the former (a result we are able

Our predictions on job satisfaction are collected in the following prediction:

Prediction 2.4.

a. More reciprocal employees are more satisfied with their jobs.

b. Job satisfaction increases over time.

c. The positive effect of reciprocal preferences on satisfaction becomes stronger over the course of the employment relationship.

The proof of Prediction 2.4 can be found in Appendix B.1.

Generally, agents' larger utility is higher if one unit of effort is implemented by reciprocal rather than by relational incentives. The intuition for Prediction 2.4 is then similar to the intuition underlying Predictions 2.1–2.3: In any period t, reciprocity-based incentives are more important for larger values of η, therefore the agent also is better off in this case. Furthermore, because reciprocity-based incentives assume a larger role later on, the positive effect of η on an agent's utility becomes stronger over time.

Having established these predictions, we now turn to evaluating them empirically.

2.4 Empirical Analysis

To explore our predictions, we use data from the German Socio-Economic Panel (SOEP). The SOEP is a yearly panel survey that is representative of the German population and goes back to 1984. It contains a wide range of questions on the personal and socioeconomic situation as well as labor market status and income of respondents. We use the same data set as Dohmen et al. (2009) and largely keep their empirical specification for the sake of comparability, however are particularly interested in exploring how the

to replicate). Our model would also predict such an outcome if unsatisfied agents were regarded to have had a bad experience with their employer, and in particular perceive the latter to have reneged on the relational contract (after which η drops to zero). But our model also allows to go beyond this rather straightforward prediction, and assess the on-the-equilibrium-path implications of observing different satisfaction/utility levels.

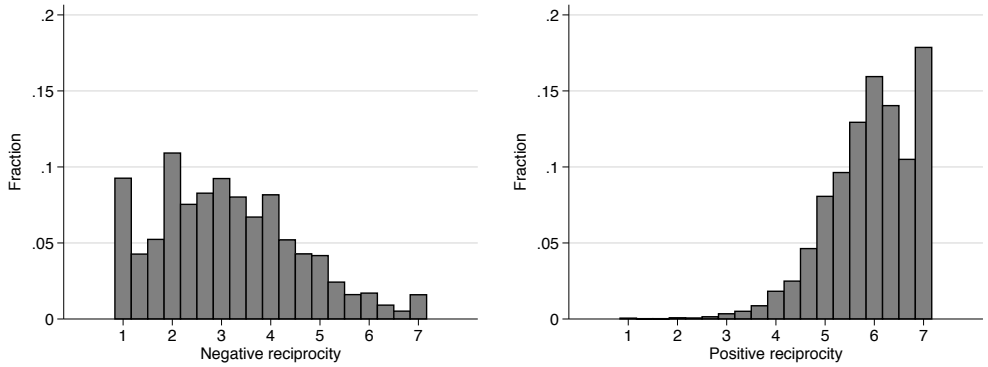
existence of intrinsic reciprocity affects dynamic incentive systems. Our focus on employment relationships lets us restrict our analysis to a subsample of all SOEP respondents and only consider employees. This excludes individuals who are unemployed, retired, self-employed, in compulsory military or community service, or in training and education. We further exclude employees below the age of 25 and above the age of 65 to avoid sample selection issues due to endogenous retirement decisions, leaving us with a sample of 9,221 individuals who participated in the 2005 wave of the survey. For our purposes, the 2005 wave is of particular interest as it contains a total of six questions that are designed to capture individual reciprocal inclinations. Note that this measure has also been included in the 2010 and 2015 waves of the SOEP; however, as those do not include other measures that we regard as important for the relation to our theoretical model, for example retirement propensity (see below), we decided to use these waves only for additional analyses, which can be found in Section 2.4.2.

The reciprocity items developed by Perugini et al. (2003) capture what they define as a personal norm of reciprocity, that is, to what extent an individual has internalized the norm of reciprocal behavior. Participants are asked to rate how well six statements (three for positive, three for negative reciprocity) apply to themselves on a seven-point Likert scale.[10] The item average then determines each person's strength of reciprocity. Figure 2.2 shows the distribution of positive and negative reciprocity among survey participants, revealing that while there is quite some variation in negative reciprocity, positive reciprocity is strongly pronounced.

It is important to note that positive and negative reciprocity constitute different traits. This is supported by the observations that the correlation between the two is rather low ($p = .052$), that the traits have different determinants (Dohmen et al., 2008), and that the six items can be represented by two distinct orthogonal principal components (Dohmen et al., 2009). In our study, we focus on the effects of positive reciprocity and

[10] Specifically, the items measuring positive reciprocity are "If someone does me a favor, I am prepared to return it", "I go out of my way to help somebody who has been kind to me before", and "I am ready to undergo personal costs to help somebody who helped me before", while the items "If I suffer a serious wrong, I will take revenge as soon as possible, no matter what the cost", "If somebody puts me in a difficult position, I will do the same to him/her", and "If somebody offends me, I will offend him/her back" are meant to capture negative reciprocity.

Figure 2.2: *Distribution of negative and positive reciprocity*



Notes: The figure plots the distribution of negative and positive reciprocity for the 9,221 employee respondents of the 2005 SOEP wave who answered all six reciprocity questions. The individual inclination for negative and positive reciprocity are calculated by taking the average of the three questions that are targeted at the respective dimension.

discuss this choice below, in Section 2.5.

To study the effect of positive reciprocity on (non-contractible) effort, we use overtime work as proxy for the latter (following Dohmen et al., 2009). More specifically, overtime is a binary variable indicating whether the employee has worked overtime hours in the month preceding the interview or not. Two aspects are important. First, this measure does *not* exclude the possibility that an employee is compensated for working overtime, which one might argue does not capture our notion of non-contractible effort. However, effort in our model does not only capture actual working time, but also other aspects that benefit the firm and are costly to the employee, such as flexibility in one's working arrangements. Therefore, even if overtime is compensated, it can very well include non-contractible aspects of effort. Moreover, in Section 2.4.2, we replicate our results using *unpaid* overtime as the dependent variable, as well as overtime that cannot be accumulated in a work-time account. Second, the overtime measure is binary, whereas one might argue that our predictions are better addressed by analyzing overtime hours (i.e., the intensive margin). This is taken care of in Section 2.4.2 as well by exploring the effect of reciprocal preferences on overtime hours. Finally, we rule out that our results are entirely driven by cohort effects. There, we include later waves of the survey that also include the reciprocity measure and conduct a panel analysis.

2.4.1 Main Results

Reciprocity and Effort

We first examine the effect of reciprocity on effort. We do so by estimating cross-sectional regressions and controlling for several other influence factors. In detail, our controls include gender, age, years of education, full-time and part-time work experience, tenure in the recent position (all included in a Mincer-type fashion), a dummy variable for part-time employment, and indicator variables for industry sector, firm size, and occupational status. To estimate the effect of reciprocity on the propensity to work overtime, we employ a logit specification with standard errors clustered at the household level. Column (1) of Table 2.1 gives the results of a regression of the binary dependent variable indicating whether the employee has worked overtime in the month preceding the interview on reciprocity and the set of controls. Our results confirm Prediction 2.1 (and are similar to those of Dohmen et al., 2009) – the propensity to work overtime is increasing in the degree of positive reciprocity.

Next, we test Prediction 2.2 and explore how effort evolves over time. In particular, we examine the effect of age on the propensity to work overtime. Assessing the net relationship in the regression in column (1) of Table 2.1, the age coefficient in the effort regression is negative and significant, indicating that the propensity to work overtime is decreasing with employee age, confirming Prediction 2.2. Figure 2.3 further explores the relationship between employee age and effort by depicting a local polynomial regression of the residual variance of overtime on age, thus illustrating the adjusted, non-linear relationship between the two variables. It can be seen that the pattern largely overlaps with the linear regression line with a few small exceptions at the lower and upper end of the age continuum.

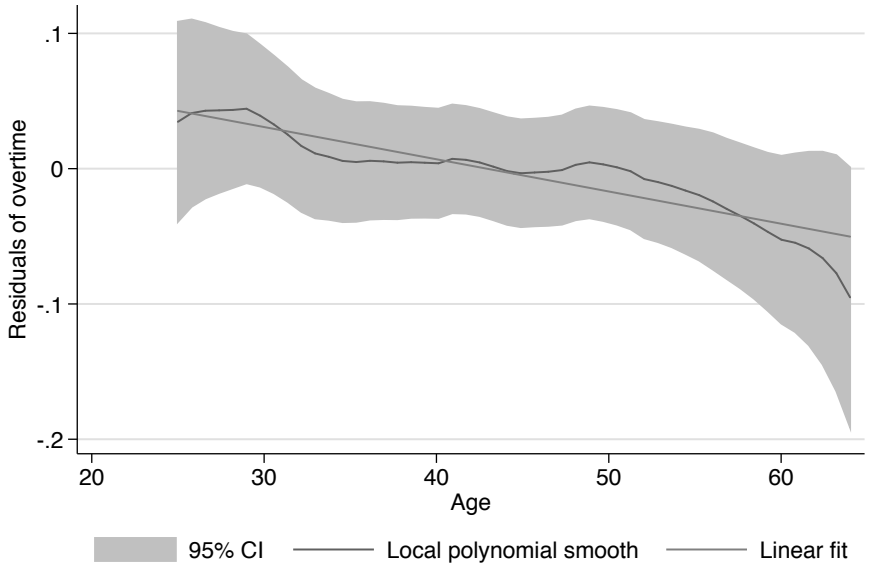
Having explored the impact of positive reciprocity and age on overtime, we now turn to our central question: how is the effect of positive reciprocity on effort changing over the course of an employment relationship? To approach this, we modify the estimation equation used so far by adding an interaction term of positive reciprocity and a dummy variable indicating proximity to the last periods of an employment relationship. In

Table 2.1: *Effect of reciprocity on effort*

DV: Overtime (Y/N)	(1) Overall	(2) Age cutoff	(3) Retirement propensity
Positive reciprocity	0.0176***	0.0184***	0.0175***
	(0.00641)	(0.00643)	(0.00642)
Age (in years)	-0.00785***		
	(0.00166)		
Age $\geq$ 60		-0.0812**	
		(0.0332)	
Age $\geq$ 60 # Reciprocity		0.0838***	
		(0.0317)	
Retire = 1			-0.0690**
			(0.0291)
Retire # Reciprocity			0.0551**
			(0.0278)
Observations	7,019	7,019	7,019
Pseudo R^2	0.0691	0.0679	0.0677

Notes: Estimations are based on the 2005 survey wave of the SOEP. Robust standard errors in parentheses clustered at household level. Table reports marginal effects at the mean calculated after logit regressions, with marginal effects for interactions reflecting the difference in slope for reciprocity between the groups. Reciprocity score standardized. Controls include years of education, gender, years of full time and part time work experience (linear and squared terms), a dummy for part-time employment, job tenure in current position (linear and squared), indicator variables for industry sector (services, agriculture, energy or mining, manufacturing, construction, trade, transport, bank or insurance) and firm size (less than 100, 100-199, 200-1999, more than 2000 employees), and an indicator variable for occupation status (white collar, blue collar, public sector). *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Figure 2.3: *Local polynomial smooth of effort on age*

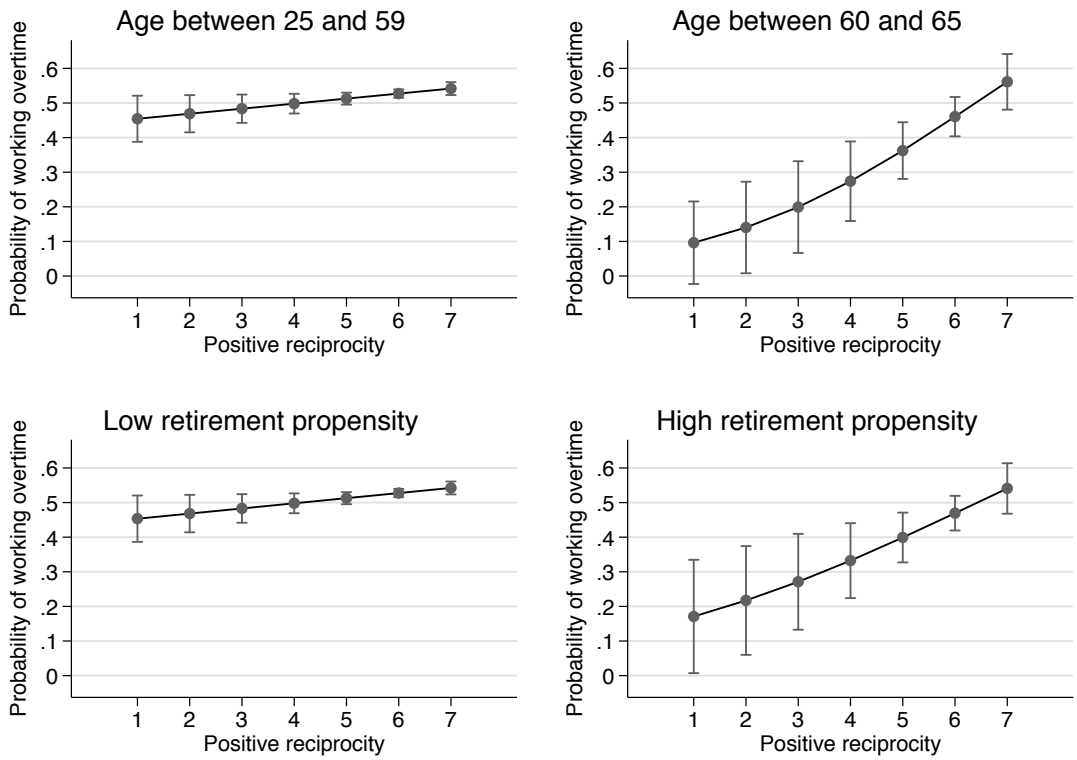


Notes: Figure plots a linear fit and local polynomial smooth of the residual variance of the propensity to work overtime on age. Estimations are based on the 2005 wave of the SOEP.

accordance with our model predictions, a positive interaction coefficient would indicate that reciprocity is indeed more important in later stages.

As a first approach, we create a dummy variable indicating whether an individual is at least 60 years old. We argue that this subgroup is sufficiently close to the last periods of their employment relationships such that an insufficient future surplus restricts the enforceability of the relational contract (or, in more technical terms, such that the (DE) constraint binds), thus making intrinsic reciprocity relatively more important. Estimation results are reported in column (2) of Table 2.1. It can be seen that while the influence of reciprocity on the propensity to work overtime is positive for all workers, the positive relationship is significantly more pronounced for workers above the age cutoff, as indicated by the positive marginal effect of the interaction. Thus, it seems like the positive effect of positive reciprocity on the probability of working overtime is mostly driven by older workers, confirming Prediction 2.3. Note that these results are qualitatively robust to choosing other age cutoffs. Figure B.1 in Appendix B.2 depicts the estimated interaction coefficients for varying cutoffs. It can be seen that the inter-

Figure 2.4: *Predictive marginal effects of positive reciprocity on effort*



Notes: The figure plots predicted marginal changes of the propensity to work overtime at different levels of positive reciprocity depending on age group (upper panel) and a high vs. low propensity to retire within the next two years (lower panel), holding all other factors constant. Estimations are based on the 2005 survey wave of the SOEP. Error bars indicate 95% confidence intervals.

action coefficient is increasing in the cutoff value for age and that there is a jump at age 60, indicating that there is a more pronounced difference for this cutoff.[11]

To further illustrate this result, the upper panel of Figure 2.4 plots the predictive marginal effects of positive reciprocity on the propensity to work overtime (i.e., the expected propensity to work overtime depending on reciprocity score, holding all other influence factors constant) calculated separately for workers who are younger than 60 and those

[11] Of course, while we present results for one specific age cutoff throughout this section, we are aware of the fact that this constitutes a somewhat arbitrary choice, and keep other possible cutoffs and specifications in mind throughout. Generally, the overall pattern of interactions at different cutoffs is as expected from our model (i.e., while the influence of intrinsic reciprocity is always more pronounced for relatively older people, the difference between the groups becomes smaller in size when the division is made earlier).

who are at least 60 years old. While the effect of reciprocity on the propensity to work overtime seems overall positive, it is substantially more pronounced for employees who are at least 60 years old. This further confirms the main intuition of our model, namely that intrinsic reciprocity mostly matters near the end of employment relationships.

However, even though there is an officially regulated age for retirement in Germany, there might still be differences in individual retirement ages, and thus age might only be a noisy proxy for the stage of employees' careers. To address this concern, we utilize an additional question from the SOEP that asks employees how likely it is that they are going to retire within the following two years. This results in a binary variable that takes on the value of one if the respondent indicates that his/her probability of retiring within the following two years is at least 50 percent. We use this variable as an alternative dummy for the interaction with positive reciprocity. The estimation results can be found in column (3) of Table 2.1. It is evident that this alternative specification produces similar results to generating the dummy based on age: again, the marginal effect of positive reciprocity is significant, but the interaction is as well, indicating that the positive relationship between intrinsic reciprocity and effort is more pronounced among among employees who expect to retire within the following two years. Once more, we further illustrate our finding by plotting the predictive marginal effects of positive reciprocity on effort separately for low and high retirement propensity in the lower panel of Figure 2.4. It can be seen clearly that while the slope of both curves is positive, it is much steeper for employees with a high likelihood of retirement.

To sum up, assessing influences on propensity to work overtime among our sample of German employees, we find that positive reciprocity is positively related to overtime (confirming Prediction 2.1). While the propensity to work overtime decreases with age (confirming Prediction 2.2), the positive effect of reciprocity on overtime is stronger for employees towards the end of their careers (confirming Prediction 2.3).

Reciprocity and Job Satisfaction

To explore the effect of reciprocity on an employee's job satisfaction as reflected in Prediction 2.4, we make use of a survey question asking employees to rate their overall

job satisfaction.[12] We regress job satisfaction from the 2006 survey wave on reciprocity, controlling for age and our usual control variables. The reason for taking job satisfaction from the subsequent wave is to address concerns of reverse causality. Note that we do not control for an individual's (monetary) income because it constitutes a substantial part of the agent's utility (together with other, non-monetary, components of an employee's compensation package which we cannot observe).

The results shown in column (1) of Table 2.2 indicate that reciprocity indeed influences job satisfaction as the marginal effect of positive reciprocity is positive and significant.

Next, we explore the development of job satisfaction over time by examining the effect of age. Analogous to the relationship between age and the propensity to work overtime shown in Figure 2.3, Figure 2.5 depicts a local polynomial regression of the residual variance of job satisfaction on age to illustrate the adjusted, non-linear relationship. Here, it is evident that while in our theoretical model, job satisfaction increases over time, the relationship between job satisfaction and age in the data is not approximated very well by a linear relationship. Indeed, it seems like job satisfaction is more or less the same for workers under the age of 50 to then first decrease and subsequently increase to its highest level.

Having explored the overall effect of positive reciprocity and age on job satisfaction, we now again add an interaction term between reciprocity and a dummy variable indicating that the individual is at least 60. Column (2) of Table 2.2 presents the results of this estimation. It can be seen that the marginal effect is positive but not significant. The same holds when we use retirement propensity as indicator variable for the interaction, as can be seen in column (3) of Table 2.2. Nevertheless, we also explore different cutoff values by depicting the interaction coefficients resulting from employing our specification and varying the age cutoff. As can be seen in Figure B.2 in Appendix B.1, the marginal effects are positive and relatively constant in value for cutoffs above 53, but the confidence intervals become much wider as the cutoff approaches 60, which is probably why we are only able to qualitatively detect the pattern predicted by our model.

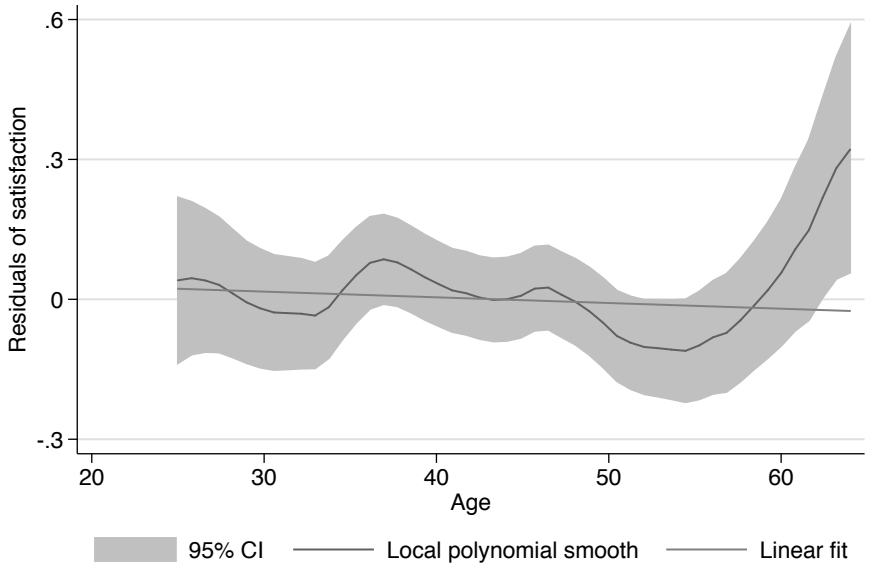
[12] The question is "How satisfied are you with your job?" and is to be rated on a scale from 0 (totally unsatisfied) to 10 (totally satisfied).

Table 2.2: *Effect of reciprocity on job satisfaction*

DV: Job satisfaction	(1) Overall	(2) Age cutoff	(3) Retirement propensity
Positive reciprocity	0.176***	0.172***	0.165***
	(0.0269)	(0.0274)	(0.0271)
Age (in years)	-0.00748		
	(0.00657)		
Age $\geq$ 60		0.390***	
		(0.149)	
Age $\geq$ 60 # Reciprocity		0.122	
		(0.152)	
Retire = 1			-0.267*
			(0.140)
Retire # Reciprocity			0.205
			(0.145)
Observations	6,218	6,218	6,218
Adjusted R^2	0.028	0.029	0.029

Notes: Estimations are based on the 2005 survey wave of the SOEP. Job satisfaction is taken from the 2006 wave. Robust standard errors reported in parentheses. Standard errors are clustered at household level. Table reports coefficients of OLS regressions. Reciprocity measure standardized. Controls include negative reciprocity, years of education, gender, years of full time and part time work experience (linear and squared terms), a dummy for part-time employment, job tenure in current position (linear and squared), indicator variables for industry sector (services, agriculture, energy or mining, manufacturing, construction, trade, transport, bank or insurance) and firm size (less than 100, 100-199, 200-1999, more than 2000 employees), and an indicator variable for occupation status (white collar, blue collar, public sector). *** p<0.01, ** p<0.05, * p<0.1.

Figure 2.5: *Local polynomial smooth of job satisfaction on age*



Notes: The figure plots a linear fit as well as a local polynomial smooth of the residual variance of job satisfaction on age. Estimations are based on the 2005 survey wave of the SOEP. Job satisfaction is taken from 2006.

However, when setting the cutoff to 52, 53, 54, 55, or 56, for instance, the interaction effect is significant. Therefore, provided job satisfaction is regarded as a proxy for the total utility an individual enjoys from his or her job, Prediction 2.4 is largely confirmed.

Recall that the results in this section are supposed to also capture the effect of reciprocity on compensation (which naturally assumes a large part of the utility stemming from employment). We do not include an analysis of the interaction between reciprocity and compensation because only its monetary part, which captures just a limited part of an individual's total compensation, is available in our data. Analyzing these, we see that an individual's annual labor income is increasing in the degree of positive reciprocity as predicted by the theoretical model. Furthermore, also in line with our theoretical approach (which would predict total compensation, the sum of wage and bonus payments, to go down over time), an individual's annual labor income is decreasing with age and particularly low above the age threshold. The effects of reciprocity for individuals above the cutoff and for those who expect to retire soon, however, are not significant.[13]

[13] Results are available from the authors upon request.

2.4.2 Robustness

Having mostly confirmed our predictions, we now show that our empirical results concerning effort are robust to different specifications. We first apply different overtime specifications and show that our results still hold. Then, we make use of data from two subsequent SOEP survey waves that also include reciprocity measures to verify that our results are not driven by cohort effects.

Unpaid overtime as effort measure

Up to now, we have – for the sake of comparability – relied on the same overtime measure as Dohmen et al. (2009). Since this measure also includes compensated overtime, one might question whether it is a good proxy for non-contractible effort. Thus, we rerun our main analysis using only *unpaid* overtime as a dependent variable. Doing so, our results are even more pronounced: for both the age cutoff and retirement propensity interactions, the main effect of positive reciprocity is positive but insignificant while the interaction terms are significant, indicating that the positive relationship between reciprocity and unpaid overtime is driven by workers who are relatively close to retirement. In addition, we use an even more restrictive overtime measure that only accounts for overtime that is neither compensated monetarily or non-monetarily (e.g., through work-time accounts). While the coefficients are not significant, they still point in the right direction. In sum, the overall picture provides further evidence for our main result. The full regression table and the predicted margins of the interactions can be found in Table B.1 and Figure B.3 in Appendix B.1.

Overtime hours

So far, our analyses have focused on the extensive margin of effort, i.e., whether employees have worked overtime or not. To further explore the influence of reciprocity on effort, we look at the intensive margin of overtime work by exploring the number of total and unpaid hours of overtime worked. This implies that our empirical measure of effort now is not binary anymore and therefore even closer to the continuous speci-

fication used in our theoretical model. Here, we estimate negative binomial models to account for the fact that we are now utilizing overdispersed count data and otherwise stick to our empirical strategy of estimating the overall effect of reciprocity and age before including interactions with age group and retirement propensity. Overall, we find that the pattern predicted by the theoretical model is also visible when we account for additional time spent at work as all interactions are positive, indicating that reciprocity has a stronger positive influence on the amount of overtime hours for workers who are at later stages of their careers. The full regression results can be found in Table B.2 in Appendix B.1.

Panel structure

While we have focused on the 2005 SOEP wave so far, the survey measures of reciprocity have been included in the SOEP questionnaire in the years of 2010 and 2015 as well. Note that, because there is some turnover among survey participants in every year, only 53% of the 2005 respondents participate in the 2010 wave, and 33% of them participate in the 2015 wave. Furthermore, in these additional survey years, not all variables that we use for our analysis have been included again, retirement propensity being one example. However, we are able to rerun our main analysis of the effect of positive reciprocity on overtime using data from all three waves.[14]

Theoretically, positive reciprocity is supposed to be a stable trait and any variation over time spurious. The SOEP data support this presumption as the intraclass correlation coefficient, a measure of consistency between more than two measurements, equals 0.418, which can be interpreted as fair (Cicchetti, 1994) – in particular taking into account the time distance between the waves, and that the SOEP contains self-reported measures. In addition, when we evaluate changes in positive reciprocity between the waves, we see that about 77% of changes are smaller than 1, and that the mean of changes is not significantly different from zero.

[14] In contrast, we are not able to do the same for job satisfaction as the 2016 measures are not available yet. As we take job satisfaction from the subsequent year to avoid issues with reverse causality, we are not able to use the 2015 data.

We estimate a random effects regression and a pooled logit regression.[15] The results are shown in Table B.3 in Appendix B.1. To account for cohort effects, we included birth year dummies in the panel estimations. First, if we stick to the age cutoff of 60, we find that we still see the expected patterns: Positive reciprocity has a significantly positive influence on the propensity to work overtime, while age has a negative influence. The interaction between being above the age cutoff and positive reciprocity is positive, even though insignificant. If we change the age cutoff to 55, though, the interaction coefficient is positive and significant.

To further explore the available data, we finally estimate the arguably most general specification. There, we not only include all three waves, from 2005, 2010, and 2015, but also use overtime hours as introduced in the previous section. The results can be found in Table B.4 in Appendix B.2. Here, the interaction coefficients are positive and significant at the 1% level for total as well as for unpaid overtime hours. These patterns are robust for changing the age cutoff from 60 to 55.

To sum up, we are able to replicate our main result when focusing on unpaid overtime instead of overtime in general, using overtime hours instead of a binary measure, as well as including two additional waves of the SOEP. This indicates that our results are not due to cohort effects, but indeed driven by the optimal design of incentive systems for employees with reciprocal preferences.

2.5 Discussion & Conclusion

We have shown that repeated-game incentives and preferences for positive reciprocity can interact in intricate ways. The two are dynamic substitutes, but complements once a specific point in time is considered. We have provided strong empirical support for this notion. No matter which specification we apply, the effect of reciprocity is always

[15] We refrain from estimating a fixed effects regression although it would enable us to isolate within-person effects. The reason is that we would need to focus on employees who switch age categories between 2005 and 2015 because only those have variation needed for using a fixed effects approach which gets rid of all stable effects. Furthermore, recall the turnover among survey participants in every year. Therefore, this procedure would only leave us with about 800 observations.

positive, and this effect is more pronounced for individuals who are relatively close to retirement – only significance levels differ. We think that this outcome is remarkable given the self-reported nature of the survey data we use.

We have not pursued the question of how negative reciprocity affects an optimal dynamic incentive scheme, though, and did this for two reasons. First, positive and negative reciprocity seem to describe quite different personality traits (recall that the correlation between the two only equals 0.052). Second, while the general *positive* effect of positive reciprocity on outcomes is well established and also straightforward from a theoretical perspective, the same is not true for negative reciprocity. There, the interaction is rather ambiguous.

On the one hand, the effect of negative reciprocity on effort and profits is negative if we assume that it is triggered by a fixed wage that is below a given reference wage. For example, assume that we give up the limited liability constraint and set this reference wage to zero. Then, the optimal spot contract still takes the form of either a bonus or a reciprocity contract, only the threshold $\overline{\eta}$ is different than in our main analysis.[16] Negative reciprocity negatively affects the bonus contract, because the fixed wage is generally negative in this case in order to extract the agent's rent. The reciprocity contract and the relational contract contain a positive fixed wage and consequently are not affected by negative reciprocity – hence the impact of the latter on effort and profits is negative on average. This negative effect is the dominant presumption in the literature (for a theoretical investigation, see Netzer and Schmutzler, 2014), and can also be found in the data (Dohmen et al., 2009).

On the other hand, negative reciprocity might also *positively* affect a relational contract. Right now, η drops to zero after a deviation by the principal. Instead, we might assume that η does not necessarily drop to zero, but that the reduction is a function of the agent's negative reciprocity. If a larger degree of negative reciprocity implies a bigger reduction of η (following a deviation by the principal), it would actually allow to implement a higher effort level (and increase profits) – because the principal's outside

[16] Omitting the limited liability constraint rather increases $\overline{\eta}$ (because using a bonus contract does not make it necessary to grant the agent a rent anymore), whereas negative reciprocity reduces $\overline{\eta}$.

option would go down accordingly, and consequently increase her commitment in the relational contract. This positive effect would now be more pronounced for *younger* employees where incentives are mostly provided via relational contracts.

We think that a thorough investigation of the effect of negative reciprocity on an optimal dynamic incentive scheme deserves more attention, but leave it for future research.

Chapter 3

Individual Differences and Contribution Sequences in Threshold Public Goods*

3.1 Introduction

As "[s]ome of the most fundamental questions about the organization of society center around issues raised by the presence of public goods" (Ledyard, 1995, p.112), public goods games are among the most studied topics in experimental economics in the lab as well as in the field. Indeed, the plethora of studies in economics and neighboring disciplines like sociology, political science, or psychology reflects the variety of situations characterized by the social dilemma inherent in public goods provision: the free-rider problem. The reason for the persistent interest stems from the fact that abstract features of the game serve as a metaphor for many economic situations spanning from defense, public infrastructure, or health to more specific problems in organizational economics like intellectual property rights or team production (Ledyard, 1995).

The leading paradigm in most of this literature has been to focus on final outcomes, which is why the public goods situation is mostly modeled as a simultaneous-move game. Within this general framework, a variety of adaptions have been suggested to solve the free-rider problem. A non-exhaustive list of those variations includes partner design, communication, monetary and non-monetary punishment, or heterogeneous

* This paper is based on joint work with Michael Schüßler and Daniel Mühlbauer.

endowments (Chaudhuri, 2011; Zelmer, 2003). While these approaches often alleviate the free-rider problem and facilitate cooperation, some of the most interesting public goods in social and organizational reality have a different temporal structure: processes like team work in organizations or fund raisers have a real-time aspect instead of simultaneous contributions.

In arguing that it is crucial to focus on the process of public goods provision instead of final outcomes, our study contributes to several strands of literature. First, in ameliorating the free-rider problem, some experimental public goods studies have started to deviate from the simultaneous-move setup and introduced the real-time protocol. However, relatively few studies employed this mechanism, reporting mixed results – while public goods implementation is enhanced in some studies, others fail to find this effect (Dorsey, 1992; Goren et al., 2003; Friedman and Oprea, 2012; Oprea et al., 2014).

Second, there has been a growing interest in looking at the effects of (endogenous) leadership (Dannenberg, 2015; Potters et al., 2005, 2007; Rivas and Sutter, 2011; Sutter et al., 2010). These studies are concerned with exploring if and why individuals volunteer to make the first move (i.e., leading by example without formal authority), finding that while having individuals volunteering to go first increases implementation success, few individuals are willing to do so. However, little is known about the theoretical rationale to explain an individual's position in the resulting contribution sequence. Our point of departure is that we are explicitly interested in those sequences. Thus, we contribute to the literature on the real-time protocol by examining the effect of visible contributions while not only focusing on the ultimate result, that is, the implementation of the public good, but also on the process leading to it. We contribute to the literature on endogenous leadership by exploring who volunteers to go first in a richer setup where contribution sequences are endogenous.

Third, while economic theory is silent concerning those contribution sequences, the focus on the role of heterogeneous tendencies to participate in collective action has long been discussed in sociological theories of collective behavior (Oliver, 1993). For our purpose, we draw on a specific model, referred to as threshold theory (Granovetter, 1978; Granovetter and Soong, 1983). This theory delivers utility-based arguments

why some individuals tend towards earlier and others towards later decisions. In fact, Braun (1994) attributes individual threshold heterogeneity to differences in behavioral determinants. While he mainly focuses on the benefits, costs, and network properties of the decision context, individual differences are neglected (Braun, 1995; Watts and Dodds, 2011). The present study aims at closing this gap between sociological theories of collective behavior and economic experiments by investigating the role of individual differences in endogenous contribution sequences in a real-time public goods context.

Fourth and related, we also add to the literature investigating how individual differences influence behavior in public goods situations. More specifically, several studies have investigated how heterogeneity in personality (Hilbig et al., 2014, 2012; Zhao and Smillie, 2015), distributional preferences (Balliet et al., 2009; Dijkstra, 2013; Murphy and Ackermann, 2014), or risk aversion (Fung et al., 2012; Teyssier, 2012) affect contribution decisions. We build on and extend these studies by measuring participants' personality traits and distributional preferences and relating them to individual heterogeneity in contribution behavior in the real-time public goods game.

Thus, we employ a simple provision-point mechanism in the laboratory and use the real-time protocol; that is, participants do not make decisions simultaneously, but are able to freely decide when to make their choice within a given time frame. This way, we are able to examine the effect of having information about others' actions on the one hand while being able to identify individual differences in behavior in these real-time public goods situations and relate them to personality traits and economic preferences on the other hand. Indeed, we are able to show that if we provide information about others' behavior in real time, implementation of the public good is facilitated, at least when the target is sufficiently difficult to reach. In addition, we find endogenous contribution sequences, i.e., individuals reacting to real-time information about their fellow group members. While distributional preferences alone explain participation decisions, the sequencing of contributions is influenced by a broader array of individual differences.

The rest of the paper is structured as follows. First, we discuss related literature by elaborating on how individual timing and individual differences are intertwined. Second, we explain both the laboratory design and procedure. After having presented our main

experimental results, we discuss their relevance in light of potential limitations. Finally, we conclude.

3.2 Related Literature

The conflict between individual and collective rationality has drawn the attention of various research disciplines. Sociologists have long been interested in the formation of collective behavior, such as social movements (Smelser, 1963; Udéhn, 1993), psychologists have been studying helping and organizational citizenship behavior (Organ, 1994; Smith et al., 1983) as well as various forms of mixed-motive social dilemmas (Dawes, 1980; Komorita and Parks, 1995; Lange et al., 2013), and economists have been examining different versions of inter-individual coordination or behavioral cascades (Anderson and Holt, 1997; Bikhchandani et al., 1998; Chaudhuri, 2011; Ledyard, 1995). Despite important differences, these approaches all tackle the problem of interpersonal coordination and the conflict between individual and collective rationality (Dawes and Messick, 2000; Kollock, 1998).

Apart from the more general examples mentioned above, specific applications to organizational reality have been discussed, arguing that "in fostering cooperation among employees, managers usually face a public good dilemma" (Bridoux et al., 2011, p.711). Relevant phenomena of organizational behavior and work group productivity, such as strategic initiatives (Lechner and Kreutzer, 2011) or organizational task forces (Gersick, 1988), closely resemble those social dilemmas for three reasons. First, strategic initiatives and organizational task forces evolve along a collective sequence made up of individual contribution decisions and both require concerted voluntary efforts of the individuals involved, while not allowing for perfect observability of individual behaviors without incurring high monitoring costs (Goren et al., 2003). Second and related, due to their highly interdependent nature, it may be hard to exclude any individual from enjoying the (long-term) benefits of the public good, while the short-term costs are only incurred by the cooperating individuals (Bridoux et al., 2011; Lechner and Kreutzer, 2011). Third, their success typically increases with the investments of time and effort by their members and they are of temporary nature, which means that they typically

have to be successfully accomplished until a certain deadline, or else fail.

In the following paragraphs we will develop two distinct, yet related sets of hypotheses. First, we will discern how institutional changes influence public goods provision. Second, we derive a set of hypotheses to explain endogenous contribution sequences.

3.2.1 The Influence of Institutional Changes on Public Goods Provision

Abstract versions of these phenomena described above have been implemented in the experimental laboratory. In standard experimental public goods contexts, the voluntary contribution mechanism (VCM) has emerged as the dominant way of representing public goods provision in experimental studies in economics and psychology (Chaudhuri, 2011; Ledyard, 1995; Zelmer, 2003). Under the VCM, all individuals voluntarily decide to either invest resources in the public account or to keep them in their private account. Afterwards, all investments to the public account are multiplied by a given factor and distributed among all individuals, including the defectors. Hence, in this simple game, keeping all resources in the private account is the individually dominant strategy, while investing everything in the public account is socially optimal. It has been shown that, on average, contribution rates under the VCM are between 40% and 60% of the social optimum, while contributions tend to decrease over repeated periods of the game (Chaudhuri, 2011; Goren et al., 2003).

As already pointed out in the previous section, to alleviate this prevalent free-rider problem, several institutional solutions have been proposed. While several variants of the standard VCM have been discussed in that regard, the provision point mechanism (PPM) is of particular interest for studying public goods because it is more comparable to organizational reality, keeping in mind the above-described strategic initiatives and organizational task forces. We argue that using the PPM enhances both mundane and experimental realism because it differs from the VCM in two notable features (Abele et al., 2010; Croson and Marks, 1998; Marks and Croson, 1999).[1] First, the public good

[1] While the former aims at resembling the real world in an experimental investigation, the latter is concerned with how well the experimental situation captures the theoretical constructs under scrutiny (Colquitt, 2008). Most often critique of experimental settings is targeted at low mundane realism, while experimental realism should be the focus of attention (Highhouse, 2009).

is not provided unless a given threshold t of contributions is reached (i.e., for a project to be conducted, a certain number of people need to participate). Second, while under the VCM contributions to the public account are typically continuously divisible, the PPM mostly employs all-or-nothing contributions (i.e., team members either participate in a proposed project or not). More specifically, under the PPM, participants play in groups of n players where every player has an endowment of E units. Subsequently, the players choose a contribution level $c_i \in \{0; 1\}$. If t or more players choose to invest their endowment in the public account, the public good of value G is provided and every group member receives a reward of $\frac{G}{n}$ units. Additional contributions do not further increase the value G. If t is not reached, the public good is not provided and the contributors lose their investment to the public account. Non-contributors keep their initial endowment in both cases. Thus, in adding a contribution target to be reached, a coordination problem arises. Typically, all parameters of the game (n, t, G, and c) are common knowledge.

However, there are also two notable differences from a game theoretic perspective. First, in comparison to the VCM, there is not just a single Nash equilibrium for contributing nothing, but additional equilibria for exactly t players contributing exist. Second, while in the continuous form everyone contributing is the Pareto efficient solution, the second Nash equilibrium where everybody is keeping their endowment is not. In the step-level form, any coordination solution that reaches the provision point is Pareto efficient (Abele et al., 2010). Consequently, under the PPM, contribution behavior has been shown to be much more rugged in comparison to the VCM: while overall efficiency and total contributions are increased, it is found to be more susceptible to changes in induced value (Rondeau et al., 2005). In addition, contributions do not always decrease over repeated periods of the game (Abele et al., 2010).

One important driver of contribution decisions to public goods is the value of the public good relative to that of the forgone private good, which has been formalized as marginal per-capita return and demonstrated to influence the provision of linear public goods. A comparable feature of threshold games is the step return. Empirical results show that comparing the step returns between games yields comparable results to uti-

lizing the marginal per capita return under the VCM: a higher step return encourages contributions (Croson and Marks, 2000). As the step return increases in t, we propose the following hypothesis:

Hypothesis 3.1. *Implementation rates are lower when the provision point is increased.*

Moreover, while most public goods experiments have been conducted under the simultaneous protocol of play where decisions are made simultaneously, anonymously and in private (Chaudhuri, 2011), this does not resemble strategic initiatives and organizational task forces very well. This because in organizations, individuals normally possess at least partial information about the decisions of other group members and consider this information before making their own decisions. As an alternative, the real-time protocol is better suited to capture the inherent decision sequence of endogenous individual timing present in organizations. Until now, comparably few studies employed this mechanism under the VCM and the PPM. Using the real-time protocol, the differences between coordination and cooperation are unclear and multiple equilibria exist, which raises the question whether participants are able to coordinate efficiently or not (Abele et al., 2010; Oprea et al., 2014). While most studies find enhanced implementation rates, others fail to find this effect (Berninghaus and Ehrhart, 2001; Dorsey, 1992; Goren et al., 2003, 2004; Kurzban et al., 2001). Those empirical inconsistencies are due to two reasons: First, the results stem from different games, thereby inhibiting comparability. Second, while most studies on coordination use revocable investments to the public account (e.g., Leng et al., 2016), we employ irrevocable commitments to avoid cheap talk (Farrell and Rabin, 1996). Theoretically, it has been argued that continuous games approximate an array of discrete games with the grid length going to zero (Simon and Stinchcombe, 1989). Thus, this form of repeated interaction should increase cooperation.[2] Therefore, we posit the following hypothesis:

[2] Theoretical predictions are difficult at this point because extant reasoning in economics does not allow for clear-cut hypotheses. However, following social interdependence theory (Abele et al., 2010; Kelley and Thibaut, 1978), one can argue that if participants are missing any form of information used to infer the behavior of others, coordination failure is more likely (Puranam et al., 2012).

Hypothesis 3.2. *Implementation rates are lower when individuals have no information about others' behaviors.*

However, behavior is not only driven by institutional arrangements or incentives, but also by individual heterogeneity. Albeit being unified in these central building blocks underlying behavior formation, economists and psychologists differ with regard to how this heterogeneity is conceptualized. While psychologists use the different facets of personality (Rabin, 1998; Rustichini, 2009), economists rely on using heterogeneity in stable and exogenous preferences like risk or time, and have long been skeptical about the usefulness of those personality dimensions. Albeit they have long questioned their stability, extant research suggests that these qualms are uncalled-for (Almlund et al., 2011). Indeed, it has been shown that personality could be a worthwhile adjunct to experimental measures of economic preferences in increasing predictive power (Rustichini et al., 2012) and are thus complementary (Becker et al., 2012). Adding to that literature, we argue that while the pertinent situation is a substantial influencing factor for behavior formation, another focus of our paper is on predicting contribution sequencing using individual differences, which will be elaborated on in the following.

3.2.2 Explaining Endogenous Contribution Sequences

There has been substantial research to predict public goods contributions under the VCM using individual differences, which can be either distal or proximal to overt behavior (Kanfer and Ackerman, 1989). While the former include broad factors, such as general personality structure (Hilbig et al., 2014; Hilbig and Zettler, 2009), the latter comprises more specific determinants of behavior like social value orientation (SVO) (Murphy et al., 2011) or risk preferences (Fung et al., 2012). As mentioned before, the game-theoretic predictions are fundamentally different under the PPM, yielding two types of pure-strategy Nash equilibria and thus a coordination problem to solve. Hence, the role of determining factors is less clear and evidence is scarcer. While using the real-time protocol of play should mitigate coordination failure, the availability of real-time information about contributors adds an additional strategic dimension to the game: moving early can be used to signal a cooperative social norm, but waiting

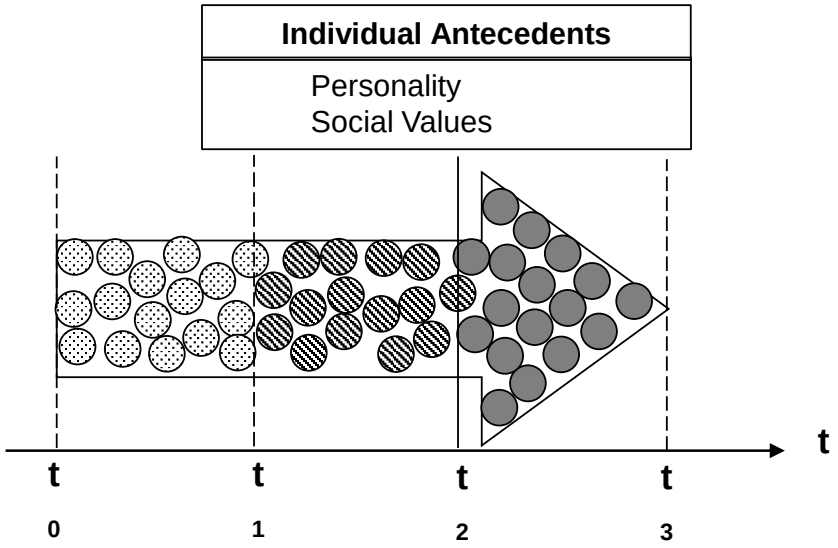
for others to move first and updating beliefs about the probability of implementation with and without one's own contribution on a rolling basis is also possible. In contrast, without real-time information about contributions available, subjects have no way of updating beliefs and less need to behave strategically, i.e., based on what others are doing. In essence, all group members base their decision on less information and thus take less time to choose whether to contribute or not. Therefore, we propose the following hypothesis:

Hypothesis 3.3. *The time to contribute is shorter when individuals have no information about others' decisions.*

In the case of symmetric agents, successful coordination is a stochastic process, as it is unknown ex ante which people are going to contribute and at which position in the behavioral sequence (Au and Budescu, 1999). While the role of asymmetric agents under the VCM has been studied extensively (e.g., Reuben and Riedl, 2013), economic and game-theoretic reasoning has mostly treated this kind of variance in coordination games as noise (Mäs and Nax, 2016). As such, differential effects of individual heterogeneity on coordination have only seldom been studied (e.g., Diekmann and Przepiorka, 2016). However, in sociology and psychology, there exists a broad range of theories specifically concerned with these processes. Specific variants of such models of collective behavior, such as bandwagon or hazard models, are often used to conduct research on conformity behavior, behavioral contagion, or on diffusion of innovation in a business setting (Centola and Macy, 2007; Oliver, 1993). In psychology and sociology, it is often assumed that individuals prefer social conformity, equality, or identity affirmation (Ashforth and Mael, 1989; Simon et al., 2008; Simpson, 2006), while in economics and economic sociology, the underlying assumption is that information is derived from behavior of (relevant) others (Anderson and Holt, 1997; Bikhchandani et al., 1998). While those accounts differ with regard to the driving force behind social interdependence, they are unified by their basic tenet of proposing that the uptake of beliefs or the adoption of an innovation depends on other people's behavior. Therefore, we posit the following hypothesis:

Hypothesis 3.4. *There is a cascading process in contribution decisions.*

Figure 3.1: *Individual differences, behavioral thresholds, and contribution decisions*



A specific specimen of the broad family of theories described above is the threshold model of collective behavior (Granovetter, 1978; Granovetter and Soong, 1983). Here, individual heterogeneity in contribution timing is assumed to manifest as a behavioral threshold which is defined as "the number or proportion of others who must make one decision before a given actor does" (Granovetter, 1978, p.1420). As shown in Figure 3.1, some individuals actively contribute very early in the process (i.e., instigators, dotted circles), others may wait and see whether or not any collective movement begins to form (i.e., moderates, striped circles), while again others may need to see a lot of other individuals acting before contributing (i.e., reluctant, filled circles). Broadly speaking, an individual's behavioral threshold represents that person's behavioral tendency to participate in a collective action.

While the original model suggests that thresholds are solely based on cost-benefit calculations and availability of information, it has been pointed out that individual differences also drive threshold formation. This way, individuals are assumed to differ in their behavioral tendency to cooperate because of heterogeneous personality factors, diverse ideologies or beliefs, varying economic motives as well as differing susceptibility to conformism. In the context of this paper, we focus on personality factors, as various studies have related personality to public goods contributions as well as to a broad range

of work- and non-work-related attitudes and behaviors (cf. Ozer and Benet-Martînez, 2006). While the so called five-factor model has emerged as the most widely used model of personality (Barrick and Mount, 2012), the HEXACO model is more suitable for research on social dilemmas (Hilbig and Zettler, 2009; Ross et al., 2003) as it clearly distinguishes the domains of honesty and agreeableness. This reflects an intensive debate in personality psychology whether the five-factor model should be complemented with a sixth factor capturing honesty or integrity (Hough et al., 2015). While this general issue has not been fully resolved yet, this distinction is important for the present study because it enables disentanglement of active (i.e., honesty in HEXACO) and reactive cooperation (i.e., agreeableness in HEXACO), both captured by agreeableness in the five-factor model (Hilbig et al., 2014). Individuals with a tendency for active cooperation are likely to contribute earlier, whereas individuals with a tendency for reactive cooperation are likely to contribute later because of their rather responsive nature.

Honesty. Honesty "represents the tendency to be fair and genuine with others, in the sense of cooperating with others" (Ashton and Lee, 2007, p.156). Empirical evidence shows that honesty is positively related to prosocial behavior, for instance in dictator (Hilbig et al., 2014; Hilbig and Zettler, 2009) and public goods games (Hilbig et al., 2012). More specifically, honest individuals are prone to cooperate early when the decision timing is left to their choice (Hilbig et al., 2013b). Therefore, we propose the following hypothesis:

Hypothesis 3.5. *Honest individuals contribute earlier in the contribution sequence.*

Emotionality. Emotionality is typically described as being "vulnerable, sensitive and anxious", while individuals scoring low on this trait are rather "fearless, tough, independent, and unemotional" (Ashton et al., 2014, p.140). Thus, it is linked to empathy and attachment (Hilbig et al., 2013a). As empirical investigations have found a negative link between emotionality and contributions in a public goods context (Clark et al., 2014; Hilbig et al., 2012), we propose:

Hypothesis 3.6. *Emotional individuals contribute later in the contribution sequence.*

Extraversion. Also belonging to the proactivity domain, extraverted individuals show proactive behaviors and have been described as being outgoing, lively, sociable, talkative, cheerful and active (Ashton and Lee, 2007). Empirical investigations yielded mixed results regarding extraversion and behavior in social dilemmas, which has been attributed to divergent theoretical linkages of social dominance and sociability. While the former is likely to foster free riding (Zhao and Smillie, 2015), the latter suggests that extraversion enhances contributions to the public good (Clark et al., 2014; Hilbig et al., 2012). As extraversion has also been found to increase proactive behavior (Onyemah, 2008) and organizational citizenship behavior (OCB) (Bourdage et al., 2012), we propose:

Hypothesis 3.7. *Extraverted individuals contribute earlier in the contribution sequence.*

Agreeableness. In the HEXACO model, agreeableness is linked to traits like tolerance and forgiveness even in the face of being exploited (Ashton and Lee, 2007), as well as flexibility and patience (Ashton et al., 2014). While agreeable individuals are also inclined to cooperate, it is a rather reactive form of cooperation (Hilbig et al., 2012, 2013a; Zhao and Smillie, 2015). Empirical research supports this differentiation, since agreeableness was positively related to reactive cooperative behavior in ultimatum games and only weakly related to active cooperation (Hilbig et al., 2013a). Therefore, we propose the following hypothesis:

Hypothesis 3.8. *Agreeable individuals contribute later in the contribution sequence.*

Conscientiousness. Conscientiousness is seen as "[...] socially prescribed impulse control that facilitates [...] thinking before acting" (John et al., 2008, p.120). Thus, conscientious individuals often show traits, such as thoughtfulness, deliberateness, and planning (John et al., 2008; McCrae and Costa, 1987). Tan and Tan (2008) report a significant negative (positive) relation between conscientiousness and social loafing (OCB). While conscientiousness has also been positively linked to individual contributions in public goods games (Clark et al., 2014; Hilbig et al., 2012), others have found no correlations to cooperative intentions (Ross et al., 2003). Taken together, we propose the following hypothesis:

Hypothesis 3.9. *Conscientious individuals contribute later in the contribution sequence.*

Openness. As a trait, openness to experience depicts a person's artistic or aesthetic character, her curiosity, or creativity (Ashton and Lee, 2009). Due to the inherent nature of the measurement of openness in the HEXACO model, it seems not plausible to form a specific hypothesis. While we include openness in our analysis, this is rather exploratory.

Social Value Orientation. Apart from distal personality traits, behavioral thresholds should also be influenced by more proximal individual preferences. Under the general label of social preferences, heterogeneity in distributional preferences has been studied in a variety of settings in economics (Fehr and Schmidt, 2006). Psychologists, however, have measured these preferences as SVO or closely related constructs (e.g., prosocial values, other orientation, or collective orientation). Building on decomposed dictator games, SVO offers a more fine-grained approach and has been shown to explain heterogeneity in individual behavior in public goods games (Balliet et al., 2009; Murphy and Ackermann, 2014). SVO measures the relative importance of own and others' payoffs (McClintock and Allison, 1989). Typically, specific types (i.e., altruistic, cooperative, competitive, and individualistic) are identified, which are in turn used to explain individual heterogeneity in behavior formation. This way, it has been shown that prosocials (i.e., cooperative or altruistic types) are more likely to cooperate in social dilemmas than proselfs (i.e., individualistic or competitive types) (Abele et al., 2010; Balliet et al., 2009; Bogaert et al., 2008). Albeit SVO has been shown to have a smaller effect in step-level public goods compared to continuous public goods (Abele et al., 2010), we propose the following hypothesis:

Hypothesis 3.10. *Prosocial individuals contribute earlier in the contribution sequence.*

Having described the hypotheses, we now turn to the laboratory design and procedures before we report the results of our study.

3.3 Experimental Design

The experimental procedure consists of four parts: Upon arriving at the laboratory, participants were randomly assigned to cubicles. Then, having been introduced to the general procedures of the experiment, the public goods game was explained in detail by using neutrally framed instructions.[3] After that, participants were asked to answer four control questions to ensure common understanding of the proceedings of the public goods game. Arising questions were dealt with by the experimenter in private. The next stage only started after all 24 participants in a session had managed to answer all control questions correctly. After that, ten rounds of a public goods game using the PPM under the real-time protocol were played. In our specific case, participants faced a PPM as described above with $n = 8$, $G = 3500$ tokens, and $c = 100$ tokens. After every round, groups were randomly rematched to mitigate reputation effects. To make strategic behavior and end-of-period effects less pronounced, every round lasted between 60 and 90 seconds with equal probability of termination in between, thus leaving it unclear ex ante how long the round would exactly be. For similar reasons, subjects were not told in advance that there would be exactly ten rounds. In each round, investment decisions of each player in a group were made known to all others in real time. Additionally, all subjects were informed that one round would be randomly selected to become payoff-relevant at the end of the experiment.

To explore the causal effect of institutional changes on our variables of interest, we used a 2x2 design altering the provision point, the level of information provided to participants, and a combination of both. While the former could be either low or high (i.e., $t = 5$ vs. $t = 7$), the latter could be either full or none (i.e., full but anonymized feedback vs. no feedback).[4] This yields the different combinations depicted in Table 3.1 to which subjects were randomly assigned.[5]

After having played ten rounds of the public goods game, two other behavioral measures

[3] A translated version of the full instructions can be found in Appendix C.3.

[4] We decided to use extreme cases instead of subtler changes to the institutional environment to receive maximum treatment effects (List, 2011).

[5] Due to administrative issues, we conducted twice as many sessions of the *Baseline*.

Table 3.1: *Overview of the treatment conditions*

		Information	
		Full	None
Provision Point	5	Baseline	No Info
	7	High Provision Point	Combined

were employed. Specifically, participants' SVO and risk attitudes were elicited. For the former we used the slider-measure approach (Murphy et al., 2011) where each participant has to indicate distributional preferences by allocating a given endowment between herself and another person. In comparison to older measures of the construct, this approach does not only yield behavioral types, but the SVO angle, a much more fine-grained measure on a ratio scale. The participants faced 15 of such decomposed games yielding a position on a circular plane of all possible allocation-based positions and defining the SVO (see Murphy et al., 2011, for a detailed description).

Following the approach by Holt and Laury (2002), individual risk aversion was assessed by having the participants choose between ten pairs of investment options A and B, one having a greater payoff spread than the other. Basically, the measure allows to identify a person's risk aversion by observing the point at which an individual switches to choose the riskier option B over option A. For example, a person is classified as risk-neutral when choosing option A for the first four times and then switching to option B, because then the expected payoff difference between option A and B turns negative for the first time. In general, the earlier an individual chooses the riskier option B over option A, the more risk seeking this person is. Again, participants were informed ex ante that one decision would be randomly selected to determine payoff from this part.

As a last part, participants were asked to fill in several questionnaires. First, the German version of the HEXACO-60 personality inventory (Ashton and Lee, 2009; Moshagen et al., 2014) was used. Additionally, besides risk aversion, several control variables were elicited as collective behavior may also be influenced by other individual characteristics. Specifically, the controls included participants' gender and age, their study major, math

grade as crude proxy for general mental ability, and measures for generalized trust as well as positive and negative reciprocity norms. While the former has been extensively studied as an antecedent of cooperativeness (Ostrom and Walker, 2003), the latter are also frequently mentioned, since they account for rather reactive forms of behavior (Chaudhuri, 2011). For example, generalized forms of retaliation for past behavior of anonymous others may also explain an individual's behavior over time. These measures deploy additional control for non-specific retaliation caused by general anger concerning defection of others. For reasons of parsimony, these controls were measured using three-item scales from the German Socioeconomic Panel (SOEP) (Caliendo et al., 2012; Dohmen et al., 2008). Then, individual payoffs from all parts were determined and all subjects received their payment in private.

3.4 Results

The experiment was conducted at MELESSA, the experimental laboratory of the University of Munich, in September and October 2015. 360 participants recruited using Orsee (Greiner, 2015) took part in 15 sessions which lasted about 75 minutes each. The average age of participants was 25.7 years; the youngest participant was 18 and the oldest 62 years old. The majority of participants was female (60%) and students (90%); 31% pursued a major in business and economics. All sessions were computerized using the software z-Tree (Fischbacher, 2007). During the experiment, participants could earn experimental points that were exchanged into € at the end. On average, subjects earned 13.71 €, including a show-up fee of 4 €.

In the next section, we present our main results. Our findings are structured as follows: First, we present descriptive results. We do so by describing our data set and investigating group-level outcomes before scrutinizing individual-level decisions, which are the foundation for the macro-level outcomes described before. After that, we turn to test our hypotheses. We start by looking at treatment effects of the institutional changes described above. Then, we delineate whether we actually observe cascading in our data as predicted and test which antecedents may be used to predict the position in the behavioral sequence.

3.4.1 Descriptive Results

Information about the descriptive statistics are displayed in Table 3.2, while individual differences and intercorrelations across treatments are depicted in Table C.1 and Table C.2 in Appendix C.1, respectively. By and large, the patterns we find for our variables correspond to those in the literature.

Personality was measured on a five-point scale with 60 items (i.e., six dimensions, ten items each) of the established and validated HEXACO questionnaire (Moshagen et al., 2014) to assess the extent to which respondents were honest (α = .76), emotionally stable (α = .76), extraverted (α = .80), agreeable (α = .74), conscientious (α = .77), and open to experience (α = .69). While there is some variation in internal consistency, all alpha values are comparable to those reported by Moshagen et al. (2014).[6] Figure C.1 in Appendix C.1 depicts the distribution of scores on the six HEXACO dimensions in our sample.

Concerning SVOs, our results and distributions are comparable to those reported by Murphy et al. (2011), who also indicated clustering around the individualistic type and no representation of purely altruistic individuals. See Figure C.2 in Appendix C.1 for the distribution of angles in our sample.

Regarding our control variables reciprocity, trust and risk aversion, patterns are also quite similar to the literature. While positive reciprocity (α = .61) only has questionable internal consistency, negative reciprocity (α = .77) and trust (α = .74) show acceptable levels of internal consistency. Compared to Caliendo et al. (2012), who also report questionable internal consistencies for negative and positive reciprocity and good internal consistency for trust, our results are in line considering the differences in sample size.[7] As the internal consistency depends on the number of items used and considering the fact that we include reciprocity and trust only as control variables, this is a caveat, but should not limit the interpretation of our results in general.

[6] Furthermore, we conducted a confirmatory factor analysis to evaluate the underlying factor structure of the HEXACO model. The results are described in Appendix C.2.

[7] A confirmatory factor analysis was conducted to evaluate the factor structure (RMSEA = .049, SRMR = .038, CFI = .987); model fit is acceptable (Bentler and Bonett, 1980; Brown, 2006).

Table 3.2: *Descriptive statistics of the explanatory variables*

	Mean	SD	Min	Max	Alpha
Honesty	3.30	0.73	1	5	0.76
Emotionality	3.17	0.68	1	5	0.76
Extraversion	3.47	0.65	1	5	0.80
Agreeableness	3.13	0.63	1	5	0.74
Conscientiousness	3.74	0.61	1	5	0.77
Openness	3.64	0.62	2	5	0.69
SVO angle	17.18	12.73	-16	46	–
Risk aversion	6.38	1.52	0	10	–
Trust	3.20	0.90	1	5	0.74
Negative reciprocity	2.55	0.92	1	5	0.77
Positive reciprocity	4.43	0.57	1	5	0.61

Regarding risk preferences, the average participant in our sample is risk averse, while 40 participants showed either an inconsistent choice pattern (i.e., more than one switching point) and/or failed to choose the higher monetary outcome when confronted with a choice between 200 points and 385 points with probability one, respectively. As excluding these subjects from our analyses does not change the results substantially and risk preferences are a control rather than a variable of interest, we decided to include all subject in our analysis.

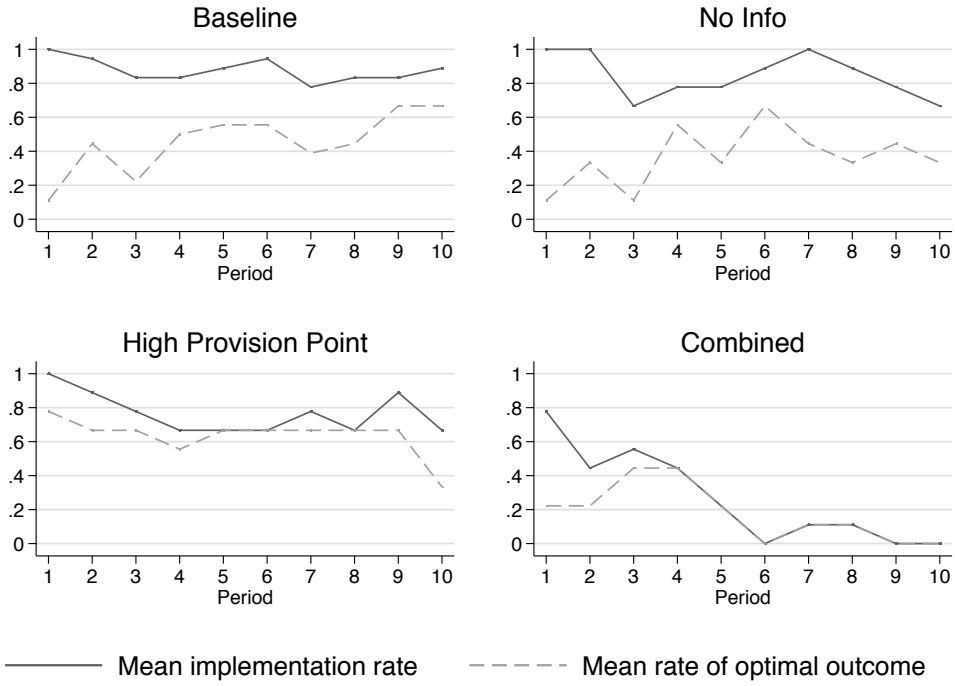
Overall, the independent variables measured in our sample resemble those established in the literature, encouraging use of all of the variables for further analysis.

Group-Level Outcomes

Before explicitly testing our hypotheses, we give a general description of the group-level outcomes and how they are shaped by individual contribution decisions. We do so by describing the implementation rates as well as the mean rate of optimal outcome across all ten periods.

Figure 3.2 displays the rate of cases out of all group observations in each of the ten

Figure 3.2: *Average implementation rates by treatment*



Notes: Figure depicts average rates of successful (public good provided) and optimal implementation (threshold reached exactly) by period. Calculations based on 450 groups.

periods in which the public good was successfully implemented (solid line) across all four treatments (i.e., at least five (seven) out of eight group members contributed their endowment). It also contains the rate of cases in which the theoretically optimal outcome (dashed line) was reached (i.e., exactly five (seven) out of eight group members contributed). Note that due to the way these two measures are constructed, the mean rate of optimal outcome is a proportion of the total mean implementation rate, which also includes the cases in which six, seven or even all eight group members contributed. These cases lead to the implementation of the public good as well, but to a lower total payoff because the additional contributions are lost. Thus, the upper limit for the mean rate of optimal outcome is given by the mean implementation rate.

Overall, it is evident that in our setting, the mean implementation rate is quite high. This rate of successful public good provision is substantially larger than in other experiments using the PPM (e.g., Croson and Marks, 2000). This indicates that the real-time protocol sustains cooperativeness in socially complex settings. In general, implementation

does not seem to become more unlikely over the course of the experiment. Moreover, most of the participants seem determined to contribute in the first period, resulting in implementation in all groups, but also in a very low proportion of groups with exactly five (seven) individuals contributing (i.e., the optimal solution). In the *Baseline*, implementation rates stay high at all times, while the number of groups that manage to coordinate efficiently increases. Judging by the graphs alone, differences in availability of information about contribution decisions of others seem to have little influence on the groups' ability to effectively coordinate for collective success as long as the step return remains constant (*Baseline* $\rightarrow$ *No Info*). In comparison, only manipulating the step return seems to affect collective success more strongly (*Baseline* $\rightarrow$ *High Provision Point*), as these groups' average implementation rate across all rounds is lower. When stakes are high for the groups, unavailability of information has a strong effect on their ability to effectively coordinate for collective success (*High Provision Point* $\rightarrow$ *Combined*). Moreover, the groups' coordination ability drops dramatically when simultaneously confronted with a lower step return and unavailability of information (*Baseline* $\rightarrow$ *Combined*). Although there is some variation over periods, the only treatment with a clear decline over time is *Combined*.

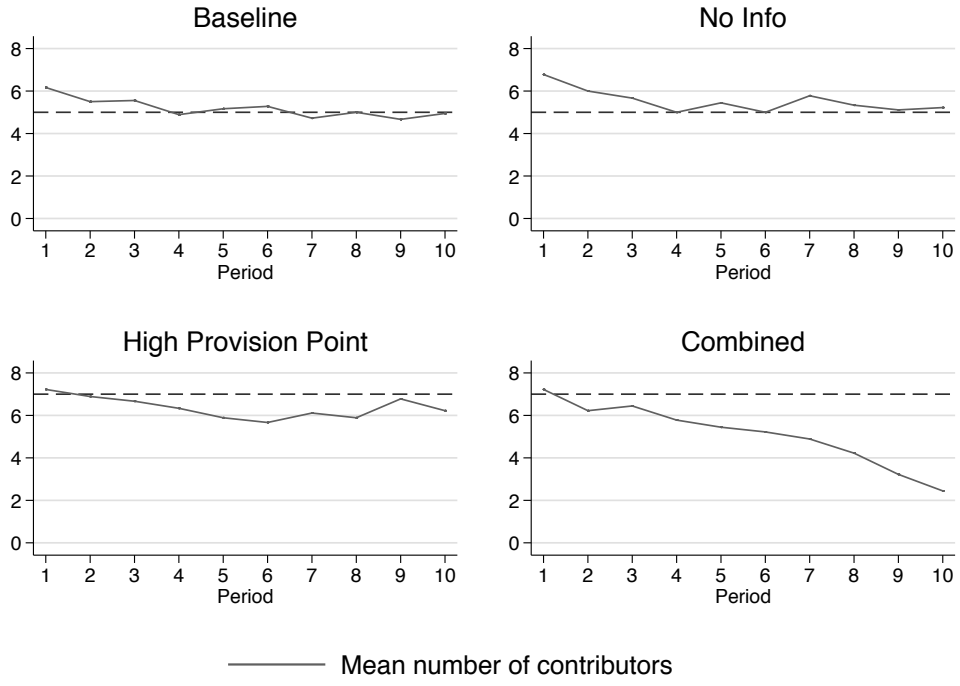
While there is more variation in *No Info* compared to the *Baseline*, there are still some groups in the *Baseline* with excess contributions although subjects should know that they burn money. To check whether subjects were behaving irrationally, we take a closer look at seemingly irrational contributions in the *Baseline* and *High Provision Point* treatments, where subjects receive real-time information about whether there are enough contributions for the provision point to be reached. Out of 1510 observations in the two relevant treatments, 84 are seemingly irrational (i.e., it was visible there were already enough contributions for the public good to be implemented), which is a percentage of 5.5%. Of these contributions, 56 were made more than one second after the provision point was reached. Thus, only 3.7% of contribution decisions were not in accordance with standard assumptions. In light of the fact that this is only a very small fraction of choices and that one could possibly rationalize these choices when non-standard preferences like inequity aversion are considered, we use all observations for our analysis.

Individual Contributions

Figure 3.3 depicts the average number of contributors for each treatment across the ten periods. Notably, albeit the number of contributors fluctuates, it is quite high in general. While the average number of contributors slightly decreases over time across all treatments, this does not automatically indicate a negative trend of collective success. For instance, in the *Baseline* treatment, the high average number of contributors in the beginning (rounds one to three) drops to converge around the optimal rate of contributors indicated by the dotted line. In addition, the graph provides some suggestive evidence that only in the *Combined* treatment the number of contributors strongly decreases below the provision point over time.

Having given a first impression of what our data look like, we now turn to test our hypotheses in the next paragraph.

Figure 3.3: *Average number of contributors by treatment*



Notes: Figure depicts average number of contributions made by period. Dashed line indicates threshold. Calculations based on 360 subjects.

3.4.2 Test of Hypotheses

Institutional Changes

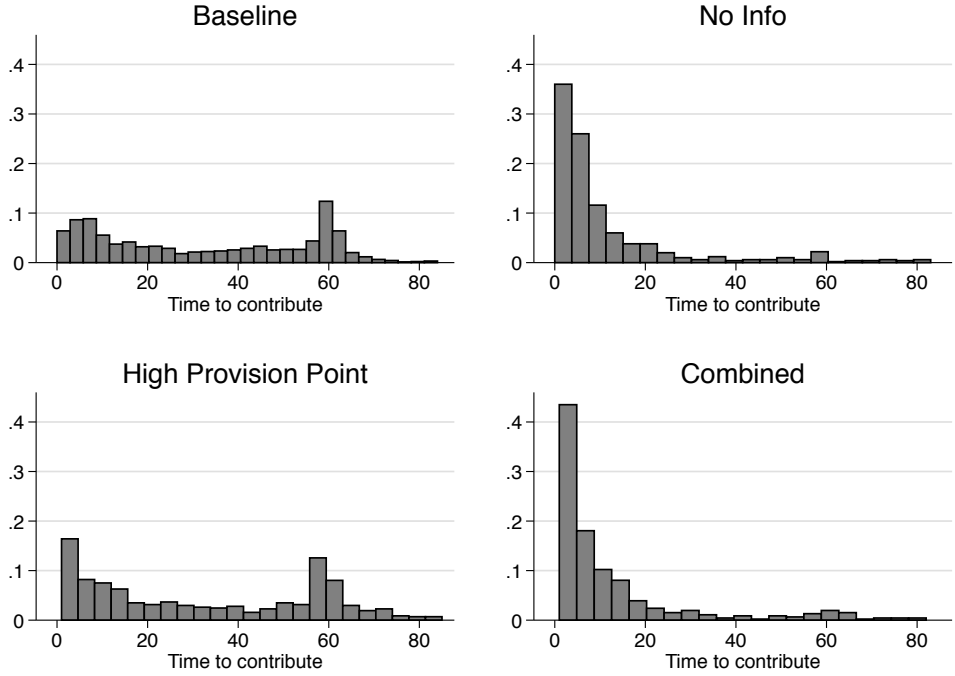
We now investigate differences in implementation of the public good across the different treatment conditions described above. While overall average implementation rates are quite high in all treatments except for *Combined* (*Baseline*: 88%; *No Info*: 84%; *High Provision Point*: 77%; *Combined*: 27%), the patterns described above and depicted in Figure 3.3 already provide suggestive evidence for Hypothesis 3.1 and Hypothesis 3.2.

The general impressions regarding the group-level success trajectories are supported by Fisher exact tests comparing the mean implementation rates of the public good across treatments. In particular, we find significant differences between all treatments except for the comparison between the *Baseline* and *No Info* treatments. In particular, the differences between *Combined* and all other conditions are highly significant ($p < 0.001$). The difference between the *Baseline* and the *High Provision Point* treatments is significant as well ($p = 0.022$), while the Fisher exact test fails to detect a significant difference between the *Baseline* and *No Info* ($p = 0.453$). In addition, we ran a probit regression predicting implementation success at the group level by provision point and availability of information. Details can be found in Table C.3 in Appendix C.1. The results mirror the Fisher exact tests: while a high provision point significantly decreases implementation success, the coefficient for lack of real-time information is negative, but insignificant. Thus, we find support for Hypothesis 3.1 but not for Hypothesis 3.2. However, we find a negative and significant interaction effect, indicating that the interplay of a high provision point and no information leads to an additional decrease in implementation success. Summing up, we find that lowering the step return leads to lower implementation rates, while removing information about what the other group members do does not. In addition, the combination of a lack of information and a high provision point also significantly reduces implementation rates.

Explaining Endogenous Contribution Sequences

In Hypothesis 3.3, we posited that removing information about group member's behavior would decrease individual decision timing. In principle, the coordination problem with the real-time protocol allows for the groups to exhibit different behavioral trajectories at the individual level, which then may or may not cause similar levels of collective success. Therefore, in Figure 3.4, we graphically explore individual decision timing.

Figure 3.4: *Average timing of contributions by treatment*



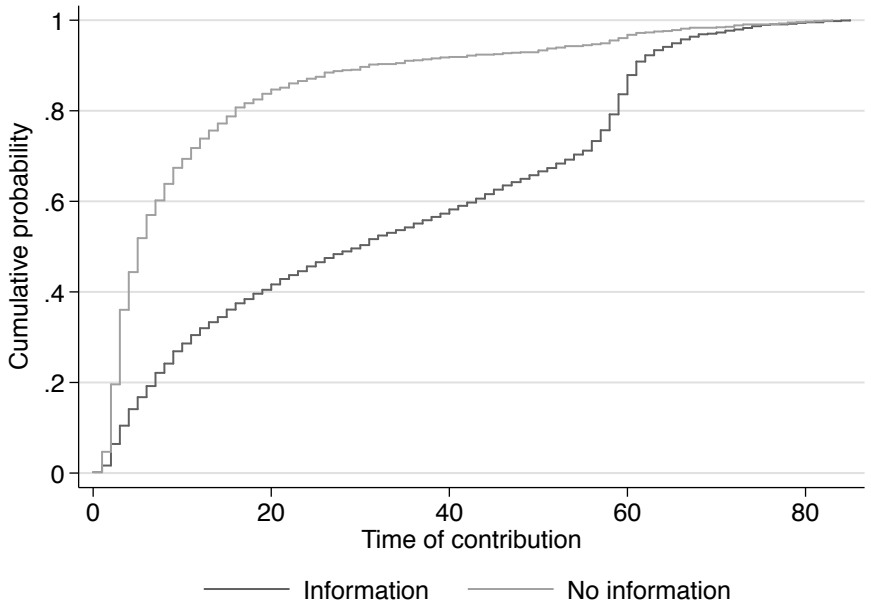
Notes: Figure depicts the distribution of time in the period when a contribution was made. Calculations based on 2,470 observations.

Some features are especially noteworthy when comparing the distribution of individual decision timing between the treatments with and without available information (i.e., *Baseline* and *High Provision Point* vs. *No Info* and *Combined*). In the treatments without information (*No Info* and *Combined*), almost all of the contributions happen within the first 20 seconds. In the other two treatments (*Baseline* and *High Provision Point*), the underlying pattern is different. In the latter case, a bimodal distribution of individual decision timing unfolds. Two moderate peaks distinguish very early contributions

and those made around 60 seconds, which marks the earliest time where each round could possibly be over according to our design. To explore this further, we compare the average time to contribute using Mann-Whitney U tests. While there is a significant difference between the treatments with and without information about others' behavior (*Baseline* vs. *No Info*: z = 2.32, p = 0.02; *High Provision Point* vs. *Combined*: z = 1.96, p = 0.05), this is not the case between the respective treatments with a low and high provision point (*Baseline* vs. *High Provision Point*: z = -0.26, p = 0.80; *No Info* vs. *Combined*: z = -0.66, p = 0.51).

In addition, to get at the difference in distributions between the contribution patterns, we compare the cumulative distributions of contribution times depending on whether information on others' decision is available or not. Figure 3.5 depicts the two functions. It is evident that the cumulative distribution function of times to contribute with information first-order stochastically dominates the cumulative distribution function without information. Summing up, we find support for Hypothesis 3.3.

Figure 3.5: *Cumulative distribution functions of contribution times*



Notes: Figure depicts cumulative distribution functions of contribution times for treatments with information (dark line) and without information (light line). Calculations based on 2,470 observations.

In Hypothesis 3.4, we claimed that there is non-random variation with regard to the sequence of decision timing. As explained above, the central notion of threshold models is that one's own contribution depends on the number of active others. For our setting, this implies that there is actually a contribution sequence observable rather than most contribution decisions being made more or less simultaneously right after the start of the period. Thus, for the threshold model to be applicable to our setting, we should observe sufficient variation in the timing of contributions within each round. To explore this, focus on the left side of Figure 3.4. As already stated above, contributions are spread over the whole period, with two peaks at the beginning of the period as well as around second 60, which marks the boundary between the first minute in which there is still time to contribute and the last 30 seconds in which the round is possibly terminated. The mean time to contribute across all periods is 24.52 seconds with a standard deviation of 23.26 seconds. This shows that we find considerable variation in contribution timing, which is also consistent with the assumption of behavioral heterogeneity in the threshold theory of collective behavior (Granovetter and Soong, 1983). Summing up, contributions are indeed made sequentially when real-time information about others' contributions is available, providing the prerequisite for the analysis of antecedents for decision timing.

To subsequently test the rest of our hypotheses, we estimated the effects of personality and SVO on contribution decisions and contribution order in a regression framework. The individual behavioral threshold, our main dependent variable, was created following Granovetter (1978). We measure an individual's behavioral threshold "[...] by the exact number of others who have made a decision before he does" (Granovetter, 1978, p.1440). Thus, for each round, an individual-level variable (obsCon) was created, which represents the number of other group member's contributions to the public account before the respective individual contributed him- or herself. For example, before making a contribution, a random individual may have observed nobody in the first round, two contributions in the second round, four contributions in the third round and so on. This individual's values of obsCon would then be 0, 2, and 4, respectively. Thus, each individual can have a maximum of 10 values for obsCon over the course of each

session, which can range from 0 to 7 within each round, as all groups contain exactly eight participants. As this procedure only yields values for cases in which individuals actually contributed to the public good, our analyses are limited to those cases. When conducting our analyses, we generally controlled for risk aversion, trust, reciprocal inclinations, and several socioeconomic factors.[8] In addition, we included period fixed effects to capture time trends and overall learning in our regressions.

As a benchmark, we start by regressing a binary variable indicating whether an individual made a contribution or not on our variables of interest and set of controls. Table 3.3 reports the model using the whole sample including treatment indicators in column (1) as well as separate regressions for each of the four treatments in columns (2) to (5). Here, it can be seen that an individual's SVO angle significantly increases the probability to make a contribution to the public account, while the HEXACO dimensions turn out to be non-significant predictors for contribution propensity. Thus, distal personality measures do not influence contribution decisions on the extensive margin in our setting. However, as we are primarily interested in the intensive margin, that is, how personality factors influence contribution order, this is not discouraging per se.

Next, we turn to the exploration of antecedents of contribution order. As obsCon can take eight possible values, but is an ordinal variable, we employ a panel ordered probit approach. As with contribution decisions, we report an estimation using all observations and including an indicator variable for the provision point as well as separate estimations for the *Baseline* and *High Provision Point* treatments. The estimation results for all specifications are summarized in Table 3.4. Looking at the overall model in column (1), four predictors turn out to be significant: The treatment indicator is significant by design. As predicted, a larger SVO angle, that is, a more prosocial orientation, is associated with earlier contributions. Interestingly, while the coefficients for both agreeableness and honesty-humility turn out to be insignificant, two other HEXACO personality dimensions are significant predictors for contribution sequence: while extraversion and the number of observed contributions are significantly negatively related, meaning

[8] These include gender, age, subject of study, and general mental ability proxied by math grade.

Table 3.3: *Panel probit regressions for contribution decisions*

	(1)	(2)	(3)	(4)	(5)
DV: Contributor	Overall	Baseline	High Provision Point	No Info	Combined
Openness	6.55e-05	-0.0250	-0.00156	0.0460	-0.0149
	(0.0268)	(0.0436)	(0.0479)	(0.0588)	(0.0525)
Conscientiousness	-0.0374	-0.0358	0.00419	-0.0646	-0.0539
	(0.0252)	(0.0348)	(0.0497)	(0.0602)	(0.0555)
Agreeableness	0.000234	-0.00553	0.00630	0.0215	0.0372
	(0.0275)	(0.0450)	(0.0569)	(0.0643)	(0.0603)
Extraversion	0.00903	0.0247	0.00263	-0.0333	-0.0217
	(0.0246)	(0.0383)	(0.0484)	(0.0561)	(0.0604)
Emotionality	0.0136	-0.0248	-0.0109	0.107**	0.0158
	(0.0254)	(0.0366)	(0.0551)	(0.0532)	(0.0595)
Honesty	0.0132	0.0718**	-0.0119	0.0240	0.0399
	(0.0238)	(0.0354)	(0.0487)	(0.0634)	(0.0424)
SVO angle	0.00523***	0.00603***	-0.000656	0.00544*	0.00717***
	(0.00114)	(0.00175)	(0.00225)	(0.00299)	(0.00273)
High PP = 1	0.0508*				
	(0.0285)				
No Info = 1	-0.0197				
	(0.0309)				
High PP	-0.191***				
# No Info	(0.0616)				
Observations	3,600	1,440	720	720	720
Groups	360	144	72	72	72
Wald Chi^2	139.1	75.31	45.14	57.12	103

Notes: The table reports marginal effects after panel probit regressions. Controls include trust, positive and negative reciprocity, risk aversion, gender, age, subject of study, mental ability, and period fixed effects. Standard errors are clustered at the individual level. *** p<0.01, ** p<0.05, * p<0.1.

Table 3.4: *Panel ordered probit regressions for contribution order*

	(1)	(2)	(3)
DV: obsCon	Overall	Baseline	High Provision Point
Openness	0.0581	0.00525	0.155
	(0.118)	(0.143)	(0.214)
Conscientiousness	0.261***	0.306***	0.221
	(0.0998)	(0.117)	(0.203)
Agreeableness	-0.103	-0.0805	-0.0701
	(0.123)	(0.131)	(0.229)
Extraversion	-0.217**	-0.345***	-0.0414
	(0.109)	(0.126)	(0.186)
Emotionality	0.0687	0.0560	0.116
	(0.0986)	(0.111)	(0.207)
Honesty	-0.123	-0.281**	-0.00172
	(0.107)	(0.139)	(0.197)
SVO angle	-0.0198***	-0.0243***	-0.0152*
	(0.00505)	(0.00595)	(0.00900)
High PP = 1	0.527***		
	(0.151)		
Observations	1,510	937	573
Subjects	214	142	72
Wald Chi^2	60.82	53.39	47.23

Notes: Table shows raw coefficients after panel ordered probit regressions. While all relevant observations are used in column (1), estimation in column (2) uses *Baseline* treatment observations only, and column (3) uses *High Provision Point* Treatment observations only. Controls include trust, positive and negative reciprocity, gender, age, subject of study, mental ability, and period fixed effects. Standard errors are clustered at the individual level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

that individuals with higher extraversion scores contribute earlier, conscientiousness has a positive impact, meaning that individuals with higher conscientiousness scores contribute later in the sequence. Thus, the estimation results partially support our hypotheses: while we find support for Hypotheses 3.7, 3.9 and, 3.10, the rest (Hypotheses 3.5, 3.6, and 3.8) cannot be supported. Hence, in our sample, more prosocial individuals are more likely to contribute earlier, indicating lower behavioral thresholds. With respect to the more distal personality traits, emotionality as well as honesty and agreeableness have no impact on contribution behavior. However, we find that extraversion and conscientiousness influence the propensity to contribute early.

3.4.3 A Closer Look at Behavioral Types

Finally, we take a closer look at different behavioral types to gain further insights into what might determine contribution behavior in our setting.

First, we rerun our main analysis of contribution order, but exclude all pivotal contributions, because these contributions are strategically fundamentally different from the others as unlike for every contribution before that, there is no uncertainty about whether the provision point is going to be reached or not – if the pivotal contribution is made, the public good is going to be implemented for sure. The results are presented in column (1) of Table 3.5, revealing a similar pattern to the analysis comprising all contribution decisions without accounting for their distinct strategic nature. Again, the coefficients for SVO angle and extraversion are significantly negative, while the coefficient for conscientiousness is significantly positive. Hence, even when only looking at contributions under uncertain conditions, we still find that individuals with a more prosocial attitude and more extraverted individuals contribute earlier in the sequence, while more conscientious individuals contribute later.

Next, we take a closer look at pivotal contributions only, following the rationale explained above. More specifically, we compare all non-pivotal contributions to those that directly lead to implementation of the public good using a probit regression. As can be seen in column (2) of Table 3.5, the propensity of making a pivotal rather than another contribution is significantly increasing in the degree of conscientiousness of decision

Table 3.5: *Subgroup regressions*

DV:	(1) Uncertain obsCon	(2) Pivotal (Y/N)	(3) Leader (Y/N)
Openness	0.0521	0.0190	-0.0247
	(0.123)	(0.0255)	(0.0299)
Conscientiousness	0.182*	0.0581**	-0.0626***
	(0.101)	(0.0227)	(0.0243)
Agreeableness	-0.0723	-0.0209	0.0201
	(0.124)	(0.0277)	(0.0313)
Extraversion	-0.184*	-0.0163	0.0383
	(0.109)	(0.0236)	(0.0244)
Emotionality	0.153	0.000157	-0.0495*
	(0.0948)	(0.0250)	(0.0258)
Honesty	-0.00538	-0.0289	0.0108
	(0.108)	(0.0254)	(0.0262)
SVO angle	-0.0148***	-0.00412***	0.00344***
	(0.00497)	(0.00114)	(0.00109)
High PP = 1	0.758***	-0.0316	-0.0207
	(0.158)	(0.0295)	(0.0344)
Observations	1,193	1,510	1,510
Number of Subjects	200	214	214
Wald Chi^2	56.11	37.15	26.75

Notes: Column (1) shows raw coefficients after panel ordered probit regression, while columns (2) and (3) depict marginal effects after panel probit regressions. Controls include trust, positive and negative reciprocity, gender, age, subject of study, mental ability, and period fixed effects. Standard errors are clustered at the individual level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

makers, while it is significantly decreasing in SVO angle. Thus, more conscientious and less prosocial individuals are more likely to make pivotal instead of uncertain contributions to the public account.

In addition, we also explore which factors influence the decision to make the first move in the group, that is, to contribute before anyone else does. As expected, the coefficients for SVO angle and conscientiousness depicted in column (3) of Table 3.5 mirror those for the other obsCon regressions, indicating that individuals with a higher prosocial attitude and lower conscientiousness are more likely to be first movers. Surprisingly, however, the coefficient of emotionality is also significant and negative, indicating that higher scores in emotional stability decrease the propensity to make the first move.

3.5 Discussion

Summing up, our laboratory experiment yields some interesting insights. First, while implementation success is higher when the provision point is lower, i.e., only five rather than seven contributing individuals are needed for the public good to be implemented, availability of real-time information about others' behavior does not automatically facilitate efficient coordination and implementation success. While this is somewhat surprising in light of our hypotheses, it is also in line with what has been reported in the literature in slightly different settings. A specific explanation in our setting might be a possible ceiling effect: As compared to the literature our implementation rates are very high, especially when the provision point is set at five contributions. Consequently, the relatively easy goal of finding these five individuals implies that real-time information cannot make a big difference in facilitating implementation success. However, we do find that the combination of a high provision point and a lack of availability of real-time information leads to a steep decrease in group success. Thus, overall, our results reveal that the institutional environment matters for group outcomes.

Furthermore, our experiment shows that the coordination problem our subjects face leads to contribution sequences as predicted by threshold theory. While some subjects are instigators and actively contribute very early, moderates wait for some collective

movement to begin, while again others are quite reluctant and need to see a lot of other individuals contribute before contributing themselves. Furthermore, our results show that while participation in collective action can be predicted by SVO alone, the sequencing is influenced by a broader array of individual differences (i.e., conscientiousness, extraversion, and SVO). Yet, not all of these individual differences affect behavioral thresholds in the same direction: while extraversion and SVO further individual proactivity, conscientiousness hinders it.

However, the results concerning emotionality, agreeableness, and honesty are somewhat surprising in light of our hypotheses. While both the theoretical rationale as well as the empirical results regarding the VCM were convincing ex ante, their applicability under the PPM seems to be limited. Albeit agreeableness is conceptualized as a rather reactive trait domain, the context may have suppressed this effect. Moreover, our setting does not provide extensive possibility to show strong reactions, such as direct punishment of others or retaliation in future sessions. It might be the case that our treatment interventions were strong enough to override the behavioral tendencies due to distal individual heterogeneity in personality structure. In contrast to the standard trait-based approach, other scholars in psychology have been in favor of more complex person-situation interactions (Hough et al., 2015; Mischel and Shoda, 1995, 2010; Schein, 2015). Using the concept of situational strength, it has been argued that cues provided by the environment heavily influence behavior formation (Dalal et al., 2015; Judge and Zapata, 2015). While in strong situations, personality has only a modest influence on overt behavior, in weak situations which are unstructured, without salient, or very ambiguous cues, personality plays a much stronger role (Meyer et al., 2010). This could be the reason why agreeableness cannot explain additional variation in our experimental setting. A somewhat similar argument can be made for emotionality and honesty. While we are not able to discern an emotionality and agreeableness effect at all, it is interesting that the effect of honesty is at least visible in the *Baseline* of our setting, suggesting an interaction of situation and personality.

However, our results have to be seen in the light of some limitations that offer various opportunities for future research. First, the advantage of controlling for situational

influences in laboratory settings naturally comes at a cost. Random rematching and irrevocable decisions rule out within-group learning, coalition building, and retaliation. Future research may use the real-time protocol including revocable commitments (Goren et al., 2004) in order to extend our results to these decision processes. Second, we investigate behavior formation when individuals are confronted with all-or-nothing contribution decisions. Thus, our results are limited to the basic decision of supporting a project or task by exerting costly effort at all or not. In addition, various forms of teamwork are characterized by a continuous contribution to the public good, which our design does not capture. Third, our design does not include heterogeneity of players with regard to network centrality. It has been reported in the literature on network effects that bandwagon dynamics can be severely impacted if individual contributions of central peers in their respective network have larger impact than more distal players (Chiang, 2007). Fourth, some of our analyses and consequently our implications are based exclusively on those individuals who contributed to the public good. This limitation is becaue albeit we are able to tell that the respective behavioral threshold of individuals choosing not to contribute was not reached, we cannot infer what their exact behavioral threshold would have been. Fifth, while we take into account that personality is a more useful explanation of behavior in weak situations than in strong situations, our treatments may interfere with the situational strength of the social dilemma. Future studies may directly compare group behavior with weak and strong situational cues and interact it with the group's personality mix (e.g., ratio of prosocials). Sixth, future studies may take a closer look at self-selection of individuals into the respective situational contexts. The usage of self-selection rather than authority as a means to allocate tasks throughout the firm has been argued to be one of the true novelties of new forms of organizing (Puranam et al., 2014). In order to build implications for practice, it would be interesting to know if individual differences yield substantial differences for self-selection. Subsequently, the performance implications of sorting may be analyzed by comparing self-selected to exogenously selected groups. According to the attraction-selection-attrition framework (Scheider, 1987) or the person-organization-fit literature (Kristof, 1996), selection into (and out of) organizations is not a random process. Thus,

employee self-selection may severely limit an organization's ability to staff task forces or project teams with the individuals prone to take lead in contribution sequences. As Bridoux et al. (2011) point out, a given mix of motives demands for a differential and nuanced approach regarding the motivational system that needs to be employed to yield maximum contributions. Finally, our study investigates the general causal effect of institutional manipulations on contribution behavior. However, we cannot tackle the exact mechanism responsible for collective results at the individual level. According to Zafar (2011), the general effect of manipulating information about others' behaviors may be due to individuals trying to infer the best choice (social learning), actually gaining utility from acting alike (social comparison), or wanting to gain a positive self-concept (social image).

3.6 Conclusion

The findings of this experiment suggest interesting implications for theory development, practitioners and future research. For practitioners and applied scholars, it is noteworthy that institutional factors eminent to the environment or malleable by the individual manager matter a great deal towards success and efficiency of project teams or strategic initiatives. It has been shown by the results of our experiment that information provision about group member's behavior can significantly facilitate coordination. What is more, our results support the basic rationale of the threshold model of collective behavior. In fact, individuals show substantial heterogeneity in threshold formation when confronted with the production of a provision point public good with endogenous decision timing. More importantly, our study provides empirical support for the anecdotal argument that this heterogeneity may, in part, be caused by individual differences. Furthermore, our study generally supports the notion that heterogeneous motives matter in the context of collective behavior formation. Given the fact that coordination has been deemed as one of the central tasks of any organization (Okhuysen and Bechky, 2009; Puranam et al., 2014), those mechanisms should be tested under real organizational conditions using for example field experiments or insider econometric studies.

Chapter 4

The Influence of Overconfidence and Competition Neglect on Entry into Competition

4.1 Motivation

Traditionally, in contract theory, the rationale for providing incentives is to align the interests of principals and agents in order to overcome contracting problems in the presence of information asymmetries (Hölmstrom, 1979; Grossman and Hart, 1983). However, this view underestimates the importance of worker self-selection, i.e., that different types of workers are attracted by different kinds of incentive schemes and therefore systematically select into different environments. When considering both aspects, performance does not only depend on an incentive effect, but also on a sorting effect (e.g., Lazear, 2000; Dohmen and Falk, 2011; Larkin and Leider, 2012).

Selection into competitive environments has proven to be an important factor in many economic settings like labor markets (based on rank-order tournaments; cf. Dechenaux et al., 2015) or entrepreneurship (Camerer and Lovallo, 1999). Findings reveal that factors other than productivity influence selection into competitive environments (e.g., Gneezy et al., 2003; Niederle and Vesterlund, 2007; Vandegrift et al., 2007; Bartling et al., 2009; Eriksson et al., 2009; Dohmen and Falk, 2011; Balafoutas et al., 2012). Papers investigating determinants of entry decisions from different angles have identified several economic preferences that influence entry into competitive environments: risk aversion and distributional preferences matter as well as heterogeneity in attitudes

towards competition. Aside from preferences, differences in overconfidence have also been brought forward as explanatory mechanism. However, in this area, central blind spots remain as overconfidence comprises different facets that need to be carefully measured, and alternative mechanisms have not been systematically explored.

First, rather than being a clearly defined mechanism, overconfidence is an umbrella term containing several different biases in beliefs or, more precisely, errors in performance assessments, that have been discussed controversely. Two different facets of overconfidence identified by Moore and Healy (2008), the overestimation of one's ability (overestimation) and the belief to be better than others (overplacement), might play a role for entry decisions. Different aspects have been emphasized by different literature streams: while some have focused on overestimation of own ability (the literature related to Camerer and Lovallo, 1999), others concentrate on relative overconfidence or overplacement while effectively eliminating uncertainty about own performance (the literature buliding on Niederle and Vesterlund, 2007), with comprehensive approaches being scarce.[1]

Second, overconfidence in beliefs has mostly been described as being caused by subjective priors or updating biases like attaching excessive precision to signals (Jehiel, 2016). However, other mechanisms that lead to biased perceptions might play a role as well. Specifically, it has been argued in research on bounded rationality that individuals do not realize that others' actions depend on their private information. This notion is incorporated in solution concepts like cursedness (Eyster and Rabin, 2005) or behavioral equilibrium (Esponda, 2008) and has been utilized as driver of biased beliefs using different terminologies – for instance, Enke and Zimmermann (2017) show that correlation neglect leads to biased information processing, and Jehiel (2016) derives overoptimism from selection neglect. Closest to the setting of this paper, Camerer and Lovallo (1999) identify competition neglect, that is, that individuals tend to neglect self-selection of (potentially high-performing) others into competition as a potential mechanism biasing entry decisions. However, since this mechanism has been brought

[1] The experiment by Kamas and Preston (2012) constitutes a notable exception as the authors explicitly relate to the proposed systematization by Moore and Healy (2008).

up as a potential cause for excess entry, only a handful of papers have followed up on it. On the contrary, most studies rule self-selection of competitors out by design, and systematic identification of the role of this bias for competition entry remains sparse.

This study aims at investigating whether overconfidence and competition neglect influence decisions to enter competitive environments. I use a controlled laboratory setting that allows to elicit belief distributions related to absolute as well as relative overconfidence to study it comprehensively while influencing it via performance feedback. In addition, my setting allows me to introduce systematic variation in the possibility for competition neglect. Studying these two mechanisms simultaneously is relevant as both overconfidence and competition neglect potentially lead to the same behavioral pattern, namely excess entry. However, for de-biasing, it is essential to know whether individuals overestimate their chances of success because they overestimate their performance in absolute terms or relative to others, or because they do not realize that they face a selected set of competitors.

With this in mind, I let individuals decide whether they want to enter a competition in a real-effort experiment while eliciting their beliefs with respect to their own and others' performance as well as their winning probability. To answer my research question, I introduce two treatment variations: First, some participants receive detailed performance feedback addressing absolute and relative overconfidence before making their decision. Comparing entry decisions of individuals who have received feedback to those who have not enables me to investigate whether overconfidence plays a role for selection into competition – if there is a difference between those groups, it does. Second, I vary whether the competition group consists of all potential competitors or only of individuals who also chose to compete. Here, finding no difference in entry between the groups implies that decision makers fail to adjust their decisions to the selected sample, thus exhibiting competition neglect.

The main findings of this paper are as follows: Individuals exhibit pronounced heterogeneity with respect to their performance beliefs, with low-performing individuals overestimating their own performance and their chances of success while underestimating performance in the competition group, and the opposite being true for high-performing

individuals. These biases in performance beliefs are ameliorated by feedback, but the composition of the competition group is persistently disregarded. Investigating determinants of entry decisions to tackle the key question of this paper whether competition neglect and overconfidence influence entry into competition, I find that both influence individuals' decisions. However, competition entry closely tied to previous performance and assessments, and in contrast to the literature, there are no gender differences in beliefs and entry decisions.

The remainder of this chapter is organized as follows. Section 4.2 gives an overview of related literature. Section 4.3 presents a simple theoretical framework illustrating the main argument of this paper that overconfidence and competition neglect lead to the same behavioral changes while constituting different mechanisms. This is the prerequisite of the experimental design, which is presented in Section 4.4. Section 4.5 describes and analyzes the experimental results. Section 4.6 discusses and concludes.

4.2 Related Literature

This study relates to several strands of literature. First, there is a wide variety of research focusing on overentry into markets of entrepreneurs, which are highly competitive environments (see, e.g., Camerer and Lovallo, 1999, for an early lab experiment). In this setting, empirical evidence suggests that an entrepreneur's confidence in his own skill, ability and knowledge has a crucial impact on new business creation and might explain the high failure rate of new business owners (Koellinger et al., 2007; Cain et al., 2015; Bolger et al., 2008). Apart from overconfidence, the neglect of selective entry of competitors into a market may be a leading cause of failure: Even if an entrepreneur assesses his skills correctly, he may still be not aware of the fact that his competitors are not a random sample of the population, but have abilities that are at least as good as his. The competitors' skills may be even higher, since their business has survived years of constant pressure from market powers that drive out non-competitive participants (Camerer and Lovallo, 1999; Moore et al., 2007). Competition neglect has been shown to also matter for market entry decisions in other situations like selling on eBay (Simonsohn, 2010). However, systematic identification of competition neglect remains

scarce. Thus, my study contributes to this strand of literature by experimentally varying the possibility of competition neglect.

Second, entry into competition has been extensively studied from a different angle: over the course of the last decade, a large number of studies has explored gender differences in competition decisions (see Niederle and Vesterlund, 2011, for an overview). In particular, one worrying finding is that high-performing women appear to opt out of competitive environments (Niederle and Vesterlund, 2011). Building on the seminal work by Niederle and Vesterlund (2007), who focus on individual-level competition decisions and find a large gender gap, it has been suggested that women show a lower willingness to engage in competition, a finding that has also been replicated in real-world occupation settings (Flory et al., 2015). Regarding the mechanisms driving this, a plethora of laboratory evidence has been produced. At first, differences were mostly attributed to heterogeneity in attitudes towards competition as well as risk and feedback aversion (Eckel and Grossman, 2008). Lately, it has been argued, however, that differences in beliefs could explain the above-described outcomes. Psychology research suggests that women exhibit less overconfidence than men, a finding that has been recently incorporated in economic investigations (see Niederle and Vesterlund, 2011). My study thus contributes to the literature on gender and competition by investigating more closely how biased beliefs about the performances of oneself and the competition contribute to differences in selection into competition.

Third, my study is related to the literature addressing the potentially benefiting role of feedback in reducing overconfidence effects (see Arkes et al., 1987, for an early paper on this). Several papers study how subjects update their relative performance beliefs and the relation to competition decisions. A general finding is that while there is belief updating, after noisy feedback, individuals tend to update less than a Bayesian would (Möbius et al., 2014). For instance, Buser et al. (2016) show that there is heterogeneity in this conservatism, with women being more conservative than men. Wozniak et al. (2014) find that providing relative feedback causes more high-performing women and less low-performing men to select into competition, thus reducing the gender gap in competition decisions. Berlin and Dargnies (2016) use a setting close to mine (except

for only giving binary feedback) and show that different aspects of performance matter for men and women. In contrast to most of this literature, I give full performance feedback. In this setup, I also see sensible, albeit conservative belief adjustment.

Before going into details of the experiment, I now describe a simple theoretical framework of how overconfidence and competition neglect influence decision making.

4.3 Theoretical Framework

The goal of this section is to stress the central argument of this paper that multiple causes of biased beliefs leading to biased entry into competition exist that need to be addressed in different ways: overconfidence and competition neglect. Thus, the model illustrates how the two mechanisms lead to the same changes in behavior albeit being conceptually different.

4.3.1 Model Setup

I assume that an individual i can be of two types, $\theta_i \in \{\theta_L; \theta_H\}$, where $\theta_H > \theta_L$. θ_i represents the productivity of the individual in some task. There are N individuals who all face an identical game with the following structure that is common knowledge: First, nature draws the type of all players $i \in \{1, .., N\}$. Here, type θ_H is assigned with probability p. Second, every agent receives a private signal $S = \theta_i + \epsilon$ with $\epsilon \sim \mathrm{N}(0, \sigma^2)$, which determines the posterior probability q_i the individual assigns to being of type θ_H.[2] Hence, there is individual heterogeneity with respect to posterior beliefs q_i. Third, every individual is facing a choice between two options: to enter a competition or to receive an individual payment. If competition is chosen, the individual is matched with a randomly drawn opponent j who has also chosen to compete.[3] She receives a piece rate of p_C if her productivity is higher than that of her opponent and nothing otherwise. If both individuals are of the same type, the winner is randomly chosen. If the individual

[2] This structure was chosen to generate continuous posterior beliefs about being of type θ_H. This has the advantage that it allows to identify threshold equilibria and analyze comparative statics.

[3] Note that I assume here that there is always at least one competitor as allowing for the possibility that there is none would make the problem more complicated while not affecting the main conclusions. Furthermore, this case did never occur in the experiment.

payment is chosen, a lower piece rate of $p_{NC} < p_C$ is received for sure.

4.3.2 Rational Benchmark

Assuming risk neutrality, individual i's expected utility from choosing competition equals

$$\mathbf{E}_C = p_C \times [q_i\theta_H \times (0.5q_j + (1-q_j)) + (1-q_i)\theta_L \times 0.5\,(1-q_j)],$$

where $q_j = \Pr(\theta_j = \theta_H |\ j \text{ enters})$. In this case, her expected utility equals

$$\mathbf{E}_{NC} = p_{NC} \times [q_i\theta_H + (1-q_i)\theta_L].$$

The individual decision is made by comparison of the expected utilities in both cases: if the expected utility from entering the competition exceeds the expected payoff from the individual payment, the individual enters. Then, she is indifferent between entering the competition and receiving an individual payoff at

$$q_i^*(s,\xi,q_j) = \frac{2-(1-q_j)s}{\xi[s(2-q_j)-2]-[(1-q_j)s-2]}$$

where $s = \frac{p_C}{p_{NC}} > 1$ and $\xi = \frac{\theta_H}{\theta_L} > 1$.

I focus on the symmetric Bayesian Nash equilibrium, which is characterized as follows:

Proposition 4.1. *For any q_i and $q^* = \frac{2-(1-q_j)s}{\xi[s(2-q_j)-2]-[(1-q_j)s-2]}$, there exists an equilibrium in pure threshold strategies where $\forall i \in \{1,...,N\}$, individual i enters the competition if $q_i > q^*$, and chooses the individual payoff if $q_i < q^*$.*

To establish $q_i \lessgtr q^*$ as sensible decision criterion, q^* has to exist, be monotone and unique. Proofs for existence and uniqueness of q^* are straightforward. For monotonicity, I investigate the effects of s, ξ, and q_j on q^* and show that all of them are unambiguous.[4] First, the comparative static of q^* with respect to the piece rate ratio s is

$$\frac{\partial q^*}{\partial s} < 0.$$

[4] The proofs for existence and uniqueness as well as the derivatives and the determination of their signs can be found in Appendix D.1.

In terms of the model, this means that as the piece rate ratio increases, the cutoff for entering the competition decreases. Intuitively, this makes sense as the piece rate ratio is a measure for the relative attractiveness of the competition that increases in the difference between the competitive and individual piece rates. Hence, as competition becomes more attractive compared to the individual incentives, the cutoff value decreases and more individuals enter the competition. Second, the comparative static of q^* with respect to the productivity ratio ξ is

$$\frac{\partial q^*}{\partial \xi} < 0.$$

This means that as the performance difference between the high type and the low type increases, the decision cutoff decreases, which makes intuitive sense: as relative productivity associated with being of type θ_H increases, expected utility from choosing competition rises faster than expected utility from choosing the individual pay as the piece rate is higher for the former than for the latter. Third, the comparative static of q^* with respect to the probability that the competitor is of type θ_H, q_j, is

$$\frac{\partial q^*}{\partial q_j} > 0.$$

Thus, the decision cutoff increases in probability that the competitor is of type θ_H. This result is intuitive as well as an increase in this probability makes it less likely that the competition is won. This way, an increase in q_j decreases the expected utility from competition while it does not affect expected utility from individual payment.

4.3.3 Introducing Biased Beliefs

Having established the cutoff value q^* as an equilibrium, I now introduce overconfidence and competition neglect and derive predictions how the cutoff and thus the fraction of individuals entering competition changes. Throughout, I assume full naiveté about the biases; that is, all players assume that they themselves and all other players act

according to the rational benchmark.[5]

For overconfidence, I follow Moore and Healy (2008) and focus on two different aspects, overestimation of own performance and underestimation of others' performance, as both are potential ways to generate overplacement.

First, to introduce a simple notion of overconfidence, assume that the true probability of being of type θ_H is q_i^{true}. Rather than having realistic beliefs about own ability, i.e., $q_i = q_i^{true}$, an individual subject to overconfidence believes that $q_i = \hat{q}_i$ with $\hat{q}_i > q_i^{true}$. Then, beliefs that satisfy $q_i^{true} < q^* < \hat{q}_i$ lead to competition entry even though this is not payoff-maximizing. This leads to the following prediction:

Prediction 4.1. *Overestimation of own performance leads to overentry into competition.*

An additional way to introduce overconfidence is via the belief about the type of the competitor, q_j. More specifically, assume that an overconfident agent underestimates the competition, i.e. $\overline{q}_j < q_j^{true}$. As already discussed above, the cutoff value q^* is positively related to q_j. Thus, if the agent underestimates the probability of facing a strong competitor, she is also going to underestimate the cutoff value: $q^*(\overline{q}_j) < q^*(q_j^{true})$. This leads to the following prediction:

Prediction 4.2. *Underestimation of competitor's performance leads to overentry into competition.*

Second, to investigate the effect of competition neglect on competition choices, the belief about the probability that the random competitor is of type θ_H, q_j, is central. Remember that the cutoff value for selecting into competition increases in q_j. To understand how competition neglect influences this value, note that q_j is a conditional probability as the competitor has also made the choice to compete. Assuming rational agents, $q_j^{true} = \Pr(\theta_j = \theta_H |\ j \text{ enters}) \in [q^*, 1]$. In contrast, if the individual is subject to (full) competition neglect, she does not take into account the choice made by the competitors and considers the unconditional probability that her competitor is of type

[5] Of course, this is a very strong assumption that is made to keep the analysis as simple as possible.

θ_H, $\hat{q}_j = \Pr(\theta_j = \theta_H) \in [0; 1]$. It follows that $\hat{q}_j < q_j^{true}$, which in turn leads to a decrease in the cutoff value: $q^*(\hat{q}_j) < q^*(q_j^{true})$.

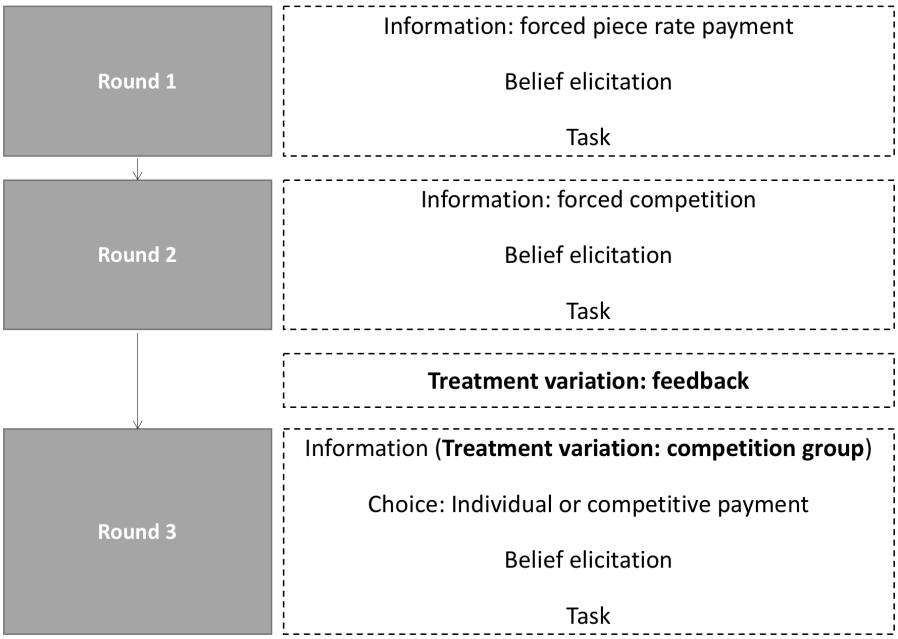
Prediction 4.3. *Competition neglect leads to overentry into competition.*

This simple model sketch illustrates the argument central to this paper, namely that overconfidence and competition neglect lead to similar changes in behavior compared to the rational benchmark – both lead to excess entry into competition. Thus, they cannot be disentangled by only looking at behavioral outcomes. As a consequence, both potential mechanisms have to be investigated carefully in an experiment exploring causes of biased entry. The following section presents the experimental design utilized in this study in detail. It is designed to help explore the role of overconfidence and competition neglect for competition decisions.

4.4 Experimental Design

The experiment proceeded as follows: Upon arriving at the laboratory and being randomly assigned to cubicles, participants were given general instructions and told that they would work on two unrelated parts. They were informed that their payment would depend on their performance and actions in the experiment, but that only one of the rounds or choices in each of the two parts would be paid out. To ensure incentive compatibility, participants would only be notified of the payoff-relevant rounds, decisions and performance at the very end of the experiment. Then, the main part of the experiment started immediately. In each of several rounds, participants had to work on different tasks: They performed a real-effort task under varying incentive schemes and had to answer several belief elicitation questions. In the last of three rounds, participants also had to make a choice between individual and competitive pay. Beliefs were elicited in every round. To experimentally vary overconfidence and competition neglect, I included two types of between-subject variation: depending on treatment, participants either received performance feedback or not, and either faced a random competitor or one that chose to compete. A summary of the procedure of this part is shown in Figure 4.1, and details about the different elements of the main part of the experiment are

Figure 4.1: *Experimental procedure in the main part of the experiment*



provided in the following subsections.[6] After completing this part, risk preferences of participants were elicited as these have been shown to be an important determinant of selecting into competitive environments. This was done using the standard choice list procedure following Holt and Laury (2002). After this second part was completed, participants were informed about their performance and payoffs and had to answer some general socio-economic questions.[7]

4.4.1 Real-Effort Task

In the real-effort task, participants had to correctly add as many blocks of five two-digit numbers as possible in three minutes. This task is commonly used in the experimental literature because it is easy to explain and produces substantial variation in individual performance (Niederle and Vesterlund, 2007). Subjects were not allowed to use a cal-

[6] Note that the experiment also incorporated a fourth round, which required subjects to choose between incentive schemes, but based on past rather than future performance. This round was conducted for robustness only and is thus not included in the main part of this paper. The results are qualitatively similar for those of round three. For the interested reader, a short description of this round and results are provided in Appendix D.4.

[7] These included gender, age, field of study, math grade as a crude proxy for general mental ability, familiarity with experiments, and mood.

culator, but were given a pen and scratch paper that they were welcome to use. In each round of the real-effort task, five randomly selected two-digit numbers were depicted on the screen in a row, with an additional field for the solution. Upon typing in their solution, participants had to confirm by clicking a "Submit" button and were then given the next five numbers. In contrast to Niederle and Vesterlund (2007), subjects were not forced to try one problem until they got the answer right, and they were not informed about the number of problems they had already solved correctly in that round. This was mainly done because this way, there is a greater scope for overestimation of own performance.[8] In addition to staying close to the related literature, I chose this task because it can be seen as ego-relevant for a student subject pool and is therefore more likely to produce overly favorable judgments (Grossman and Owens, 2012) and because it has been argued in the literature there should not be a significant gender gap in performance in this task (e.g., Niederle and Vesterlund, 2007; Buser et al., 2014).

4.4.2 Incentives and Reference Group

After being told that they would have to work on the real-effort task and being given one minute to practice, participants faced three rounds with the same basic setup: In the beginning of each round, participants were informed about the payment rules. Then, performance beliefs were elicited before the actual task was conducted.

The rounds mainly differed in incentives. More specifically, following Niederle and Vesterlund (2007), in the first round, all subjects received a piece rate of 0.50 € for every correctly solved item. In the second round, all subjects then faced a competitive pay: they received a higher piece rate of 1.00 € per correctly solved item, but only if their performance in that round was higher than that of a randomly chosen opponent from the reference group. For round three, participants then could choose between the lower piece rate of 0.50 € and the competitive payment scheme.

Here, it is important to note that the reference group for this and every other competi-

[8] In addition, the authors also state that they chose to force participants to answer a problem correctly before being able to proceed to prohibit them to skip hard problems and search for easy ones. In my case, there was no subject showing this behavior.

tive incentive scheme as well as for the corresponding belief elicitations explained in the following subsection did not consist of the other participants of the same experimental session, but of participants of a past session. This was mainly done for two reasons: On the one hand, it ensured that it was possible to inform participants about choices of the reference group before they had to make their own choices, and on the other hand, this way every observation in one session can be treated as an independent observation, which substantively increases the statistical power of the analyses. In addition, individual behavior could not affect others' payoff, ruling out any influence of social preferences on competition choices.

4.4.3 Treatment Variations

To get at overconfidence and competition neglect, two treatment variations were implemented in the main part of the experiment: First, to address overconfidence, feedback on own and others' performance was given between rounds two and three in half of the sessions. More specifically, participants in the feedback condition were told the number of items they had solved in round two as well as the average number of items solved and the whole performance distribution in the reference group. Second, the possibility for competition neglect was manipulated by design as in half of the sessions, the competitive pay was dependent on the performance of a randomly drawn participant of the reference group, while in the other half, the competition group in rounds three and four only consisted of participants of the reference group who had also chosen to work under the competitive incentive scheme in that round. This way, participants in the selection condition also had to factor in that they were facing a selected sample of better-performing individuals. Table 4.1 gives an overview of the treatment conditions.

Taken together, these treatment variations make it possible to explore how feedback about own and others' performance affects participants' beliefs and choices, if and how strongly participants are able to factor in the selectiveness of their competitors, and how this can be influenced by feedback as well.

[9] Treatment conditions are named according to the addressed mechanism. Hence, in *Selection*, competition neglect is eliminated while overconfidence is not addressed, and in *Overconfidence*, overconfidence is tackled by feedback while the possibility for competition neglect persists.

Table 4.1: *Overview of the treatment conditions* [9]

		Competition Group	
		Selected	Random
Performance Feedback	No	None	Selection
	Yes	Overconfidence	Both

4.4.4 Belief Elicitation

Because belief elicitation is crucial for addressing my research questions, beliefs are elicited in every round rather than at the very end of the experiment. In addition, instead of only eliciting participants' subjective ranks, I ask a total of three questions to get at the different facets of errors in beliefs and, in this way, at overconfidence and competition neglect.

First, it is important to keep in mind that concerns have been voiced that "overconfident" behavior is consistent with Bayesian updating in many cases. According to Benoît and Dubra (2011), this can be tracked back to the fact that the conventionally used research methodology does not allow to disentangle true overconfidence from rational information processing. Merkle and Weber (2011) identify aggregation of beliefs as greatest concern and show that eliciting belief distributions allows to rule out rational belief updating as explanation. I address these concerns by eliciting the whole belief distribution for own performance in the task as well as performance of a randomly drawn individual of the reference group. To do so, I implemented the procedure proposed by Harrison et al. (2015, 2017). More specifically, beliefs were elicited using a range of possible responses that were presented as ten intervals. Subjects then could allocate a total of 100 tokens according to their subjective beliefs. The allocation translated into payoffs according to the quadratic scoring rule with a maximum payoff of 2 € per question. Harrison et al. (2017) show that using this approach allows to elicit distributions while keeping the bias caused by risk aversion minimal.[10] Thus, I use these two questions to

[10] Subjects were given detailed instructions explaining that the allocation of tokens to intervals com-

assess absolute and relative overconfidence reflected in performance beliefs.

Second, I elicited participants' subjective probability of performing better than a randomly chosen participant of the reference group using a reservation probability or crossover mechanism as first proposed by Ducharme and Donnell (1973) and implemented in variants by Möbius et al. (2014), Buser et al. (2016) or van Veldhuizen (2017), among others.[11] In particular, subjects had to indicate the probability p_s that made them indifferent between receiving either a price of 2 € if their performance is better than that of a randomly drawn opponent or 2 € with that probability. Then, a random number z was drawn from a uniform distribution on the interval $[0, 1]$ and subjects received 2 € with probability z if $z > p_s$ and 2 € if their performance was better than that of a randomly drawn opponent otherwise. This procedure makes stating the true probability incentive compatible.[12] This question functions as compound measure for overconfidence and competition neglect.

4.5 Results

The experiment was conducted at MELESSA, the experimental laboratory of the University of Munich, in June of 2017. A total number of 240 participants were recruited using Orsee (Greiner, 2015) and took part in 10 sessions (two for deriving the performance of the reference group, and eight main sessions) which each lasted about 70 minutes.

The following analyses focus on the eight main sessions. In total, 192 subjects took part in these sessions. Apart from one subject who tried to use a cell phone calculator to solve the task, all participants with a performance of more than 20 correctly solved items per round were excluded from the analysis to eliminate outliers. This leaves 189 subjects (96 men and 93 women) for the analysis.

First of all, I check for pre-treatment differences between the treatment groups by per-

prised a trade-off between payoffs and risk. The full instructions can be found in Appendix D.5.

[11] See Schlag et al. (2015) for an overview of this and other belief elicitation mechanisms.

[12] Again, subjects were given detailed instructions and had to answer several control questions to make sure that they understood that they would maximize their expected payoff by choosing their subjective probability of winning as the threshold value.

forming Kruskal-Wallis tests on observable characteristics to test whether the samples are drawn from the same distribution. The tests detect no significant differences on observables except for biased beliefs about others' performance in rounds one and two (χ^2=8.24, p=.0413 and χ^2=11.56, p=.0009).[13] However, this means that results with respect to beliefs about others' performance can only be interpreted with caution.

In addition, I assess whether there is productivity sorting of the reference group (i.e., whether the performance of the selected sample is indeed better than the performance of the random sample) by calculating a Kolmogorov-Smirnov test for equality of distribution functions between the groups, which turns out to be significant (D=0.35, p=.086), indicating that there is indeed productivity sorting in my sample. Individuals facing the selected sample should factor that in when making competition decisions.[14]

Having established these preliminaries, I now turn to investigating the effects of overconfidence and competition neglect in my sample. I do so in two steps: First, I focus on beliefs, give an impression of how they look like in my sample and show how they are affected by the treatment variations. Then, I turn to entry decisions by again first presenting the general pattern and then addressing treatment effects.

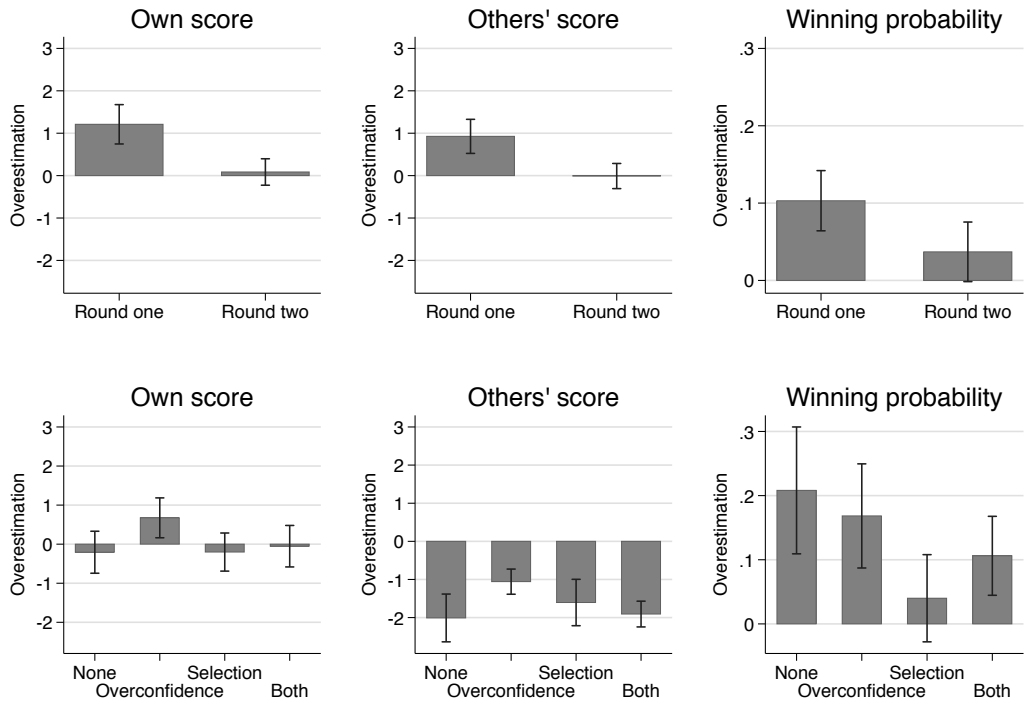
4.5.1 Beliefs

In the following paragraphs, I take a closer look at participants' beliefs across rounds. Despite having elicited whole belief distributions, I first calculate expected values of performances and focus on these. To arrive at a measure for subjects' overconfidence, I then compute the difference between expected and actual number of correctly solved items. Positive values indicate that own performance is overestimated, while negative values indicate that it is underestimated. Analogously, the difference between expected and actual number of correctly solved items of a randomly drawn individual from the competition group is calculated. For subjective winning probability, I compute the actual probability of winning for every number of correctly solved items based on reference

[13] The averages of observable characteristics (age, risk aversion, performance and beliefs in rounds one and two) can be found in Table D.1 in Appendix D.2.

[14] Note that productivity sorting can also be seen in the main sessions (D=0.38, p<.001).

Figure 4.2: *Average bias in beliefs*

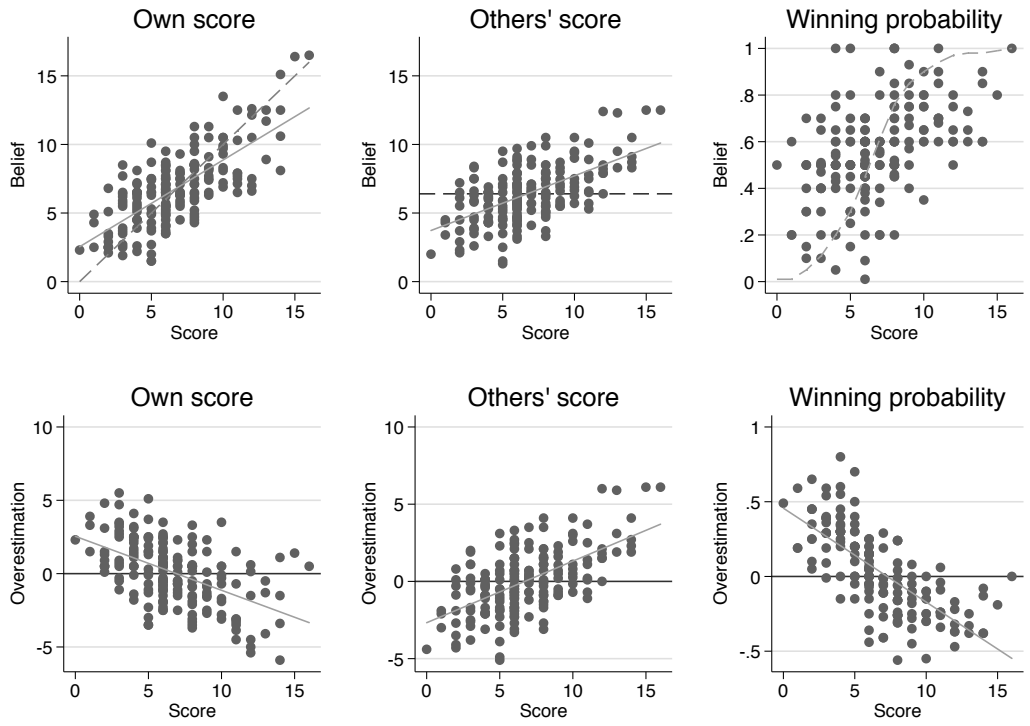


Notes: Upper row depicts average bias in beliefs about own performance (left panel), reference group performance (middle panel) and winning probability (right panel) in rounds one and two in the main sessions. Lower row depicts average bias in beliefs about own performance (left panel), reference group performance (middle panel) and winning probability (right panel) by Treatment in round three in the main sessions. Calculations are based on 189 observations. Means are depicted as bars, 95% confidence intervals as error bars.

group performance and subtract this from participants' stated winning probability. This procedure results in three values for erroneous beliefs per round.

First, I explore how accurate beliefs are and how they are distributed in my sample. To this end, Figure 4.2 depicts average deviations in beliefs from actual scores. Here, two things are especially noteworthy. First, the upper row clearly shows that on average, participants in my sample have very accurate beliefs by round two: while there is overestimation of own and (less drastically) also of others' performance as well as overestimation of the chances of being better than a randomly drawn individual from the reference group in the first round, none of these statements holds true for average beliefs in round two. In addition, the lower row of Figure 4.2 shows that in round three, there are differences in beliefs between treatments: underestimation of others' perfor-

Figure 4.3: *Beliefs and overestimations by actual performance in round two.*



Notes: Upper row of figure depicts beliefs about own performance (left panel), reference group performance (middle panel) and winning probability (right panel) depending on actual performance in round two of the main sessions. Lower row depicts bias in beliefs about own performance (left panel), reference group performance (middle panel) and winning probability (right panel) depending on actual performance. Calculations are based on 189 observations.

mance and overestimation of winning probability is stronger when facing the selected than when facing the random competition group, providing evidence for competition neglect. Interestingly, this difference is more pronounced when no feedback is provided.

Next, I explore heterogeneity in beliefs. As can be seen in Figure 4.3, there is systematic heterogeneity even in round two. Plotting beliefs and overestimation values by actual performance, it is evident that low-performing subjects are more prone to overestimate themselves while underestimating the reference group as well as overestimating their chances of performing better than a randomly drawn individual from the reference group.[15]

[15] One might argue that this effect is at least partly mechanic as high-performing individuals have little scope to overestimate and much scope to underestimate themselves, while for low-performing

Having established how participants' beliefs behave, I now investigate whether they are influenced by giving feedback and by the composition of the reference group to see whether the treatments worked as expected. This way, exploring changes in subjects' assessments serves as a prerequisite for identifying the effect of competition neglect and overconfidence on choices. Thus, I explore determinants of changes in beliefs between rounds two and three in various specifications in Table 4.2.

First, to see whether competition neglect is reflected in participants' beliefs, focus on the coefficient for the selected competition group when change in winning probability is the dependent variable, which is insignificant regardless of specification; thus, compound beliefs are not adjusted more strongly when facing the selected competition group.

Second, I investigate whether absolute and relative overconfidence are ameliorated by providing feedback. The first thing to note is that the coefficient for feedback is insignificant regardless of specification. However, since it might be important to account for the fact that subjects receiving feedback have better information, I additionally look at the interaction between feedback and round two performance. As can be seen in columns (4) to (6) of Table 4.2, the positive effect of previous performance on belief changes is indeed higher when performance feedback is provided. In columns (7) to (9) of Table 4.2, I include round two overestimations and interactions with feedback to explore whether subjects exhibiting higher pre-treatment overestimation correct their estimations more strongly. This would mean that changes in beliefs are sensible, that is, feedback ameliorates perception biases. Indeed, round two overestimations of own and others' performance have a negative influence on changes in beliefs, meaning that stronger round two overestimations are associated with changes towards lower performance beliefs. When feedback is provided, this negative influence is even stronger. For winning probability, previous overestimation only has a significantly negative influence on belief changes when feedback is provided.

---

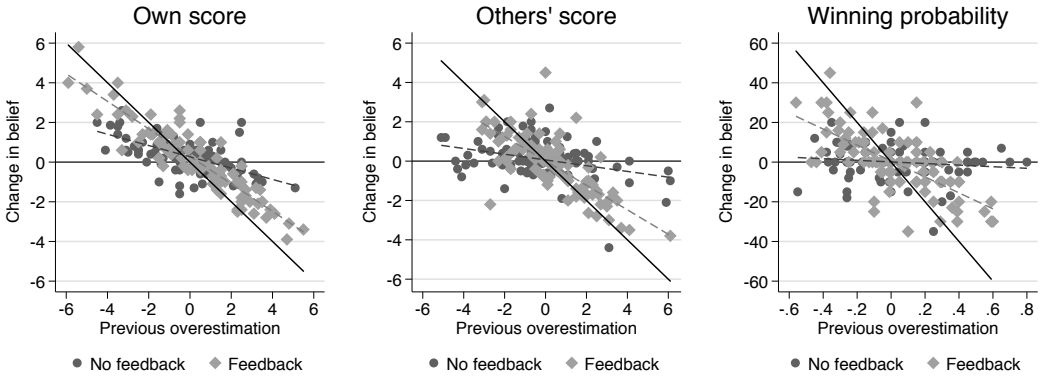
individuals, the opposite is true. However, this does not apply to beliefs about others as these should be independent of own performance. In addition, the relationship between actual and overestimation of own performance and winning probability is still significantly negative if overestimation is not considered continuously, but as binary variable (point-biserial correlations are r=–.37, p<.001 for own performance, and r=–.66, p<.001 for winning probability).

Table 4.2: *Determinants of changes in beliefs between rounds two and three*

DV: Changes in beliefs about...	(1) Self	(2) Others	(3) Probability	(4) Self	(5) Others	(6) Probability	(7) Self	(8) Others	(9) Probability
Feedback = 1	-0.237	-0.225	-0.359	-0.292	-0.195	-1.013	0.0129	0.00671	0.446
	(0.200)	(0.193)	(1.698)	(0.191)	(0.188)	(1.628)	(0.115)	(0.147)	(1.413)
Selection = 1	-0.140	-0.0784	-1.592	-0.196	-0.0485	-2.260	0.0113	0.200	-3.088**
	(0.193)	(0.190)	(1.699)	(0.186)	(0.183)	(1.646)	(0.107)	(0.133)	(1.385)
Previous Performance	0.538***	-0.160	3.230***	0.257**	-0.0101	-0.106			
	(0.119)	(0.112)	(0.941)	(0.103)	(0.114)	(0.795)			
1.feedback				0.590**	-0.315	7.018***			
#c.Score_pre				(0.233)	(0.225)	(1.473)			
Previous overestimation							-0.294***	-0.143***	-3.118
							(0.0382)	(0.0504)	(3.552)
1.feedback							-0.401***	-0.511***	-37.89***
#c.overestimation_pre							(0.0529)	(0.0719)	(5.375)
Observations	187	187	187	187	187	187	187	187	187
Adjusted R^2	0.136	0.00727	0.0783	0.175	0.0172	0.159	0.730	0.462	0.365

Notes: Table reports results of OLS regression. Controls include risk aversion, gender, age, an indicator for field of study, and math grade. Previous score is standardized. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Figure 4.4: *Changes in beliefs by overestimation in round two*



Notes: Figure depicts changes in beliefs about own performance (left panel), reference group performance (middle panel) and winning probability (right panel) between rounds two and three depending on overestimation in round two of the main sessions. Calculations are based on 189 observations.

The benefiting role of performance feedback on performance beliefs is illustrated in Figure 4.4. Here, it can be seen clearly that while there is some adjustment in beliefs, it is not perfect (if it was, observations would lie on the inverse 45 degree line). For overestimation of own and others' performance, there is adjustment even without feedback (dark dashed lines do not overlap with horizontal line), while this is not true for winning probability. More importantly, the adjustment is stronger if feedback is provided, as the comparisons of dark and light dashed lines in all three graphs indicate. Thus, providing feedback leads to sensible changes in beliefs.

To sum up, the analysis of individuals' beliefs demonstrates that there is systematic heterogeneity in beliefs about own performance, others' performance, and winning probability. While changes in beliefs are not influenced by the composition of the reference group, revealing competition neglect, they are positively affected by performance feedback, revealing overconfidence that can be ameliorated. Hence, these results establish that both of the mechanisms that are discussed to influence competition entry are prevalent in this experiment.[16]

[16] Note that in contrast to what other studies find, there are no gender differences in beliefs. For the interested reader, analyses of gender differences are provided in Appendix D.3.

4.5.2 Competition Decisions

After having explored individuals' beliefs, I now turn to the main dependent variable, decisions to compete, and thus to the main focus of this paper, namely how overconfidence and competition neglect influence selection into competition. Again, I investigate this question by first providing a general impression of competition decisions in my sample before systematically investigating determinants of competition choices.

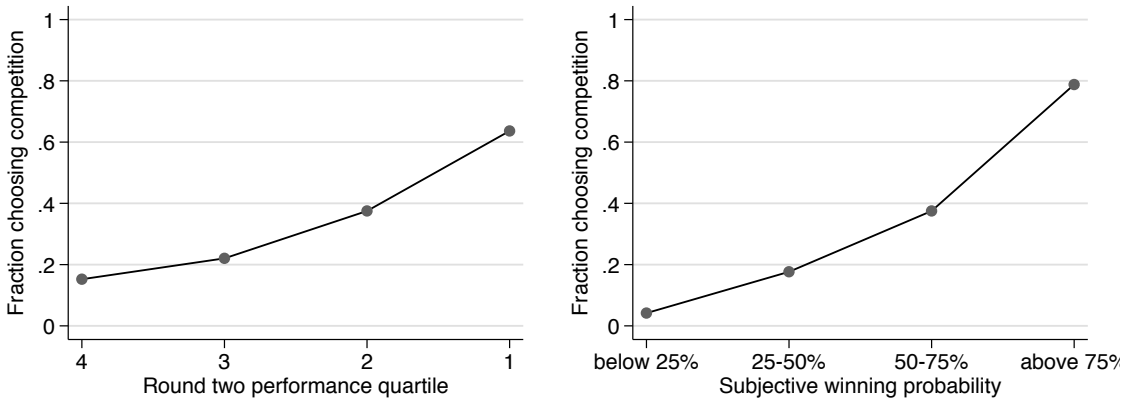
To start, I assess whether there is excess entry in my setup. Across all treatments, 33.33% of individuals choose the competitive incentive scheme. Considering the actual distribution of performances in the reference group, a payoff-maximizing individual should choose competition if her round three performance equals at least eight correctly solved items.[17] Assuming that performances in rounds two and three are equal, 35.98% of individuals have higher expected earnings when choosing competition. Thus, I do not find excess entry compared this benchmark in my sample.

However, these comparisons are based on the assumption of payoff-maximization and thus are rather strict. Therefore, I next explore competition decisions depending on round two performance and beliefs about winning probability. The left panel of Figure 4.5 depicts the proportion of participants selecting into competition by performance quartile in the initial competition. Contrary to what Niederle and Vesterlund (2007) find, there is a positive relationship between performance quartile and proportion of competition decisions. The right panel of Figure 4.5 depicts the proportion of participants selecting into competition by subjective probability of winning. There is a positive relationship between beliefs and the proportion of subjects selecting into competition that is even stronger than for previous performance. Thus, individuals' competition decisions in my sample are closely tied to their previous performance and beliefs.

To answer my central question whether competition decisions are influenced by overconfidence and competition neglect, I explore competition decisions by treatment; if competition neglect plays a role, differences between the *None* and *Selection* as well

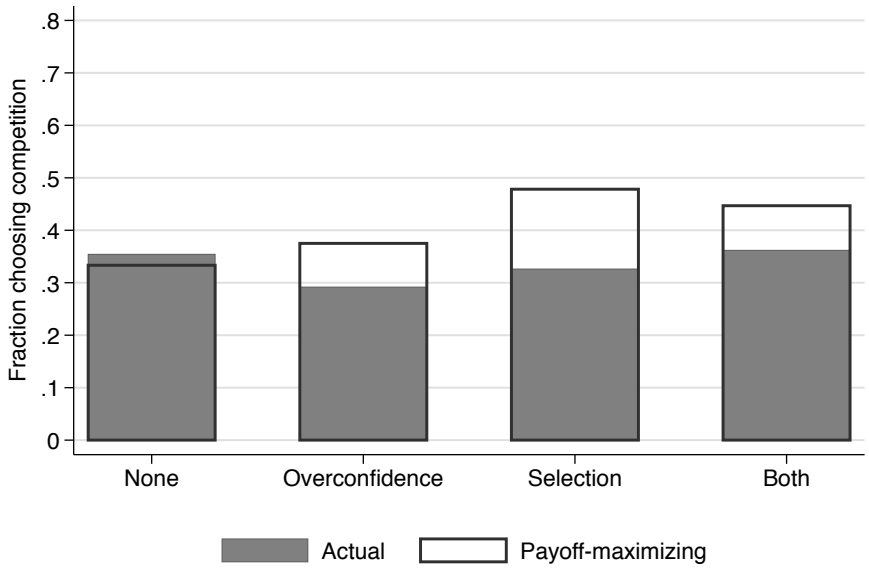
[17] Figure D.1 in Appendix D.2 plots actual chances of winning depending on performance. It can be seen that the winning probability exceeds 50% for more than seven correctly solved items.

Figure 4.5: *Competition choices by previous performance and beliefs*



Notes: Figure depicts fractions of individuals choosing the competitive incentive scheme in round three of the main sessions by performance quartile in round two (left panel) and subjective winning probability (right panel). Calculations are based on 189 observations.

Figure 4.6: *Competition choices by treatment*



Notes: Figure depicts fraction of participants choosing the competitive incentive scheme in round three of the main sessions by treatment condition. Calculations are based on 189 observations.

as between the *Overconfidence* and *Both* treatments should be smaller than expected. If errors in performance beliefs influence decisions, this should be reflected in the difference between treatments with and without feedback as well as in the influence of beliefs. The fraction of individuals selecting the competitive incentive scheme by treatment as well as the fraction of individuals for which it is the payoff-maximizing choice are depicted in Figure 4.6. Apart from the fact that absolute ratios of subjects deciding to compete seem pretty low when compared to previous literature[18] and are lower than the payoff-maximizing benchmark, there are no striking differences in competition rates between the treatments. Testing for gender differences within treatment and for treatment differences with Fisher exact tests does not detect significant effects.

Next, I take a closer look at determinants of competition decisions. The results of linear probability model estimations with decisions to compete as dependent variable are shown in Table 4.3. First, it can be seen that the main effects for performance feedback as well as for facing the self-selected competition group are nonsignificant, which is not surprising given Figure 4.6. Second, as indicated by Figure 4.5, previous performance has a positive effect on the probability to choose the competitive incentive scheme. Additionally, while the gender and competition literature has established that there is a gender difference in competitiveness (Niederle and Vesterlund, 2011), I find no effect of gender in my sample, while risk aversion has a significantly negative effect on selection into competition.[19] In columns (3) and (4), performance beliefs are included in the model. Now, the coefficients for all three performance assessments are significant, indicating that these beliefs influence competition decisions beyond performance. More specifically, holding actual round two performance constant, participants are more likely to select the competitive incentive scheme if they believe to have a higher performance, if they assess their competitors' performance to be lower, and if they believe to have a higher probability of winning. Thus, overly optimistic beliefs indeed influence participants' competition decisions.

[18] In Niederle and Vesterlund (2007), 73% of men and 35% of women enter. While their setup is slightly different from mine as participants have to outperform three others, studies also employing the two-person competition with a randomly drawn competitor find similar entry rates; for instance, in Berlin and Dargnies (2016), 63.2% of men and 35.1% of women initially chose to compete.

[19] A separate analysis of treatment effects for men and women is provided in Appendix D.3.

Table 4.3: *Determinants of competition choices*

DV: Competition (Y/N)	(1)	(2)	(3)	(4)
Feedback = 1	-0.00974	0.00317	0.00333	0.00785
	(0.0634)	(0.0656)	(0.0591)	(0.0598)
Selection = 1	-0.0519	-0.0582	-0.0311	-0.0357
	(0.0627)	(0.0665)	(0.0567)	(0.0584)
Previous performance	0.0581***	0.0613***	-0.0266	-0.0255
	(0.0108)	(0.0114)	(0.0235)	(0.0234)
Own performance belief			0.107***	0.112***
			(0.0306)	(0.0313)
Others' performance belief			-0.0797***	-0.0799***
			(0.0271)	(0.0291)
Subj. winning probability			0.628***	0.620***
			(0.168)	(0.176)
Female = 1	-0.0346	-0.00442	0.0294	0.0538
	(0.0659)	(0.0694)	(0.0587)	(0.0606)
Risk Aversion	-0.0438**	-0.0424**	-0.0305*	-0.0291*
	(0.0176)	(0.0175)	(0.0163)	(0.0166)
Constant	0.259*	0.750***	0.112	0.433
	(0.156)	(0.262)	(0.190)	(0.276)
Controls	No	Yes	No	Yes
Observations	189	187	189	187
Adjusted R^2	0.153	0.151	0.306	0.304

Notes: Table reports results of linear probability estimations. Robust standard errors in parentheses. Estimations are based on round three. Controls include age, an indicator for field of study (business, economics, STEM, humanities, law, psychology, other social sciences, other), and math grade. *** p<0.01, ** p<0.05, * p<0.1.

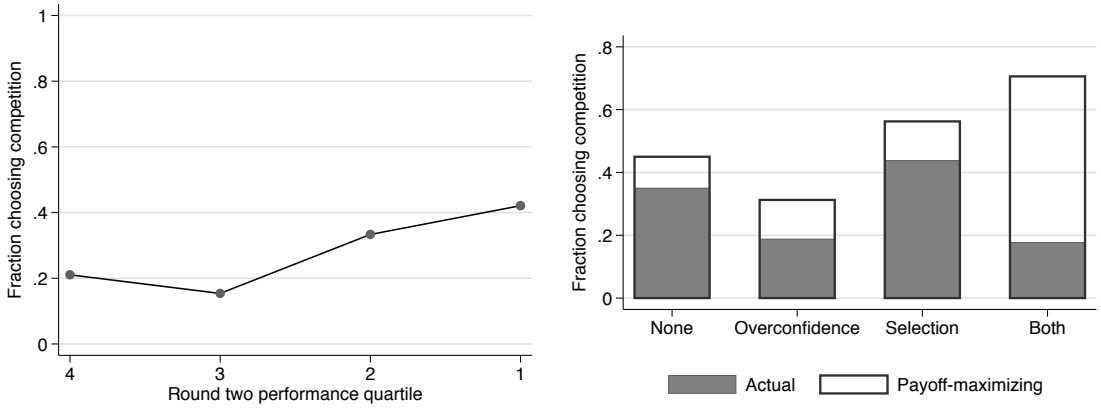
Regarding the central question asked in this paper, the following conclusions can be drawn based on the experimental results: as the proportion of participants selecting into the competitive incentive scheme does not depend on the composition of the competition group, competition neglect increases entry into competition. In addition, I also find evidence for the positive influence of absolute and relative overconfidence on entry into competition. Performance feedback does ameliorate bias in performance beliefs, but does not influence competition decisions directly.

4.5.3 The Effect of Pre-Treatment Belief Elicitation

In light of the vast literature on competitiveness, where competition rates are excessive and overconfidence is prevalent on average, it is surprising that in my setup, belief calibration works reasonably well, and that competition decisions are tied closely to performance and expectations. However, one clear point of departure of the present study is that the belief elicitation is conducted with three questions and that the belief questions are asked before the task in every round, making them salient. In this way, participants are forced to think very carefully about how they perceive their performance in the task relative to others, while other studies rely on eliciting the beliefs only after participants have made their decisions. Put differently, the belief elicitation procedure itself could cause subjects to think deliberately about the aspects relevant to the decision while not asking allows them to make intuitive "gut decisions". This is in line with dual-process theories of higher cognition in psychological research (see Kahneman, 2011; Evans, 2008; Evans and Stanovich, 2013).

To address the fact that competition decisions in the main sessions looked starkly different from the findings in the literature and to explore whether the belief elicitation procedure in itself might already change the way the competition decision is approached, I conducted four additional sessions. These sessions have the exact same setup as the main sessions, with the exception that the belief elicitation in rounds one and two is dropped so that beliefs are now elicited for the first time only after the competition choice is made. However, note that the sessions conducted so far only yield 69 additional observations (44 women and 25 men). Thus, the results presented in this Section

Figure 4.7: *Competition choices in additional sessions*



Notes: Figure depicts fraction of men and women choosing the competitive incentive scheme in round three of the additional sessions by round two performance quartile (left panel) and by Treatment (right panel). Calculations are based on 25 men and 44 women.

are preliminary and have to be interpreted with caution.

Of course, this change in procedure makes it impossible to explicitly investigate changes in beliefs as the pre-decision measures are missing. However, I am still able to investigate the relationship between competition decisions and performance and to compare competition decisions across treatments. These relationships are illustrated in Figure 4.7. As can be seen in the left panel, exploring competition choices depending on previous performance reveals a different picture than in the main sessions: there is no clear positive relationship between performance quartile and the proportion of individuals selecting into competition.

In a next step, the fraction of individuals selecting into competition depending on the treatment conditions is depicted in the right panel of Figure 4.7. Again, there is less entry than would be payoff-maximizing, and the general pattern suggests that overconfidence as well as competition neglect influence choices. Giving feedback seems to reduce the fraction of individuals selecting the competitive incentives, and the composition of the competition group does not have a clear effect – while there is a slight increase from *None* to *Selection*, the opposite is true for *Overconfidence* to *Both*. However, the fraction of payoff-maximizing competition entries does vary substantively, making clear statements about overconfidence and competition neglect difficult. While this

seems puzzling, keep in mind that treatments were between-subject and the sample size is rather small, making differences between the groups more likely. This is also why I refrain from investigating determinants of competition choices using a regression framework for now.

Finally, to investigate whether the results obtained in these additional sessions are indeed closer to the competition patterns that are found in the literature, I compare selection into competition by men and women. Overall, 44% of men choose the competitive incentive scheme, while only 20.45% of women do, making the difference in choices more extreme compared to the sessions with belief elicitation in every round. Indeed, the difference in competition choices between men and women is significant for the additional sessions (two-sided Fisher exact test: p=.054). In *Selection*, which is closest to the setting used in the gender and competition literature, the difference is even more pronounced: 66.67% of men choose to compete while only 30% of women do.[20]

To sum up, the additional sessions conducted without belief elicitation before the competition choice provide preliminary evidence for the hypothesis that the belief elicitation procedure itself might lead to changes in the way individuals process the decision.

4.6 Discussion

In this paper, I presented the results of a laboratory experiment designed to explore two distinct mechanisms underlying perception biases that influence selection into competition: overconfidence with respect to own and others' performance and competition neglect. In my setting, I find that there is systematic heterogeneity in perception biases, with low-performing individuals overestimating their own performance and their chances of success while underestimating performance in the competition group, while the opposite is true for high-performing individuals. These biases in performance beliefs are ameliorated by feedback; however, individuals persistently disregard the composition of the competition group they face. Investigating determinants of entry decisions to

[20] This is quite close to the 73% and 35% in Niederle and Vesterlund (2007) and the 63.2% and 35.1% in Berlin and Dargnies (2016).

tackle the key question of this paper whether competition neglect and overconfidence influence entry into competition, I find that both matter. However, decisions are closely tied to previous performance and assessments, and there are no gender differences.

Based on these findings, I argue that eliciting subjects' beliefs pre-treatment might have changed their mode of thinking, inducing them to think about the relevant evaluations more carefully rather than making intuitive decisions. Results from additional sessions where I refrain from eliciting subjects' beliefs pre-treatment are closer to the existing literature: in these sessions, there is a much more pronounced difference between competition choices of men and women, and choices are not as closely tied to performance.

However, this comparison implies that in the main sessions, competition neglect persists even though the belief elicitation procedure points individuals to the fact that they should evaluate their own as well as competitors' performance and factor those into their chances of winning. While this changes their choices and ties them closer to performance, it does not help to address competition neglect, which appears to be a persistent phenomenon. Hence, educating individuals about the aspects that matter for their decision does lead to less excessive entry, but does not ameliorate competition neglect. To address this bias, other, more direct ways of educating decision makers have to be applied. While this paper has mainly established competition neglect as a relevant bias in a laboratory setting, it is up to future research to come up with ways to specifically address this bias and to explore its impact in real-world settings.

On a more general note, based on these results, it seems important to keep in mind the mode of thought in which competition decisions are made. One might argue that while decisions made in the laboratory involve small stakes and are thus more prone to be gut decisions, decisions like career choices naturally involve more careful thought and are thus more closely tied to actual performance. Hence, laboratory findings might overestimate the differential in competitiveness between women and men. However, to be able to make detailed statements about the importance of the role of the mental resources deployed for decision making, one would have to design a mechanism experiment that explicitly manipulates the mode of processing. I leave this for future research.

Appendix A

Complementarities of HRM Practices: A Case for a Behavioral Economics Perspective

A.1 Detailed Exposition of the Model

Benchmark case

The principal faces the following optimization problem when implementing a_H:

$$
\begin{aligned}
\max_{w,b,\hat{a}} \quad & (1-\pi_H)\,(q_1 - w) + \pi_H\,(q_2 - w - b) \\
s.t. \quad & (1-\pi_H)\,w + \pi_H(w+b) - c + \eta\,[(1-\hat{\pi})\,w + \hat{\pi}(w+b) - c(\hat{a})]\,ER(a_H) \\
& \qquad \geq 0, && \text{(IR)} \\
& (1-\pi_H)\,w + \pi_H(w+b) - c + \eta\,[(1-\hat{\pi})\,w + \hat{\pi}(w+b) - c(\hat{a})]\,ER(a_H) && \text{(IC)} \\
& \qquad \geq (1-\pi_L)w + \pi_L(w+b) + \eta\,[(1-\hat{\pi})\,w + \hat{\pi}(w+b) - c(\hat{a})]\,ER(a_L), \\
& \quad \text{and} \quad \hat{a} = a_H. && \text{(EB)}
\end{aligned}
$$

The first and second constraint represent the standard individual rationality (IR) and incentive compatibility (IC) restrictions, respectively. The third constraint represents the restriction that beliefs have to be in equilibrium (EB). More precisely, this means that the agent can reasonably expect the kindness of the proposed contract offer as his expected rent when choosing action $\hat{a}$. Thus, destructive and "babbling" equilibria are ruled out. However, note that it is sufficient to assume (EB) and solve to implement $\hat{a} = a_H$ (cf. the discussion in Englmaier and Leider, 2012a).

For deriving the optimal contract, it is helpful to first think about the properties of the constraints at the optimum. Considering (IR), it is straightforward to see that this constraint has to be binding: suppose (IR) was slack.[1] Then, the principal could decrease w by the amount $\epsilon > 0$. If ϵ is small enough, the principal still manages to fulfill (IR) and at the same time increases her expected outcome without violating the (IC), because the (IC) depends on the bonus b (and thus on the wage spread between the states) which does not change. Thus, (IR) cannot be optimally slack.

Since (IR) is binding, the agent does not receive a rent, i.e., the reciprocal part of his utility is equal to zero and hence drops out of the problem. Thus, the optimal contract can easily be derived by solving the constraints of the standard problem for w and b.

Now, consider the (IC). It can easily be seen that since the agent is risk-neutral and does not care about the wage spread, there exist an infinite number of optimal contracts implementing a_H at first-best cost for the principal as long as $b \geq c/\Delta\pi$, where $\Delta\pi = \pi_H - \pi_L$. For reasons of simplicity and to make the first and second best solutions comparable, we focus on the case that the wages are set to solve (IC) with equality.

Limited liability

Now, the optimization problem of the principal implementing a_H (when again assuming (EB) and solving to implement $\hat{a} = a_H$) takes the following form:

$$
\begin{aligned}
\max_{w,b} \quad & (1-\pi_H)\,(q_1 - w) + \pi_H\,(q_2 - w - b) \\
s.t. \quad & (1-\pi_H)w + \pi_H(w+b) - c + \eta\,[(1-\pi_H)\,w + \pi_H(w+b) - c]\,ER(a_H) \\
& \qquad \geq 0, && \text{(IR)} \\
& (1-\pi_H)w + \pi_H(w+b) - c + \eta\,[(1-\pi_H)\,w + \pi_H(w+b) - c]\,ER(a_H) \\
& \qquad \geq (1-\pi_L)w + \pi_L(w+b) + \eta\,[(1-\pi_H)\,w + \pi_H(w+b) - c]\,ER(a_L), && \text{(IC)} \\
& \text{and} \quad w \geq \underline{w}. && \text{(LL)}
\end{aligned}
$$

[1] Note, when we add limited liability as a source of inefficiency into the model below, (IR) may be optimally slack, as in Englmaier and Leider (2012a).

It is immediately clear that the principal is not able to implement the first best contract in this case because she would have to set $w < 0$, which would violate the limited liability constraint (LL). Thus, we have to derive the second best contracts, which we do by considering two cases – selfish and reciprocal agents.

Selfish agent. First, consider the optimal contract for a selfish agent ($\eta = 0$). Then, the optimization problem boils down to the standard case. Again considering the properties of the constraints in the optimum, it is easy to see that (LL) binds at the optimum. Additionally, (IC) is binding, which provides us with two equalities to solve for the optimal contract. The properties of this contract are summarized in the following proposition.

Proposition A.1. *Under hidden action with limited liability ($w \geq \underline{w} \geq 0$), the risk-neutral principal proposes the following contract to the risk-neutral, selfish agent ($\eta = 0$):*

$$w = \underline{w}, \qquad b = \frac{c}{\Delta\pi}, \qquad \textit{and} \qquad \hat{a} = a_H.$$

In doing so, she can induce the agent to exert a_H, but only at implementation cost

$$B^{SB}(a_H) = c + \underline{w} + \frac{\pi_L}{\Delta\pi}c.$$

The agent receives a limited liability rent:

$$U(a_H) = \underline{w} + \frac{\pi_L}{\Delta\pi}c.$$

In this case, the first-best solution can no longer be employed. The (LL) constraint limits the range of possible wage payments. Comparing the boni in the first and second best cases reveals that they do not differ ($b = c/\Delta\pi$ in both cases), which makes sense when considering that the (IC) is binding in both cases and thus the wage spread cannot be reduced further; however, the principal cannot charge negative wages, thus she has to increase the base wage w and consequently pay a rent to the agent. In other words, while the principal is able to use both carrot and stick in the first best case, she loses the opportunity of using the stick with limited liability and has to rely on the "carrot" only to incentivize the agent.

Reciprocal agent. Now, consider a reciprocal agent with $\eta > 0$. We first illustrate this situation by a verbal argument: since the agent is not only motivated by monetary incentives, but also by the reciprocal part of his utility that is triggered when he receives a rent exceeding his outside option, the principal has an additional instrument to induce a_H. Stated differently, the limited liability rent has an additional incentive component and helps the principal to align the agent's interests to some degree with her own.

Indeed, solving the optimization problem in the same fashion as the standard limited liability problem yields the same base wage w, but the bonus payment b, and thus also the wage differential between the states of the world, are reduced.[2]

A.2 Additive Formulation of Worker Type and HRM Practices

Consider a simple, linear relationship between inherent and influenceable reciprocity: $\eta = \eta_1 + \eta_2$. Then, one can simply exchange η in Proposition 1.2 and again derive the comparative statics of the optimal bonus payment to see how an increase in η_2, for instance by investment in team building, influences b:

$$\frac{\partial b}{\partial \eta_2} = -\frac{\Delta\pi\Delta ER}{\left(\Delta\pi + (\eta_1+\eta_2)\,\Delta ER\pi_H\right)^2}\left(\underline{w} + \frac{\pi_L}{\Delta\pi}c\right) < 0.$$

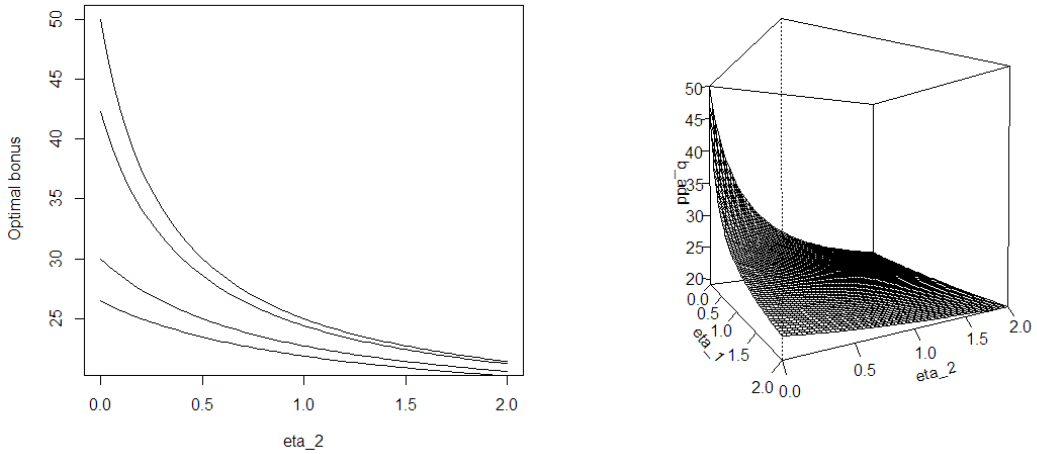
The second derivative as well as the cross-derivative are globally positive, indicating diminishing marginal returns of an increase in η:

$$\frac{\partial^2 b}{\partial \eta_2^2} = \frac{\partial^2 b}{\partial \eta_1 \eta_2} = \frac{\Delta\pi\left(\Delta ER\right)^2\pi_H}{\left(\Delta\pi + (\eta_1+\eta_2)\,\Delta ER\pi_H\right)^3}\left(\underline{w} + \frac{\pi_L}{\Delta\pi}c\right) > 0.$$

In this case, the two components of the reciprocity term act as substitutes, i.e., the firm can lower the bonus either by hiring workers with an inherent concern for reciprocity in the first place, or by triggering reciprocal concerns later on through investments in HRM practices. This relationship can also be seen in the plot of the optimal bonus dependent on η_1 and η_2 in Figure A.1.

[2] Note that the bonus is decreasing in η. Assuming $c \geq \underline{w}$ is sufficient to ensure a positive b.

Figure A.1: *Optimal bonus dependent on η in additive model*



Notes: The figure depicts the optimal bonus as function of η_1 and η_2, the parameters capturing the innate reciprocal inclination of a worker and the HRM policies targeted at increasing this reciprocal inclination towards the firm, for an additive formulation $\eta = \eta_1 + \eta_2$.

Appendix B

What Drives Reciprocal Behavior? The Optimal Provision of Incentives over the Course of Careers

B.1 Proofs

Proof of Lemma 2.1

Proof. Plugging the agent's optimal effort choice, $e^* = \sqrt{b + \eta w \theta}$, into the principal's profits gives the Lagrange function $\mathcal{L} = \sqrt{b + \eta w \theta}\,(\theta - b) - w + \lambda_b b + \lambda_w w$ and first order conditions

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial b} =& \frac{1}{2\sqrt{b + \eta w \theta}}\,(\theta - b) - \sqrt{b + \eta w \theta} + \lambda_b = 0 \\ \frac{\partial \mathcal{L}}{\partial w} =& \frac{\eta \theta}{2\sqrt{b + \eta w \theta}}\,(\theta - b) - 1 + \lambda_w = 0\end{aligned}$$

We first show that bonus and wage payments are not used simultaneously. To the contrary, assume this as the case, i.e. that $\lambda_b = \lambda_w = 0$. Then, first-order conditions are

$$\begin{aligned}\frac{1}{2\sqrt{b + \eta w \theta}}\,(\theta - b) - \sqrt{b + \eta w \theta} = 0 \\ \frac{\eta \theta}{2\sqrt{b + \eta w \theta}}\,(\theta - b) - 1 = 0\end{aligned}$$

Second-order conditions will not hold in this case, though.

The Hessian matrix of second-order partial derivatives equals

$$\begin{pmatrix} -\frac{(b+\eta w\theta)^{-3/2}}{4}(\theta-b)-\frac{1}{\sqrt{b+\eta w\theta}} & -\frac{(b+\eta w\theta)^{-3/2}\eta\theta}{4}(\theta-b)-\frac{\eta\theta}{2\sqrt{b+\eta w\theta}} \\ -\frac{(b+\eta w\theta)^{-3/2}\eta\theta}{4}(\theta-b)-\frac{\eta\theta}{2\sqrt{b+\eta w\theta}} & -\frac{\eta^2\theta^2(b+\eta w\theta)^{-3/2}}{4}(\theta-b) \end{pmatrix},$$

yielding a determinant equal to $-\frac{3}{2}\eta\theta < 0$. Hence, either $w = 0$ or $b = 0$.

First, assume that $w = 0$ and $b > 0$, i.e., $\lambda_b = 0$ and $\lambda_w > 0$. Then, only the principal's first first-order condition is relevant and yields $b = \theta/3$. Furthermore, effort is $e^* = \sqrt{\theta/3}$, profits are $\pi = \frac{2\theta}{3}\sqrt{\frac{\theta}{3}}$, and the agent's utility equals $u = \frac{2\theta}{9}\sqrt{\frac{\theta}{3}} > 0$.

Now, assume that $w > 0$ and $b = 0$, i.e., $\lambda_b > 0$ and $\lambda_w = 0$. Then, only the principal's second first-order condition is relevant and yields $w = \eta\theta^3/4$. Furthermore, effort is $e^* = \eta\theta^2/2$, profits are $\pi = \frac{\eta\theta^3}{4}$, and the agent's utility equals $u = \frac{\eta\theta^3}{4} + \frac{\eta^3\theta^6}{12} > 0$.

Note that the second-order conditions hold in both cases.

Profits using a wage contract are higher than using a bonus contract, if $\eta^2\theta^3 \geq \frac{64}{27}$, hence the proclaimed threshold $\overline{\eta}$ exists, with $\overline{\eta} = \sqrt{\frac{64}{27\theta^3}}$. □

Proof of Lemma 2.2

Proof. If the (DE) constraint does not bind in a period t, the principal maximizes profits $\pi_t = e_t^*\theta - \left((e_t^*)^3/3 - \eta w_t e_t^*\theta + 2/3\left(\sqrt{\eta w_t\theta}\right)^3\right) - w_t$, subject to $w_t \geq 0$.

The Lagrange function equals $\mathcal{L}_t = e_t^*\theta - (e_t^*)^3/3 + \eta w_t e_t^*\theta - 2/3\left(\sqrt{\eta w_t\theta}\right)^3 - w_t + \lambda_{w_t} w_t$, where $\lambda_{wt} \geq 0$ represents the Lagrange parameter for the agent's limited liability constraint, giving first order conditions

$$\frac{\partial\mathcal{L}_t}{\partial e_t^*} = \theta - (e_t^*)^2 + \eta w_t\theta = 0$$
$$\frac{\partial\mathcal{L}_t}{\partial w_t} = \eta\theta\left(e_t^* - \sqrt{\eta w_t\theta}\right) - 1 + \lambda_{w_t} = 0$$

First, we show that $\lambda_{w_t} = 0$. To the contrary, assume that $\lambda_{w_t} > 0$ and hence $w_t = 0$. Then, $e_t^* = \sqrt{\theta}$ and $\pi_t = \frac{2}{3}\left(\sqrt{\theta}\right)^3$. In this case, a small increase of the wage would raise profits: $\frac{d\pi_t}{dw_t}\,|_{w_t=0} = \sqrt{\eta^2\theta^3} - 1 > 0$, since $\eta > \overline{\eta}$ implies that $\eta^2\theta^3 > 64/27 > 1$.

Since $\lambda_{w_t} = 0$, the first order conditions allow us to obtain the values for effort and wage, yielding $w_t = \frac{\left(\eta^2\theta^3-1\right)^2}{4\eta^3\theta^3}$ and $e_t^* = \frac{1+\eta^2\theta^3}{2\eta\theta}$. $w_t > 0$ because $\eta > \overline{\eta}$ implies $\eta^2\theta^3 > 1$. □

Proof of Lemma 2.3

Proof. Taking the (DE) constraint into account, the Lagrange function of the principal's maximization problem in a period t becomes $\mathcal{L}_t = e_t^*\theta - (e_t^*)^3/3 + \eta w_t e_t^*\theta - 2/3\left(\sqrt{\eta w_t\theta}\right)^3 - w_t + \lambda_{DE_t}\left[\delta\left(\Pi_{t+1} - \tilde{\Pi}_{t+1}\right) - \frac{2}{3}\left(\sqrt{\eta w_t\theta}\right)^3 - (e_t^*)^3/3 + \eta w_t\theta e_t^*\right]$, where $\lambda_{DE_t} \geq 0$ represents the Lagrange parameter for the principal's dynamic enforcement constraint, and where we omit the agent's limited liability constraint and show ex-post that is satisfied.

First order conditions are

$$\frac{\partial\mathcal{L}}{\partial e_t^*} = \theta - (e_t^*)^2 + \eta w_t\theta + \lambda_{DE_t}\left[-(e_t^*)^2 + \eta w_t\theta\right] = 0$$
$$\frac{\partial\mathcal{L}}{\partial w_t} = \eta\theta e_t^* - \eta\theta\sqrt{\eta w_t\theta} - 1 + \lambda_{DE_t}\left[-\eta\theta\sqrt{\eta w_t\theta} + \eta\theta e_t^*\right] = 0,$$

yielding $w_t = \frac{\left(\eta^2\theta^3\left(1+\lambda_{DE_t}\right)-1\right)^2}{4\eta^3\theta^3\left(1+\lambda_{DE_t}\right)^2}$ and $e_t^* = \frac{1+\eta^2\theta^3\left(1+\lambda_{DE_t}\right)}{2\eta\theta\left(1+\lambda_{DE_t}\right)}$. $w_t > 0$ for $\lambda_{DE_t} \geq 0$ because $\eta > \overline{\eta}$ implies $\eta^2\theta^3 > 1$. Finally, it is straightforward to show that for $\lambda_{DE_t} > 0$, $\frac{\left(\eta^2\theta^3\left(1+\lambda_{DE_t}\right)-1\right)^2}{4\eta^3\theta^3\left(1+\lambda_{DE_t}\right)^2} > \frac{\left(\eta^2\theta^3-1\right)^2}{4\eta^3\theta^3}$ and $\frac{1+\eta^2\theta^3\left(1+\lambda_{DE_t}\right)}{2\eta\theta\left(1+\lambda_{DE_t}\right)} < \frac{1+\eta^2\theta^3}{2\eta\theta}$. □

Proof of Lemma 2.4

Proof. The (DE) constraint in period $T-1$ (where on- and off-path continuation profits are $\Pi_T = \eta\theta^3/4$ and $\tilde{\Pi}_T = \frac{2\theta}{3}\sqrt{\frac{\theta}{3}}$, respectively) equals $(e_t^*)^3/3 - \eta w_t\theta e_t^* \leq \delta\left(\frac{\eta\theta^3}{4} - \frac{2\theta}{3}\sqrt{\frac{\theta}{3}}\right) - \frac{2}{3}\left(\sqrt{\eta w_t\theta}\right)^3$. For first-best effort and wage levels $w_t = \frac{\left(\eta^2\theta^3-1\right)^2}{4\eta^3\theta^3}$ and $e_t^* = \frac{1+\eta^2\theta^3}{2\eta\theta}$, it becomes

$$\frac{3\eta^2\theta^3 - 1}{6\eta^3\theta^3} \leq \delta\left(\frac{\eta\theta^3}{4} - \frac{2\theta}{3}\sqrt{\frac{\theta}{3}}\right).$$

By assumption ($\eta > \overline{\eta}$), both left and right hand side are strictly positive. Therefore, the constraint is violated for first-best effort and wage levels if $\delta \to 0$.

To investigate the first part of the Lemma, rewriting the (DE) constraint gives $\frac{3-\frac{1}{\eta^2\theta^3}}{6\eta} \leq \delta\sqrt{\theta^3}\left(\frac{\eta}{4} - \sqrt{\frac{4}{27}}\right)$. Therefore, the left hand side is decreasing and the right hand side

increasing in η and θ.

However, we have imposed the assumptions $\theta < 3$ and $\frac{\eta\theta^2}{2} < 1$ in order to guarantee an interior solution. This gives an upper bound for η, $\eta = \frac{2}{\theta^2}$. Plugging this upper bound into the constraint yields $\frac{12-\theta}{48}\theta \leq \delta\left(\frac{1}{2} - \frac{2}{3}\sqrt{\frac{\theta}{3}}\right)$. This is satisfied for any positive δ provided θ is sufficiently small. Finally, note that in this case, the condition $\eta > \overline{\eta}$ also holds for θ sufficiently small. Plugging the upper bound $\eta = \frac{2}{\theta^2}$ into this condition (i.e., $\eta > \sqrt{\frac{64}{27\theta^3}}$) yields $27 > 16\theta$.

Concerning the second part, recall that the equilibrium is sequentially efficient, hence, the principal's maximization problem is equivalent to maximizing $\pi_t = e_t^*\theta - b_t^* - w_t$ in every period t, subject to the (DE) constraint $(e_t^*)^3/3 - \eta w_t \theta e_t^* \leq \delta\left(\Pi_{t+1} - \tilde{\Pi}_{t+1}\right) - \frac{2}{3}\left(\sqrt{\eta w_t \theta}\right)^3$. Defining $\Delta_t \equiv \Pi_t - \tilde{\Pi}_t$, implementable effort in period t is ceteris paribus strictly increasing in Δ_{t+1}, whereas per-period profits π_t are consequently weakly increasing in Δ_{t+1}. Furthermore, per-period profits in periods $t < T$ can solely be expressed as functions of Δ_{t+1}, i.e. $\pi_t(\Delta_{t+1})$, with $\pi_t' \geq 0$.

The profit-maximizing spot reciprocity contract is the principal's optimal choice in the last period T, hence $\pi_T = \Pi_T = \eta\theta^3/4$. In all previous periods, the principal still has the option to implement the spot reciprocity contract (by setting $b_t^* = 0$ and $w_t = \eta\theta^3/4$), therefore $\pi_t \geq \pi_T \, \forall t$. In addition, off-path profits are determined by a bonus spot contract, hence $\tilde{\pi} = \frac{2\theta}{3\sqrt{\frac{\theta}{3}}}$ in every period.

Now, we can apply proof by induction to verify that $\Delta_{t-1} > \Delta_t$. We start with the last periods in order to show that $\Delta_{T-1} > \Delta_T$:

$$\Delta_{T-1} = \pi_{T-1} - \tilde{\pi} + \delta\Delta_T \geq \pi_T - \tilde{\pi} + \delta\Delta_T = \Delta_T(1+\delta) > \Delta_T.$$

Now, assume that $\Delta_t > \Delta_{t+1}$. Since $\pi_t'(\Delta_{t+1}) \geq 0$, $\pi_{t-1} \geq \pi_t$. Therefore, $\Delta_{t-1} = \pi_{t-1} - \tilde{\pi} + \delta\Delta_t \geq \pi_t - \tilde{\pi} + \delta\Delta_t > \pi_t - \tilde{\pi} + \delta\Delta_{t+1} = \Delta_t$, which completes the proof. □

Proof of Proposition 2.1

Proof. In Lemmas 2.2 and 2.3, we established that $w_t = \frac{\left(\eta^2\theta^3\left(1+\lambda_{DE_t}\right)-1\right)^2}{4\eta^3\theta^3\left(1+\lambda_{DE_t}\right)^2}$ and $e_t^* = \frac{1+\eta^2\theta^3\left(1+\lambda_{DE_t}\right)}{2\eta\theta\left(1+\lambda_{DE_t}\right)}$, where λ_{DE_t} is the Lagrange parameter associated with the (DE) constraint in period t. Hence, $w_t = w_{t-1}$ and $e_t^* = e_{t-1}$ if $\lambda_{DE_t} = \lambda_{DE_{t-1}} = 0$. By Lemma 2.3, if $\lambda_{DE_{t-1}} = 0$ but $\lambda_{DE_t} > 0$, then $w_t > w_{t-1}$ and $e_t^* < e_{t-1}$. Finally, assume that $\lambda_{DE_{t-1}} > 0$. First, we show that in this case also $\lambda_{DE_t} > 0$: Plugging $w_{t-1} = \frac{\left(\eta^2\theta^3\left(1+\lambda_{DE_{t-1}}\right)-1\right)^2}{4\eta^3\theta^3\left(1+\lambda_{DE_{t-1}}\right)^2}$ and $e_{t-1}^* = \frac{1+\eta^2\theta^3\left(1+\lambda_{DE_{t-1}}\right)}{2\eta\theta\left(1+\lambda_{DE_{t-1}}\right)}$ into the binding (DE) constraint for period $t-1$ yields

$$\frac{3\eta^2\theta^3\left(1+\lambda_{DE_{t-1}}\right)-1}{6\eta^3\theta^3\left(1+\lambda_{DE_{t-1}}\right)^3} = \delta\left(\Pi_t - \tilde{\Pi}_t\right).$$

By the implicit function theorem, $\frac{d\lambda_{DE_{t-1}}}{d\left(\Pi_t-\tilde{\Pi}_t\right)} = \frac{2\delta\eta^3\theta^3\left(1+\lambda_{DE_{t-1}}\right)^4}{1-2\eta^2\theta^3\left(1+\lambda_{DE_{t-1}}\right)} < 0$ (since $\eta > \overline{\eta}$ implies $\eta^2\theta^3 > 1$). Hence, Lemma 2.4 yields $\lambda_{DE_{t-1}} < \lambda_{DE_t}$, which implies $\lambda_{DE_{t-1}} > 0 \Rightarrow \lambda_{DE_t} > 0$. Furthermore, if $\lambda_{DE_t} = 0$ in a period t, this also holds for all previous periods.

The wage schedule is increasing in periods $t < T$ since $\frac{\partial w_t}{\partial \lambda_{DE_t}} = \frac{\left(\eta^2\theta^3\left(1+\lambda_{DE_t}\right)-1\right)}{2\eta^3\theta^3\left(1+\lambda_{DE_t}\right)^3} > 0$, whereas the effort path is decreasing because of $\frac{\partial e_t^*}{\partial \lambda_{DE_t}} = \frac{-1}{2\eta\theta\left(1+\lambda_{DE_t}\right)^2} < 0$. Finally, wage and effort in period T are $e_T^* = \frac{\eta\theta^2}{2}$ and $w_T = \frac{\eta\theta^3}{4}$, respectively. $e_T^* < e_t^*$ for all $t < T$ follows from $\frac{\eta\theta^2}{2} < \frac{1+\eta^2\theta^3\left(1+\lambda_{DE_t}\right)}{2\eta\theta\left(1+\lambda_{DE_t}\right)}$ $(\Leftrightarrow \eta^2\theta^3\left(1+\lambda_{DE_t}\right) < 1+\eta^2\theta^3\left(1+\lambda_{DE_t}\right))$. $w_T > w_t$ for all $t < T$ follows from $\frac{\eta\theta^3}{4} > \frac{\left(\eta^2\theta^3\left(1+\lambda_{DE_t}\right)-1\right)^2}{4\eta^3\theta^3\left(1+\lambda_{DE_t}\right)^2}$ $(\Leftrightarrow 2\eta^2\theta^3\left(1+\lambda_{DE_t}\right) > 1)$, which completes the proof. □

Proof of Prediction 2.4

Proof. An agent's utility in a period $t < T$ is

$$\begin{aligned} U_t &= w_t + b_t - \frac{e_t^3}{3} + \eta w_t e_t^* \theta \\ &= \frac{\left(2+\lambda_{DE_t}\right) + 3\eta^4\theta^6\left(1+\lambda_{DE_t}\right)^3 - 3\eta^4\theta^6\left(1+\lambda_{DE_t}\right)^2 + \eta^6\theta^9\left(1+\lambda_{DE_t}\right)^3 - 3\eta^2\theta^3\left(1+\lambda_{DE_t}\right)}{12\eta^3\theta^3\left(1+\lambda_{DE_t}\right)^3} \end{aligned}$$

$\frac{\partial U_t}{\partial \eta} = \frac{3\eta^4\theta^6(1+\lambda_{DEt})^2\lambda_{DEt}+3\left(\eta^6\theta^9(1+\lambda_{DEt})^3-1\right)+3(1+\lambda_{DEt})\left(\eta^2\theta^3-1\right)}{12\eta^4\theta^3(1+\lambda_{DEt})^3} > 0$ yields that more reciprocal agents are more satisfied with their jobs.

$\frac{\partial U_t}{\partial \lambda_{DEt}} = \frac{3\left(\eta^4\theta^6(1+\lambda_{DEt})^2-1\right)+2(1+\lambda_{DEt})\left(3\eta^2\theta^3-1\right)}{12\eta^3\theta^3(1+\lambda_{DEt})^4} > 0$ (because $\eta > \overline{\eta}$ implies $\eta^2\theta^3 > 1$) yields that job satisfaction increases over time. Concerning the last period T, recall that in a spot reciprocity contract, $U = \frac{\eta\theta^3}{4} + \frac{\eta^3\theta^6}{12}$.

This is larger than the utility in previous periods if

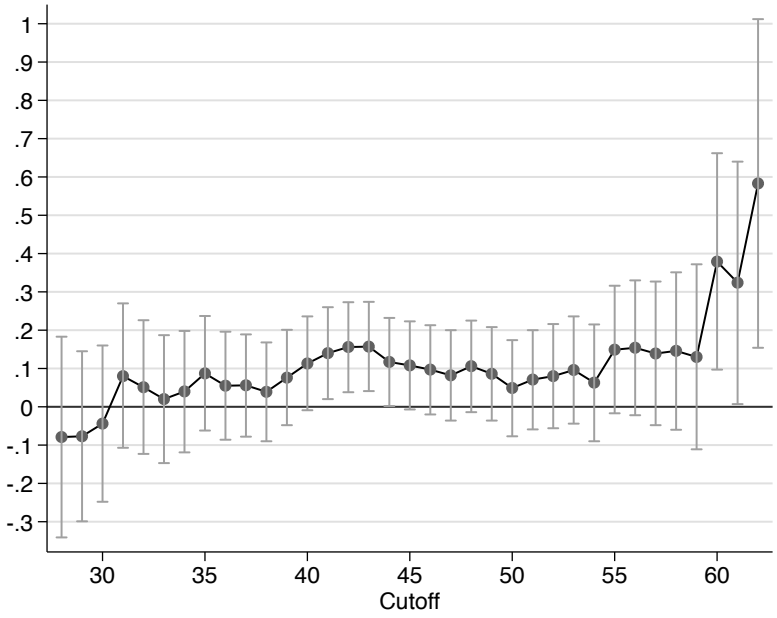
$$(1+\lambda_{DEt})\left(3\eta^2\theta^3-1\right) > 1 - 3\eta^4\theta^6\left(1+\lambda_{DEt}\right)^2$$

which always holds.

$\frac{\partial^2 U_t}{\partial \lambda_{DEt}\partial \eta} = \frac{(1+\lambda_{DEt})\left(\eta^2\theta^3-1\right)^2+\eta^4\theta^6(1+\lambda_{DEt})\lambda_{DEt}+(4+\lambda_{DEt})}{4\eta^4\theta^3(1+\lambda_{DEt})^4} > 0$ yields that the positive effect of reciprocity on job satisfaction increases over time. □

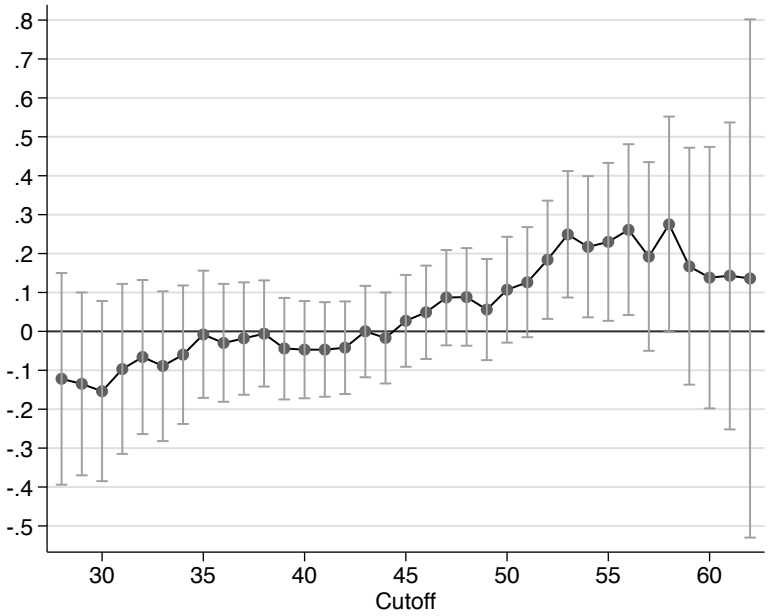
B.2 Figures and Tables

Figure B.1: *Interaction coefficients in overtime regressions*



Notes: Figure plots coefficient estimates of interaction between positive reciprocity and being at least the threshold value of years old. Error bars indicate 95% confidence intervals.

Figure B.2: *Interaction coefficients in satisfaction regressions*



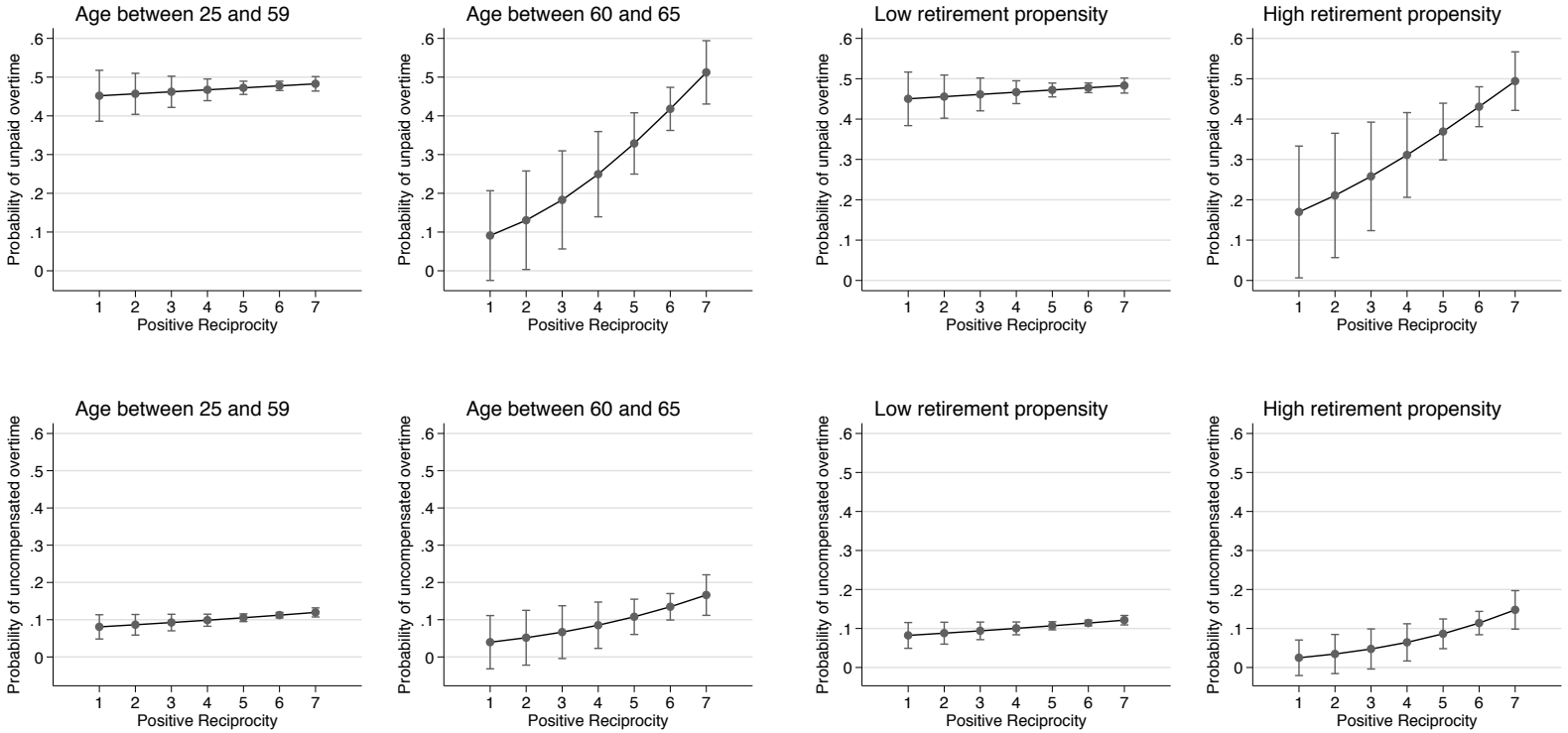
Notes: Figure plots coefficient estimates of interaction between positive reciprocity and being at least the threshold value of years old. Error bars indicate 95% confidence intervals.

Table B.1: *Effect of reciprocity on unpaid overtime*

	DV: Unpaid Overtime			DV: Uncompensated Overtime		
	(1)	(2)	(3)	(4)	(5)	(6)
	Overall	Age cutoff	Retirement propensity	Overall	Age cutoff	Retirement propensity
Positive reciprocity	0.00879	0.00973	0.00883	0.00507**	0.00494*	0.00504**
	(0.00652)	(0.00653)	(0.00652)	(0.00246)	(0.00253)	(0.00246)
Age (in years)	-0.00744***			0.000507		
	(0.00171)			(0.000712)		
Age $\geq$ 60		-0.0749**			0.0145	
		(0.0328)			(0.0142)	
Age $\geq$ 60# Reciprocity		0.0861***			0.0144	
		(0.0316)			(0.0122)	
Retire = 1			-0.0580**			-0.0348
			(0.0292)			(0.188)
Retire # Reciprocity			0.0566**			0.00137
			(0.0276)			(0.0104)
Observations	7,019	7,019	7,019	7,019	7,019	7,019
Pseudo R^2	0.0794	0.0786	0.0782	0.1824	0.1829	0.1826

Notes: Estimations are based on the 2005 survey wave of the SOEP. Robust standard errors in parentheses clustered at household level. Table reports marginal effects at means after logit regressions with marginal effects for interactions reflecting the difference in slope for reciprocity between the groups. Reciprocity measure standardized. Controls include years of education, gender, years of full time and part time work experience (linear and squared terms), a dummy for part-time employment, job tenure in current position (linear and squared), indicator variables for industry sector (services, agriculture, energy or mining, manufacturing, construction, trade, transport, bank or insurance) and firm size (less than 100, 100-199, 200-1999, more than 2000 employees), and an indicator variable for occupation status (white collar, blue collar, public sector). *** p<0.01, ** p<0.05, * p<0.1.

Figure B.3: *Predicted marginal effects for unpaid and uncompensated overtime*



Notes: The figure plots predicted average marginal changes of the propensity to work unpaid (upper panel) and uncompensated overtime (lower panel) at different levels of positive reciprocity depending on age group and a high vs. low propensity to retire within the next two years. Estimations are based on the 2005 survey wave of the SOEP. Error bars indicate 95% confidence intervals.

Table B.2: *Effect of reciprocity on overtime hours*

	DV: Total Hours			DV: Unpaid Hours		
	(1)	(2)	(3)	(4)	(5)	(6)
	Overall	Age cutoff	Retirement propensity	Overall	Age cutoff	Retirement propensity
Positive reciprocity	0.0973***	0.0871***	0.0898***	0.0637**	0.0536**	0.0553**
	(0.0205)	(0.0207)	(0.0206)	(0.0248)	(0.0252)	(0.0252)
Age (in years)	-0.0178***			-0.0188***		
	(0.00568)			(0.00695)		
Age $\geq$ 60		-0.245**			–0.327**	
		(0.116)			(0.130)	
Age $\geq$ 60 # Reciprocity		0.283**			0.306**	
		(0.119)			(0.125)	
Retire = 1			-0.246**			-0.281**
			(0.107)			(0.116)
Retire # Reciprocity			0.157			0.224*
			(0.106)			(0.115)
Observations	6,976	6,976	6,976	6,864	6,864	6,864
Pseudo R^2	0.0111	0.0110	0.0110	0.0132	0.0133	0.0133

Notes: Estimations are based on the 2005 survey wave of the SOEP. Robust standard errors in parentheses clustered at household level. Table reports coefficients of negative binomial regressions. Reciprocity measure standardized. Controls include years of education, gender, years of full time and part time work experience (linear and squared terms), a dummy for part-time employment, job tenure in current position (linear and squared), indicator variables for industry sector (services, agriculture, energy or mining, manufacturing, construction, trade, transport, bank or insurance) and firm size (less than 100, 100-199, 200-1999, more than 2000 employees), and an indicator variable for occupation status (white collar, blue collar, public sector). *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table B.3: *Effect of reciprocity on effort (2005, 2010, and 2015)*

DV: Overtime (Y/N)	(1) Overall	(2) Age cutoff = 60	(3) Age cutoff = 55	(4) RE: Overall	(5) RE: Age cutoff = 60	(6) RE: Age cutoff = 55
Positive reciprocity	0.00852**	0.00850**	0.00859**	0.00802**	0.00794**	0.00812**
	(0.00348)	(0.00349)	(0.00349)	(0.00330)	(0.00330)	(0.00330)
Age (in years)	-0.00284***			-0.00267**		
	(0.00108)			(0.00105)		
Age cutoff = 1		-0.0808***	-0.0219		-0.0776***	-0.0184
		(0.0178)	(0.0137)		(0.0175)	(0.0134)
Age cutoff # Reciprocity		0.0149	0.0156*		0.00821	0.0108
		(0.0136)	(0.00870)		(0.0130)	(0.00824)
Observations	25,653	25,653	25,653	25,653	25,653	25,653
Groups				18,128	18,128	18,128

Notes: Estimations are based on the 2005, 2010, and 2015 survey waves of the SOEP. Robust standard errors in parentheses clustered at household level. Table reports marginal effects at the mean of pooled logit regressions in columns (1) and (2) and coefficents of RE logit estimations in columns (3) and (4). Controls include years of education, gender, years of full time and part time work experience (linear and squared terms), a dummy for part-time employment, job tenure in current position (linear and squared), indicator variables for industry sector (services, agriculture, energy or mining, manufacturing, construction, trade, transport, bank or insurance) and firm size (less than 100, 100-199, 200-1999, more than 2000 employees), an indicator variable for occupation status (white collar, blue collar, public sector), and indicator variables for birth years. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table B.4: *Effect of reciprocity on overtime hours (2005, 2010, and 2015)*

	DV: Total Hours			DV: Unpaid Hours		
	(1)	(2)	(3)	(4)	(5)	(6)
	Overall	Age cutoff = 60	Age cutoff = 55	Overall	Age cutoff = 60	Age cutoff = 55
Positive reciprocity	0.147***	0.136***	0.122***	0.110***	0.0986***	0.0797**
	(0.0248)	(0.0249)	(0.0264)	(0.0297)	(0.0300)	(0.0323)
Age (in years)	-0.0515***			-0.0530***		
	(0.00706)			(0.00855)		
Age cutoff = 1		-0.645***	-0.542***		-0.739***	-0.507***
		(0.130)	(0.0974)		(0.142)	(0.110)
Age cutoff # Reciprocity		0.359***	0.188***		0.404***	0.218***
		(0.128)	(0.0701)		(0.133)	(0.0764)
Observations	15,883	15,883	15,883	15,771	15,771	15,771
Pseudo R^2	0.0149	0.0143	0.0143	0.0170	0.0165	0.0164

Notes: Estimations are based on the 2005, 2010, and 2015 survey waves of the SOEP. Robust standard errors in parentheses clustered at household level. Table reports coefficients of negative binomial regressions. Reciprocity measure standardized. Controls include years of education, gender, years of full time and part time work experience (linear and squared terms), a dummy for part-time employment, job tenure in current position (linear and squared), indicator variables for industry sector (services, agriculture, energy or mining, manufacturing, construction, trade, transport, bank or insurance) and firm size (less than 100, 100-199, 200-1999, more than 2000 employees), and an indicator variable for occupation status (white collar, blue collar, public sector). *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Appendix C

Individual Differences and Contribution Sequences in Threshold Public Goods

C.1 Figures and Tables

Figure C.1: *Distribution of HEXACO scores*

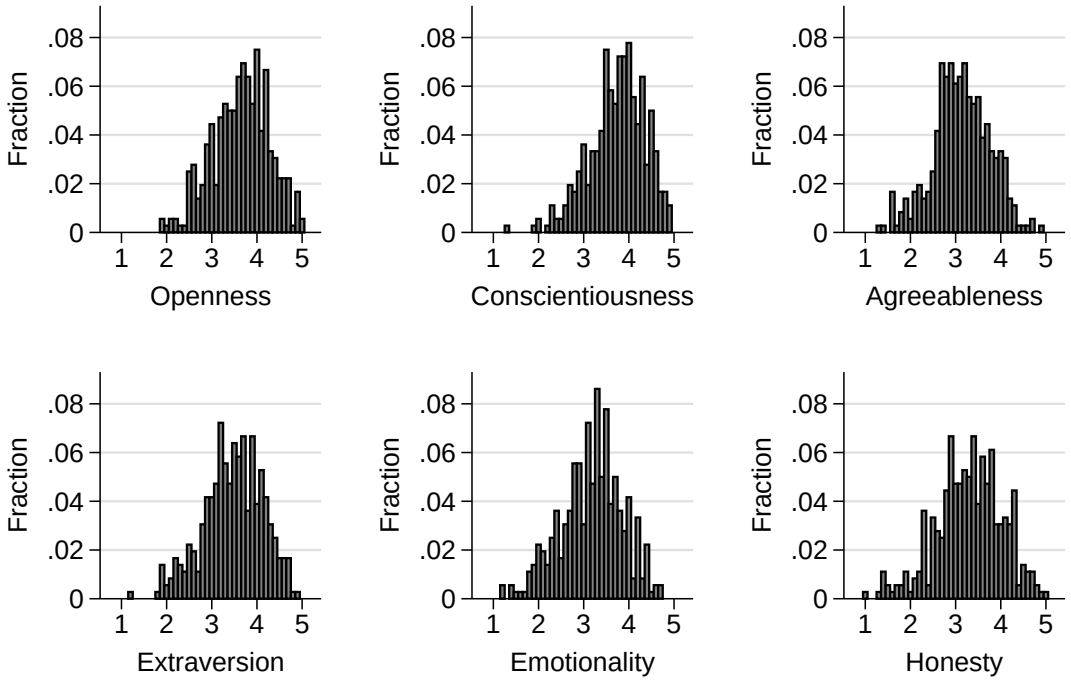


Figure C.2: *Distribution of SVO angles*

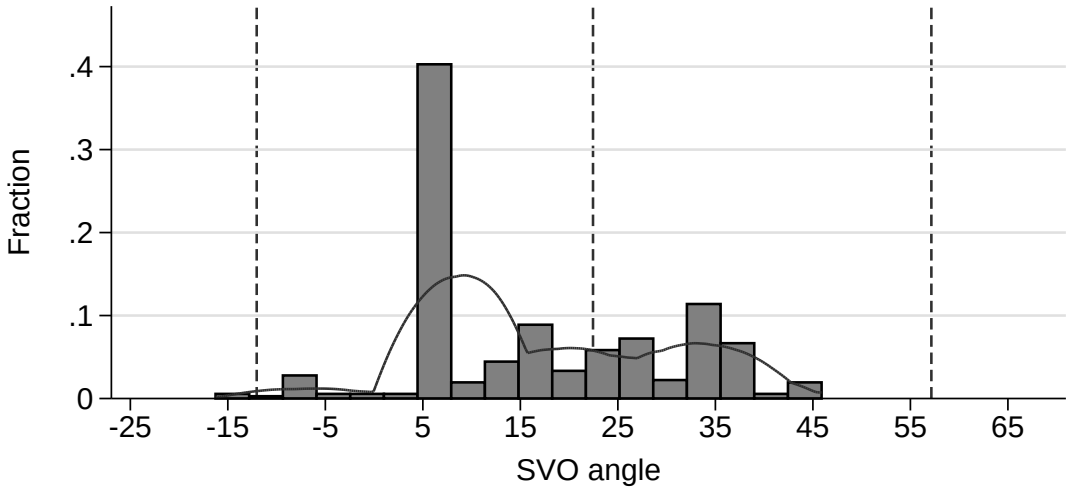


Figure C.3: *Distribution of trust and reciprocity scores*

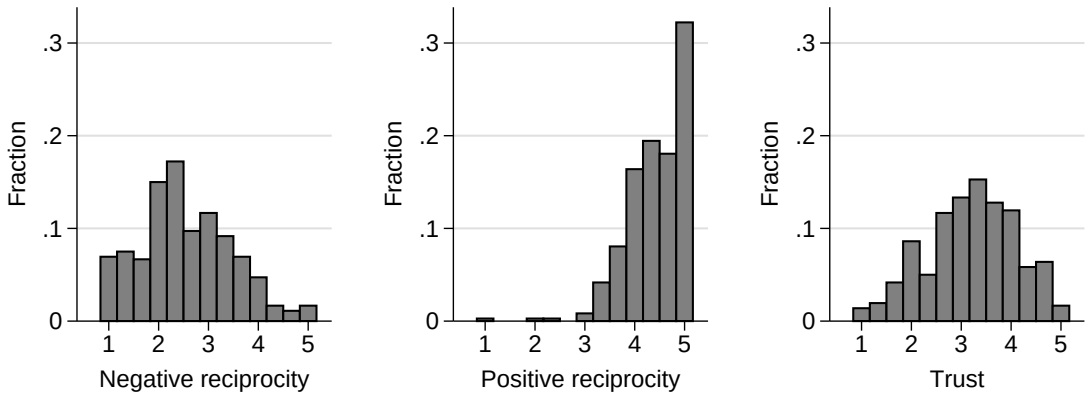


Table C.1: *Balancing table*

	Baseline	No Info	High Provision Point	Combined	Chi^2
Gender	0.66	0.56	0.50	0.61	5.75
	(0.48)	(0.50)	(0.50)	(0.49)	
Age	26.78	23.82	25.36	25.71	5.85
	(8.56)	(3.73)	(3.84)	(6.61)	
Student	0.85	0.93	0.96	0.89	6.73*
	(0.35)	(0.26)	(0.20)	(0.32)	
Openness	3.70	3.62	3.54	3.65	4.87
	(0.64)	(0.60)	(0.52)	(0.70)	
Conscientiousness	3.77	3.69	3.74	3.74	1.83
	(0.63)	(0.58)	(0.57)	(0.65)	
Agreeableness	3.07	3.16	3.30	3.06	7.23*
	(0.62)	(0.68)	(0.55)	(0.65)	
Extraversion	3.45	3.48	3.56	3.41	1.89
	(0.68)	(0.66)	(0.59)	(0.67)	
Emotionality	3.21	3.15	3.10	3.20	2.95
	(0.74)	(0.65)	(0.58)	(0.65)	
Honesty	3.25	3.33	3.37	3.31	0.80
	(0.75)	(0.71)	(0.64)	(0.79)	
SVO angle	16.39	17.87	19.28	15.95	3.78
	(12.83)	(12.99)	(12.17)	(12.77)	
Risk aversion	6.52	6.49	6.38	5.99	5.26
	(1.56)	(1.40)	(1.44)	(1.60)	
Trust	3.07	3.25	3.31	3.30	5.82
	(0.88)	(0.93)	(0.88)	(0.91)	
Negative reciprocity	2.57	2.51	2.54	2.56	0.13
	(0.96)	(0.84)	(0.90)	(0.98)	
Positive reciprocity	4.40	4.44	4.44	4.46	2.33
	(0.52)	(0.56)	(0.64)	(0.61)	
Observations	144	72	72	72	

Notes: Table reports variable means by treatment. Standard deviations are reported in parentheses. Column "Chi^2" reports Kruskal Wallis test statistics / Chi^2 test statistics for binary variables. *** p<0.01, ** p<0.05, * p<0.1.

Table C.2: *Pairwise correlations between the explanatory variables*

	OP	CO	AG	EX	EM	HH	SVO	Risk	Trust	Neg.
CO	0.059									
AG	0.115*	-0.065								
EX	0.256***	0.089	0.199***							
EM	-0.024	0.064	-0.112*	-0.199***						
HH	0.169**	0.029	0.257***	0.096	0.175***					
SVO	0.077	0.014	0.050	0.019	0.105*	0.221***				
Risk	0.044	-0.009	-0.136**	-0.040	0.107*	0.066	-0.008			
Trust	0.215***	-0.112*	0.408***	0.280***	-0.096	0.240***	0.181***	-0.105*		
Neg.	-0.174***	-0.100	-0.421***	-0.195***	-0.059	-0.459***	-0.249***	0.112*	-0.335***	
Pos.	0.219***	0.074	0.180***	0.249***	0.062	0.183***	0.072	0.003	0.188***	-0.143**

Notes: Table depicts pairwise correlations between the explanatory variables. Abbreviations: OP = Openness to experience, AG = Agreeableness, EX = Extraversion, EM = Emotionality, HH = Honesty-humility, SVO = Social value orientation, Risk = Risk aversion, Neg. = Negative reciprocity, Pos. = Positive reciprocity. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table C.3: *Implementation success dependent on treatment condition*

DV: implementation success	(1)
High PP = 1	-0.459**
	(0.197)
No Info = 1	-0.151
	(0.206)
High PP # No Info	-1.375***
	(0.296)
Observations	450
Pseudo R^2	0.282

Notes: Table shows coefficients after probit regression of implementation success on indicators for institutional conditions on the group level. Controls include period fixed effects. Standard errors are robust. *** p<0.01, ** p<0.05, * p<0.1.

C.2 Factor Analysis

Albeit the root mean square error of approximation [RMSEA] of .060 as well as the standardized root mean residual [SRMR] of .080 indicated an acceptable model fit, the comparative fit index [CFI] of .594 is fairly low, which can be troublesome (Bentler and Bonett, 1980; Brown, 2006). This finding, however, is not unique to our setting, but has been reported for other measures of personality like the five-factor model. As a consequence, the usefulness of CFA to assess the factor structure of personality measures has been questioned (Hopwood and Donnellan, 2010; McCrae et al., 1996; Oswald et al., 2013; Raykov, 1998). However, as a six-factor model yielded significantly better fit than a one-factor model ($\Delta\chi^2 = 2041.12$, $p < .01$), we decided to follow the approach suggested by Ashton and Lee (2009) and conducted an item-level factor analysis applying principal axis extraction with varimax rotation of six factors. The eigenvalues of the first eight factors derived from the 60 items exceeded the value of 1 and specifically were 5.14, 3.81, 3.43, 2.58, 1.98, 1.84, 1.39, and 1.13. There is suggestive evidence for a six-factor solution, as the eigenvalue decreased rather sharply from 1.84 for factor six to 1.39 for factor seven. A varimax rotation ensured that item loadings of all items were primarily on the factor defined by the other items of the respective scale and revealed that the primary item loadings replicated the basic structure of the HEXACO model. Regarding individual factor loadings, one item showed a slightly higher loading on a different factor than proposed by the manual. This was "I worry a lot less than most people do", which loaded higher on conscientiousness than on emotional stability. Apart from item 19, 37, and 49 connected to the openness dimension, and item 5 connected to emotional stability, all primary loadings exceeded the value of .30. As this is in line with what has been found in research on HEXACO personality in general and average scores, standard deviations and interrelations between dimensions are comparable to those reported by Moshagen et al. (2014), we decided to keep the original item structure.

C.3 Instructions

Welcome to the experiment and thank you for your participation!

Please do not speak with the other participants from now on. Please remain silent throughout the course of the experiment.

This experiment is designed to evaluate economic decision-making behavior. You will be able to earn money which will be paid to you after the experiment privately and in cash.

The entire experiment will last for about 70 minutes and consists of 3 parts. At the beginning of each part you will receive detailed instructions. The parts of the experiment are independent of each other, i.e., decisions in one part do not have any impact on your earnings in the other parts. The sum of your earnings from all parts will be added to your total earnings in this experiment. The total earnings will be paid to you upon completion of the final part, individually and in cash. If there are any questions concerning the instructions or during the experiment, press the red button on your keyboard (F11). One of the experimenter will come to you and answer your questions in private. During the experiment, you will be asked to make several decisions. Some of them will be made in interaction with other participants. This means that both your own decisions and those of the other participants may determine your payoffs.

Payoff During the course of the experiment, payoffs are calculated not in Euros, but in Experiment Points (EP). At the end of the experiment, the sum of your earned EP will be converted to Euros. Here, the following exchange rate applies: 100 EP = 1 €.

In addition to the income that you can earn during the experiment, you receive 4 € for your punctuality and 0.50 € for each page of the questionnaire at the end of the experiment.

Anonymity None of the other participants will be able to observe your choices in this experiment. In addition, the data from the experiment will be evaluated anonymously. At the end of the experiment, you have to sign a receipt for the income you have earned during this experiment. This is only due to accounting issues and cannot be used to associate your personal information with your decisions. Your name cannot be combined with your behavior in the experiment at any time.

Rounds Part 1 consists of a number of mutually independent rounds. In each round, you have between 60 and 90 seconds to make your decision. Each round lasts at least

60 seconds; after this the end of the round is equally likely at any second. A clock will inform you on elapsed time during the first 60 seconds.

Groups In each round, you form a group with seven other participants. The allocation to those groups is random. The groups are dissolved after the end of each round and re-formed randomly. None of the participants will be informed about the identity of the other members of her group.

Decision Task At the beginning of each round, each of the eight members of a group receives an initial endowment of 100 points. During the course of the round each participant can decide freely to keep this amount in her personal account, or to invest it in the group account. As soon as at least five of the eight members of the group have decided to invest their initial endowment in the group account, the investment pays off and grants a profit of 2800 points, which is distributed evenly across all eight members of the group. If fewer than five group members have decided to invest until the end of the round, the investment does not pay off. In this case, the investments in the group account are lost. Each group member then receives only the points from her private account.

During each round, every group member can anonymously follow the decisions of all other group members on the screen.

Earnings You can decide in each round if you want to invest your initial endowment in the group account. This decision you can make at any time the current round by selecting your contribution and click "OK". If you choose to keep your initial endowment and not to invest in the public account, there are two possible results:

- If at least five of the eight members of the group have decided to invest their initial endowment in the group account, in addition to your initial endowment you will receive one eighth of the profit of 2800 points. Overall, your payoff in this case equals $100 + 2800 \div 8 = 450$ points.
- If by the end of the round fewer than five group members have decided to invest their initial endowment in the group account, all you get is your initial endowment of 100 points. If you decide to invest your initial endowment, there are also two possible outcomes:
- If at least five of the eight members of your group have decided to invest their initial endowment in the group account, you will receive one eighth of the profit

of 2800 points. Overall, your payoff in this case equals $2800 \div 8 = 350$ points.

- If by the end of the round fewer than five group members have decided to invest their initial endowment in the group account, you do not earn anything in this round.

At the end of the experiment, one of the rounds you played will be randomly selected and paid out.

Appendix D

The Influence of Overconfidence and Competition Neglect on Entry into Competition

D.1 Proofs

Existence of q^*

Proof. First, it is straightforward to see that the denominator is always > 0. In addition, keeping in mind that $\xi > 1$ and $q_j \in [0; 1]$, one can see that the numerator is ≥ 0 as long as $s \leq 2$ (and thus, $q^* \geq 0$ as well): $2 - (1 - q_j)s \geq 0 \Leftrightarrow s \leq \frac{2}{1-q_j}$. This condition is strictest if $q_j = 0$, when it becomes $s \leq 2$. On the other hand, the numerator is smaller or equal than the denominator (an, in turn, $q^* \leq 1$) as long as $s \geq \frac{2}{2-q_j}$. This is strictest for $q_j = 1$, when it becomes $s \geq 2$. Thus, for $s = 2$, $q^* \in [0; 1]$ for all possible values of q_j (and ξ). □

Uniqueness of q^*

Proof. First, note that due to the common knowledge assumption, q^* is symmetric. For uniqueness, that is, for q^* to be a global threshold, one needs to show that $\forall i \in N$, $q_i' < q^* \Leftrightarrow \mathbf{E}_{NC}(q_i') > \mathbf{E}_C(q_i')$ and $q_i'' > q^* \Leftrightarrow \mathbf{E}_{NC}(q_i'') < \mathbf{E}_C(q_i'')$. To see this, plug q_i' into $\mathbf{E}_{NC} > \mathbf{E}_C$. Solving for q_i' yields $q_i' < \frac{2-(1-q_j)s}{\xi[s(2-q_j)-2]-[(1-q_j)s-2]}$, which equals $q_i' < q^*$. Similarly, plugging q_i'' into $\mathbf{E}_{NC} < \mathbf{E}_C$ and solving for q_i'' yields $q_i'' > \frac{2-(1-q_j)s}{\xi[s(2-q_j)-2]-[(1-q_j)s-2]}$, which equals $q_i'' > q^*$. Thus, q^* is unique. □

Comparative statics of q^*

$$\frac{\partial q^*}{\partial s} = -\left(\frac{(1-q_j)}{\xi[s(2-q_j)-2]-[(1-q_j)s-2]} + \frac{[2-(1-q_j)s]\times(\xi(2-q_j)-(1-q))}{\{\xi[s(2-q_j)-2]-[(1-q_j)s-2]\}^2}\right)$$

To see that this is < 0, note that the denominators of both fractions are > 0. Then, the first fraction is ≥ 0 because $(1-q_j) \geq 0$. The numerator of the second fraction consists of two parts. As established in the existence proof, the first, $[2-(1-q_j)s]$, is positive. The second, $(\xi(2-q_j)-(1-q))$, is > 0 as long as $\xi > \frac{1-q}{2-q}$, which is satisfied for $\xi > 0.5$. Thus, the whole second fraction is > 0. Finally, the whole expression is < 0 (and not ≤ 0) because even if the first fraction $= 0$, which happens for $q_j = 1$, the second fraction is > 0.

$$\frac{\partial q^*}{\partial \xi} = -\frac{[2-(1-q_j)s]\times[s(2-q_j)-2]}{\{\xi[s(2-q_j)-2]-[(1-q_j)s-2]\}^2}$$

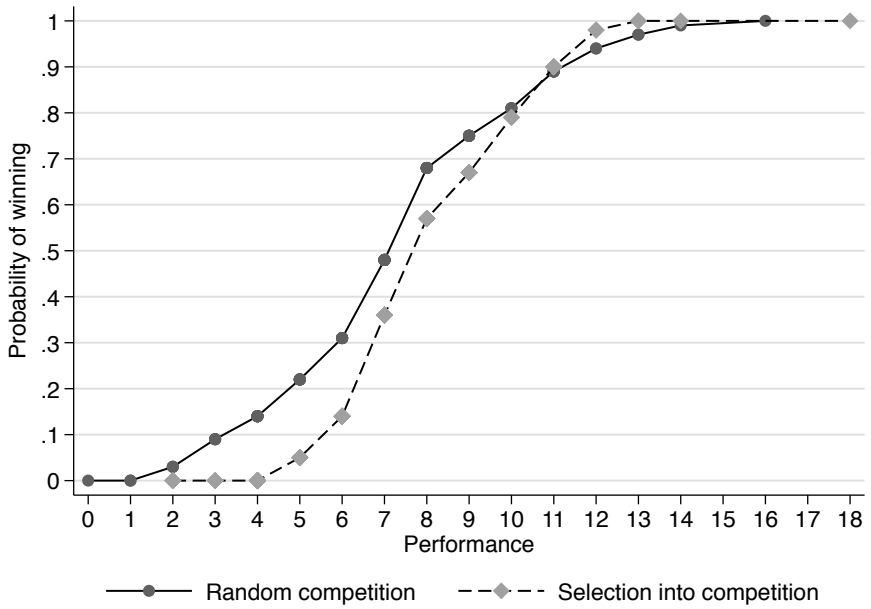
To see that this is < 0, note that the denominator is > 0. The first part of the numerator is again > 0, as established in the existence proof. The second part of the numerator is positive as long as $s \geq \frac{2}{2-q_j}$, making the whole expression negative.

$$\frac{\partial q^*}{\partial q_j} = \frac{s}{\xi[s(2-q_j)-2]-[(1-q_j)s-2]} + \frac{[2-(1-q_j)s]\times(\xi-1)s}{\{\xi[s(2-q_j)-2]-[(1-q_j)s-2]\}^2}$$

Again, both denominators are > 0. Thus, the first fraction is also > 0 as s is positive by definition. The numerator of the second fraction again consists of two parts, with $[2-(1-q_j)s]$ being positive. As $\xi > 1$, the second part is positive as well, which means that the whole expression is > 0.

D.2 Figures and Tables

Figure D.1: *Actual chance of winning by round three performance*



Notes: Figure depicts the actual chances of winning given reference group performance when the competition group consists of the whole reference group (solid line) and when the competition group consists of self-selected individuals only (dashed line). Calculations based on 48 individuals in the reference group, of which 21 chose the competitive incentives.

Table D.1: *Balancing table*

	None	Overconfidence	Selection	Both	Chi^2
Gender	0.50	0.48	0.50	0.49	0.23
	(0.51)	(0.50)	(0.51)	(0.51)	
Age	23.69	23.17	23.30	24.15	3.03
	(3.07)	(4.70)	(3.60)	(5.75)	
Risk aversion	5.67	5.94	6.30	6.38	3.34
	(2.00)	(2.23)	(1.76)	(1.55)	
Round 1					
Score	5.73	6.27	6.30	5.55	3.62
	(2.95)	(2.81)	(2.74)	(2.88)	
Overestimation of self	1.23	1.79	0.32	1.47	5.24
	(2.96)	(3.61)	(2.70)	(3.49)	
Overestimation of others	0.66	1.88	0.50	0.64	8.24**
	(3.03)	(2.97)	(2.42)	(2.56)	
Overestimation of probability	0.14	0.04	0.07	0.16	5.76
	(0.29)	(0.26)	(0.28)	(0.24)	
Round 2					
Score	6.40	7.08	6.93	6.32	2.55
	(3.06)	(3.00)	(3.06)	(3.14)	
Overestimation of self	-0.16	0.51	-0.15	0.13	3.69
	(2.17)	(2.28)	(1.72)	(2.44)	
Overestimation of others	-0.49	0.75	-0.13	-0.17	11.56***
	(2.33)	(1.85)	(2.12)	(1.77)	
Overestimation of probability	0.08	-0.01	0.01	0.07	3.29
	(0.30)	(0.25)	(0.25)	(0.27)	
Observations	48	48	46	47	

Notes: Table reports variable means in the main sessions by treatment. Standard deviations in parentheses. Estimations are based on round three. Column "Chi^2" reports Kruskal Wallis test statistics / Chi^2 test statistic for binary variable. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

D.3 Gender Differences

In this section, the analyses of beliefs and competition decisions are done separately for men and women to investigate whether the established difference from the literature are prevalent in my sample as well and whether the influence of competition neglect and overconfidence on competiton decision is different for men and women.

Beliefs

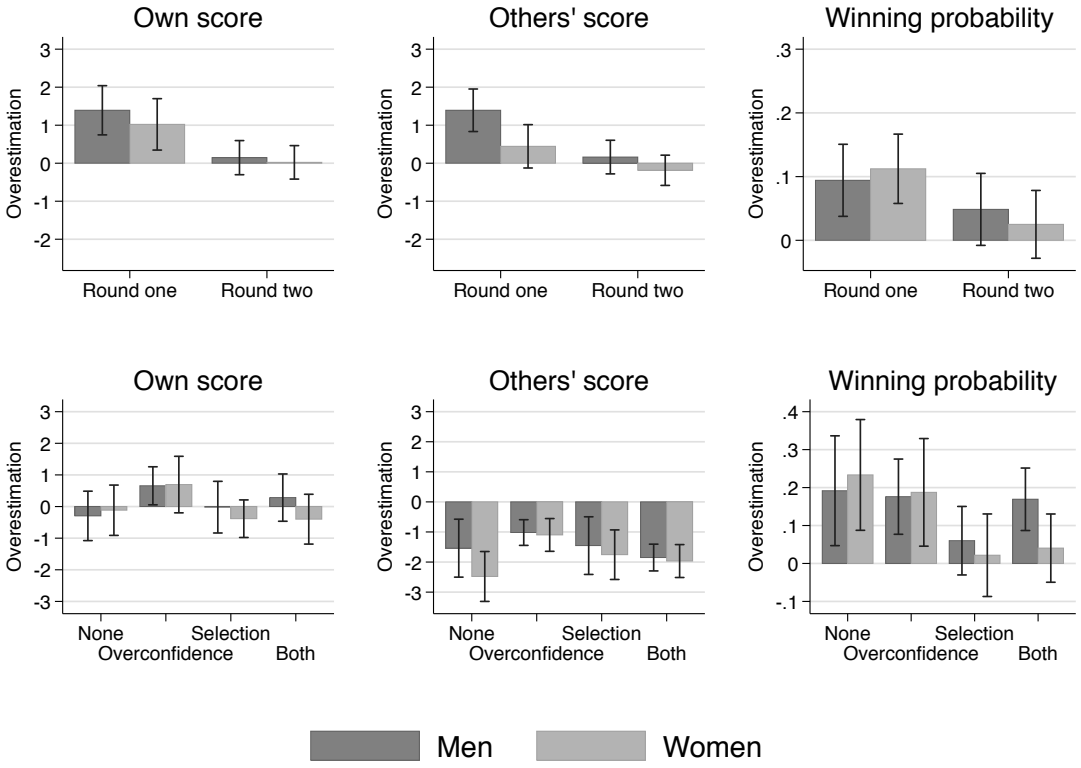
Figures D.2 and D.3 are similar to Figures 4.2 and 4.3 in Section 4.5 of Chapter 4, but depict beliefs split up by men and women. It can be seen that the patterns for men and women are more or less the same, and that there are are no significant differences between men and women in overconfidence in my sample. This is surprising as previous literature on gender differences in competitive environments has established heterogeneity in overconfidence as one important driver of gender differences in competitiveness. For instance, Niederle and Vesterlund (2007) find that 75% of men and 43% of women believe to have the highest performance in a four-person group.

Competition Decisions

Now, I analyze how competition decisions and the effects of competition neglect and overconfidence differ for men and women.

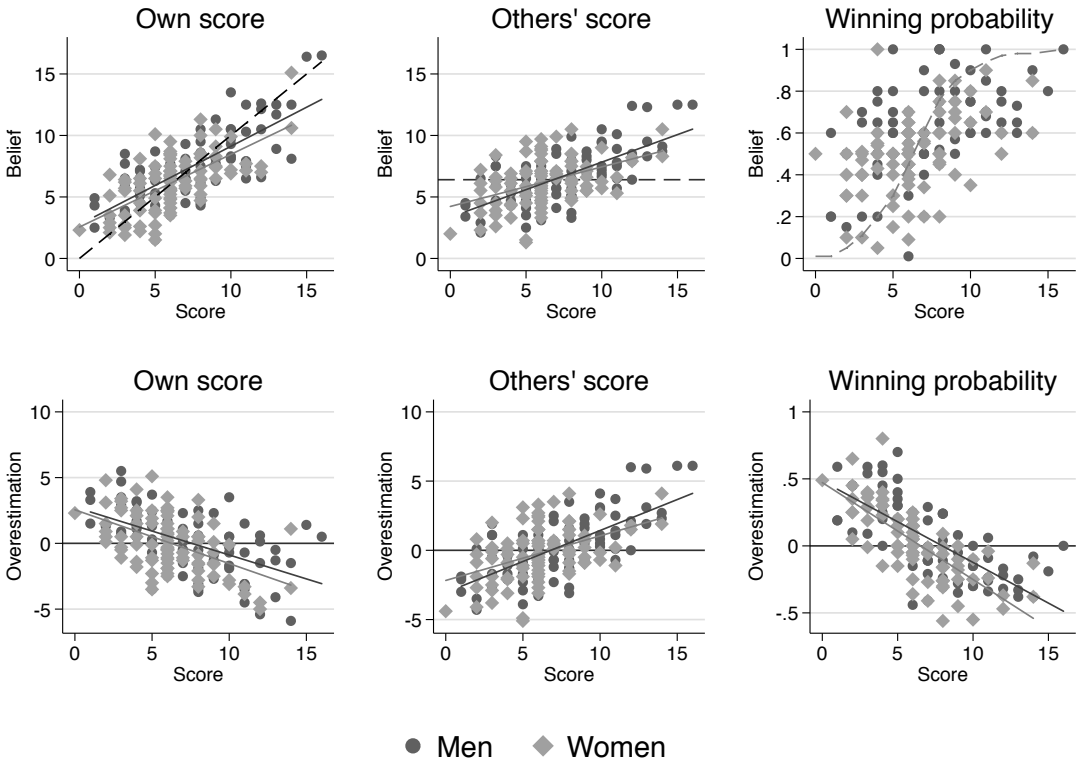
An important condition to being able to compare competition decisions of men and women directly is that there are no gender differences in performance – in this case, the money-maximizing choice is the same for both genders and, normatively speaking, competition decisions should not differ. This is also an important assumption underlying arguments for acting against gender inequality in labor markets and calls for an increased number of women in competitive professions. Thus, Figure D.4 depicts average performance by round and gender in the main sessions of the experiment. While in most papers, no significant difference in performance between men and women is found for this task, I find that in my sample, men on average solve 1.28 items more than women in round one, 0.98 items more in round two, and 0.80 items more in

Figure D.2: *Average bias in beliefs*



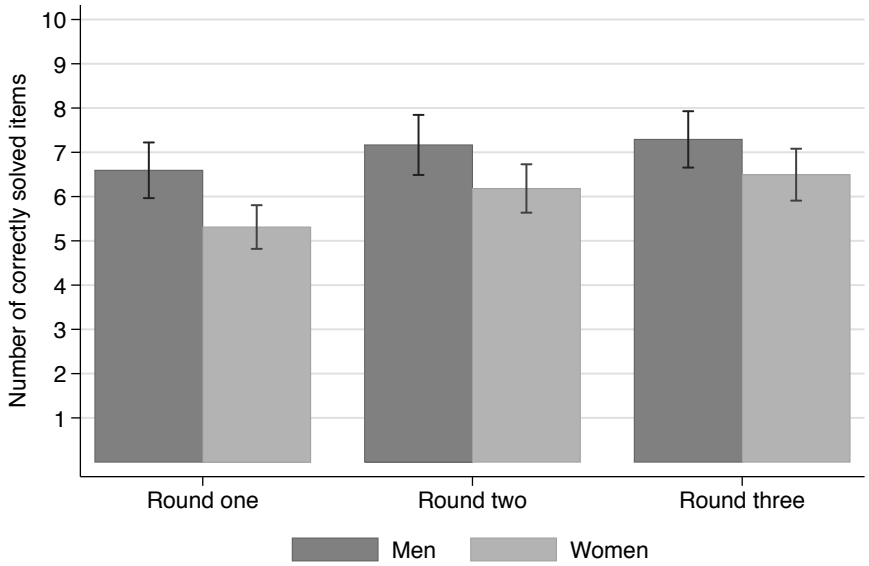
Notes: Upper row depicts average bias in beliefs about own performance (left panel), reference group performance (middle panel) and winning probability (right panel) in rounds one and two in the main sessions. Lower row depicts average bias in beliefs about own performance (left panel), reference group performance (middle panel) and winning probability (right panel) by Treatment in round three in the main sessions. Calculations are based on 189 observations. Means are depicted as bars, 95% confidence intervals as error bars.

Figure D.3: *Beliefs and overestimations by actual performance in round two.*



Notes: Upper row of figure depicts beliefs about own performance (left panel), reference group performance (middle panel) and winning probability (right panel) depending on actual performance by men and women in round two of the main sessions. Lower row depicts bias in beliefs about own performance (left panel), reference group performance (middle panel) and winning probability (right panel) depending on actual performance by men and women in round two of the main sessions. Calculations are based on 96 men and 93 women.

Figure D.4: *Average performance by gender across rounds*



Notes: Figure depicts average number of correctly solved items by men and women across rounds in the main sessions. Calculations are based on 96 men and 93 women. Means are depicted as bars with 95% confidence intervals.

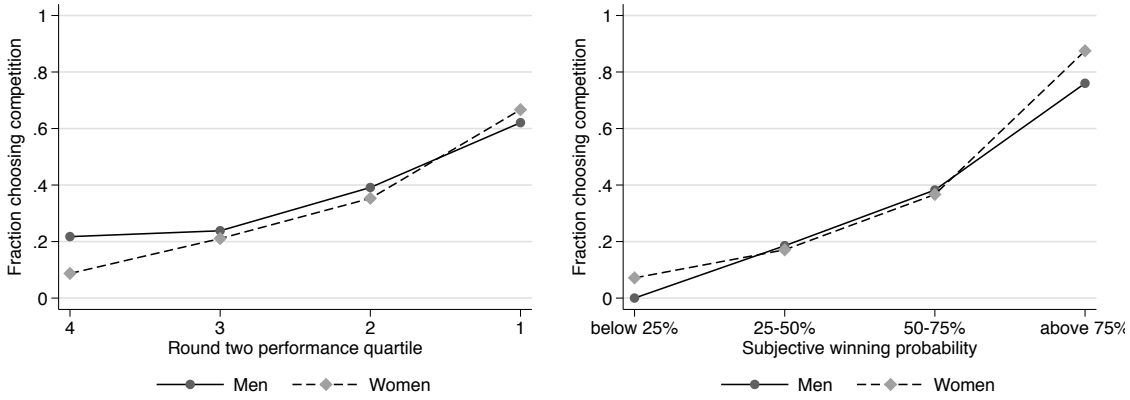
round three. Thus, they perform significantly better than women in all rounds of the experiment (two-sided Mann-Whitney tests yield z=2.43, p=.0148, z=1.98, p=.0476, and z=1.75, p=0.0803).[1]

Next, cosider entry rates. When considering average entry across treatments, contrary to Niederle and Vesterlund (2007) and most of the literature, there exists no significant gender gap: 38.54% of men and 27.96% of women chose the competitive incentive scheme (two-sided Fisher's exact test yields p=.165).

Figure D.5 depicts the proportion of participants selecting into competition by performance quartile in the initial competition. Contrary to what Niederle and Vesterlund (2007) find, there is a positive relationship between performance quartile and propor-

[1] Note that while some studies have indicated that there might be an adverse effect of having to work under competitive incentives for women Gneezy et al. (2003), I find the opposite. Of course I cannot disentangle these effects from learning; however, the fact that womens' average performance improves more strongly between rounds one and two than mens' average performance speaks against a strong negative effect and a strong positive of the competitive incentives for women and men, respectively.

Figure D.5: *Competition choices by previous performance and beliefs*

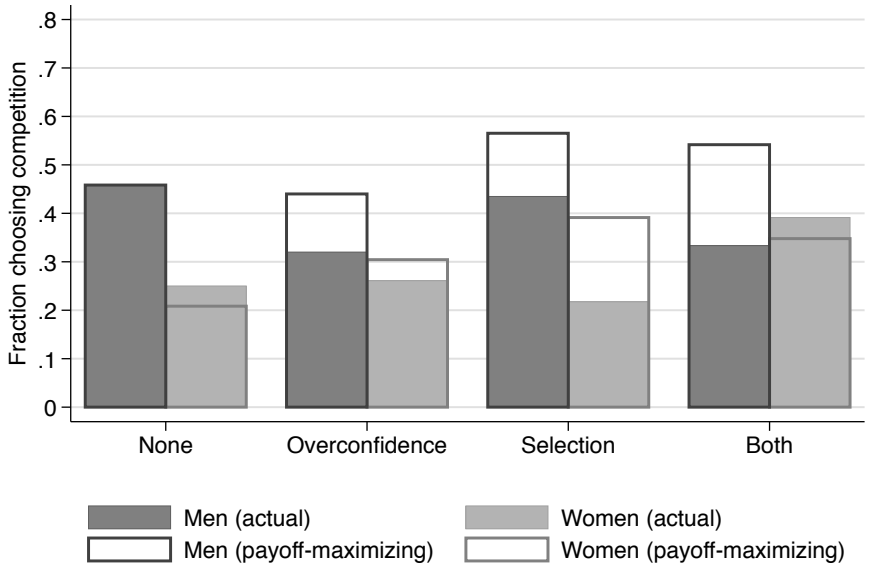


Notes: Figure depicts fractions of men and women choosing the competitive incentive scheme in round three of the main sessions by performance quartile in round two (left panel) and subjective winning probability (right panel). Calculations are based on 96 men and 93 women.

tion of individuals selecting into competition for both men and women. In addition, the relationship looks rather similar for men and women, with the exception of subjects in the lowest quartile: among these, competition rates for men are higher than for women, indicating that there is more overcompetition of low-performing men than women. The right panel of Figure D.5 depicts the proportion of men and women selecting into competition by their subjective probability of winning. There is a positive relationship between beliefs and the proportion of subjects selecting into competition that is even stronger than for previous performance. Interestingly, however, there is a small percentage of women who enter the competition even though they think that their probability of winning is below 25%.

The fraction of men and women selecting the competitive incentive scheme by treatment is depicted in Figure D.6. Apart from the fact that absolute ratios of subjects deciding to compete seem pretty low when compared to previous literature, it can also be seen that while more men than women decide to compete without feedback, gender differences seem to be ameliorated when feedback is available in round three. Competing against a selected or a random sample does not seem to make a difference to men and to women only if combined with feedback; thus, subjects appear to exhibit

Figure D.6: *Fraction of individuals selecting into competition by treatment*



Notes: Figure depicts fraction of men and women choosing the competitive incentive scheme in round three of the main sessions. Calculations are based on 96 men and 93 women.

competition neglect. Testing for differences between genders within treatment and for differences between the treatments with Fisher exact tests does not detect significant differences, however.[2]

To explore whether there are differential effects of the treatments for men and women, I build on the regression specification shown in Table 4.3 and include interactions of the treatment conditions with gender. The results are shown in Table D.2. It can be seen that while the main effect of gender still is insignificant and previous performance and beliefs still have significant influence, there is an additional effect in the full model: when controlling for beliefs and positive performance, the interaction between the feedback indicator and gender is significant. Hence, providing feedback increases the probability that women select the competitive incentive scheme.

[2] Note that while the sample size does not allow to detect differences in this experiment, the number of subjects would have been sufficient to detect gender differences based on previous results. As ex ante power calculations were based on competition ratios in existing experiments, this was unexpected.

Table D.2: *Determinants of competition choices*

DV: Competition (Y/N)	(1)	(2)	(3)	(4)
Feedback = 1	-0.106	-0.0950	-0.113	-0.0965
	(0.0944)	(0.0937)	(0.0860)	(0.0843)
Selection = 1	-0.00256	-0.0198	0.0308	0.0200
	(0.0929)	(0.0991)	(0.0845)	(0.0875)
Feedback # Female	0.196	0.208	0.240**	0.223*
	(0.127)	(0.132)	(0.112)	(0.117)
Selection # Female	-0.100	-0.0803	-0.124	-0.114
	(0.127)	(0.139)	(0.115)	(0.125)
Previous performance	0.0581***	0.0602***	-0.0262	-0.0256
	(0.0105)	(0.0113)	(0.0236)	(0.0235)
Own performance belief			0.110***	0.115***
			(0.0317)	(0.0323)
Others' performance belief			-0.0870***	-0.0866***
			(0.0273)	(0.0297)
Subj. winning probability			0.621***	0.602***
			(0.168)	(0.175)
Female = 1	-0.0820	-0.0692	-0.0278	-0.00136
	(0.115)	(0.129)	(0.0971)	(0.107)
Risk aversion	-0.0449**	-0.0434**	-0.0316*	-0.0307*
	(0.0178)	(0.0179)	(0.0165)	(0.0170)
Controls	No	Yes	No	Yes
Observations	189	187	189	187
Adjusted R^2	0.158	0.156	0.320	0.314

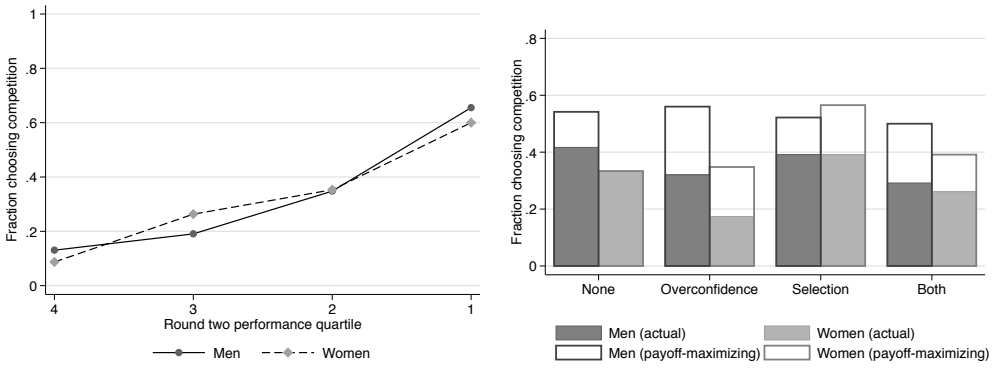
Notes: Table reports results of linear probability estimations. Robust standard errors in parentheses. Estimations are based on round three. Controls include age, an indicator for field of study (business, economics, STEM, humanities, law, psychology, other social sciences, other), and math grade. *** p<0.01, ** p<0.05, * p<0.1.

D.4 Analysis of Round 4: Competition Based on Past Performance

In the experiment, an additional round was included after the three rounds described in the paper, the difference being here that in round three, participants made the choice on their subsequent performance while in round four, they were again paid based on their performance in round two. This was done to address the fact that factors like differing attitudes against the act of competing or ambiguity attitudes are discussed to influence competition decisions (see Niederle and Vesterlund, 2011).

First, there is even less of a gender difference in competition entry than in round three: while 35.42% of men chose the competitive incentive scheme, 29.03% of women do. Considering actual performance in the reference group, a payoff-maximizing individual should choose competition if her round-two performance is at least seven without and eight with selection in the competition group. Based on round two performances, this applies to 46.81% of men and 35.42% of women.

Figure D.7: *Competition choices by performance and treatment in round four*



Notes: Figure depicts fraction of men and women choosing the competitive incentive scheme in round four of the main sessions by performance quartile (left panel) and by treatment (right panel). Calculations based on 96 men and 93 women.

Exploring competition decisions based on actual performance and by treatment in Figure D.7, it can be seen that as in round three, there is a positive relationship between performance quartile and proportion of individuals selecting into competition that looks similar for men and women. With respect to treatment variations, it can be seen that there are roughly similar entry ratios for men facing the random and the selected sam-

ple of competitors, while entry ratios for women are lower when they have to compete against competitors who themselves self-selected into the competitive incentive scheme. Feedback is associated with lower entry rates for both men and women.

D.5 Instructions

General Introduction

Welcome to the experiment and thank you for your participation!

Please do not speak with the other participants from now on. Please remain silent throughout the course of the experiment.

This experiment is designed to evaluate economic decision-making behavior. You will be able to earn money which will be paid to you after the experiment privately and in cash.

During the experiment, you will be asked to make several decisions. Some of them will be made in interaction with other participants. This means that both your own decisions and those of the other participants may determine your payoffs.

The entire experiment will last for about 60 minutes and consists of two parts. At the beginning of each part you will receive detailed instructions. The parts of the experiment are independent of each other, i.e., decisions in one part do not have any impact on your earnings in the other parts. The sum of your earnings from all parts will be added to your total earnings in this experiment. The total earnings will be paid to you upon completion of the final part, individually and in cash.

Payoff During the course of the experiment, payoffs are calculated not in Euros, but in Experiment Points (EP). At the end of the experiment, the sum of your earned EP will be converted to Euros. Here, the following exchange rate applies: 100 EP = 1 €.

In addition to the income that you can earn during the experiment, you receive 5 € for your punctuality.

Anonymity None of the other participants will be able to observe your choices in this experiment. In addition, the data from the experiment will be evaluated anonymously. At the end of the experiment, you have to sign a receipt for the income you have earned during this experiment. This is only due to accounting issues and cannot be used to associate your personal information with your decisions. Your name cannot be combined

with your behavior in the experiment at any time.

Permitted aids There are a pen and notepaper located on your desk. Please leave them on the table after the experiment.

If you have any questions after the instructions or during the experiment, please press the red button on the keyboard (F11). One of the experimenters will then answer your question in private. If you do not need help any more, please press the red button again.

Introduction of Part 1 Procedure

In this part of the experiment, you will go through four rounds. In every round, you will have to work on two different tasks: a sum-calculation task and an estimation task.

Sum calculation In this task, you are requested to calculate sums of five random two-digit numbers. You will have a time window of three minutes for calculating as many sums as possible. It is not allowed to use a calculator, but you can use the provided notepaper.

Your screen is built up as follows: On the left side, you are going to see the five two-digit numbers you should add up. On the right, there is a text box where you enter the solution. Then, you press “OK” (on the far right) to submit your answer and going to the next calculation. On the top right side you can check the time you have left in this round.

– Participants have one minute to get accustomed to the screen for the sum calculations –

Payment rules You are paid according to different rules in every round. At the beginning of each round, you are informed about how your earning is determined.

In some rounds, your payment can also depend on the performance of other participants. In this case, this does not apply to the performance of the other participants sitting in the lab with you right now, but participants of a past experiment. These were selected in the same way as you, have worked on the same tasks and made the same decisions as you. Therefore, these participants are comparable to you and the other participants of this session. In the following, whenever you read about the "reference group", these comparable participants of an earlier session are meant.

Estimation task In this task you are requested to give your own estimation. This will be explained more in detail in the following rounds.

Summary of rounds In summary, all rounds are going to follow the same procedure:

- You are informed about the payoff rule for this round
- You do the estimation task
- You have three minutes to finish the sum calculations

At the end of the experiment, one out of the four rounds will be chosen randomly and your earnings from that round paid together with the 5 € show-up fee and your payoff from Part 2.

Explanation of Estimation Task

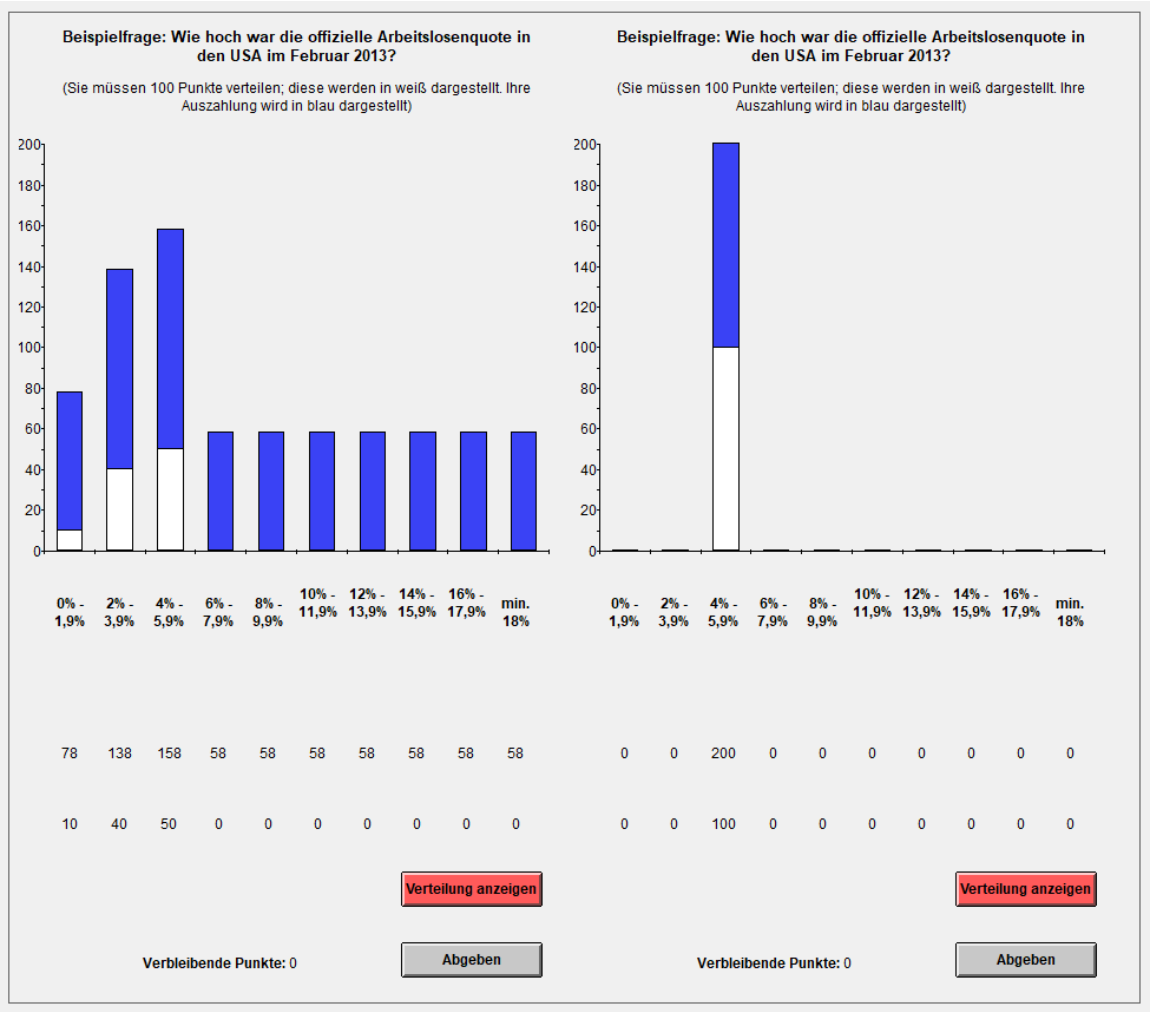
Estimation Task 1 In the following you are asked two questions. You are asked to state what you think the correct answer to the respective question is.

For each question, you can allocate a total of 100 points to 10 possible categories. To do so, select the check box below the particular category and type in how many of your 100 available points you want to allocate to that category. By pressing "show distribution" you can see your chosen distribution of points between the categories looks like (bars on the upper part of the screen). You have to allocate the total amount of 100 points available to you. Once you are satisfied with the distribution, press "submit" to get to the next screen.

In total, you can earn up to 200 EP in each question. For this, the points you have distributed are converted into EP in consideration of the full distribution. The EP you earn when you submit the currently chosen distribution are displayed to you as as bars and in numbers above the check boxes as soon as you click on "show distribution".

How much points you should allocate to each category depends on your estimation of the correct answer to the question asked. To illustrate this point, consider the following example:

How high was the official unemployment rate in the US in February 2003?



You can now freely choose how you want to spread your 100 points. The figure shows you two exemplary options.

- *Left distribution:* Assume you think it very likely that the true answer is slightly under 5%. Then you could allocate 50 points to the category "4 to 5.9%", 40 points to the category "2% to 3.9%", and the remaining 10 points to the category "0-1.9%", for example.
  Your payoff for this question is determined by the amount of points you have allocated to the category that contains the true unemployment rate. If it is actually between 4 and 5.9%, you earn 158 EP. If it is between 2 and 3.9%, you receive 138 EP, and if it is between 0 and 1.9%, you receive 78 EP. For each other true unemployment rate you receive with this distribution 58 EP. Thus, your finally

payoff depends on your given estimation and the correct answer to the questions. You can change the allocation of your points as much as you want to such that it represents your personal estimation as well as possible. As the actual unemployment rate in the USA in February 2013 was 7.7%, you would have earned 58 EP with the left distribution.

- *Right distribution:* Assume you had allocated all your available points on one category, for example on the category "4% to 5.9%". Then, the distribution of possible earnings for this question would look as follows: If the true unemployment rate lies between 4 and 5.9%, you would earn the maximum payoff of 200 EP for this question. However, since the true unemployment rate is 7.7%, you would not have earned anything with this distribution of points.

How you find a trade-off between the precision of your estimation and the risk that you are wrong is up to you. Keep in mind the following three important things:

- Your assessment of the right answer to the question is a personal estimation which is based on the information you have.
- Depending on your decision you can earn up to 200 EP per question.
- Your decisions can also depend on your willingness to take risks. The estimations you are going to make now are individual.

Estimation Task 2: Winning Probability Now we would like you to give an estimation of how likely it is that you are going to solve more calculations correctly that a randomly chosen participant from the reference group.

As a reminder: The reference groups consist of participants of a past date which have solved the same tasks you will solve now.

Your earnings from this estimation are affected by the precision with which you specify the probability that your performance is above the performance of a randomly chosen participant from the reference group. The payoff is constructed such that you have the highest chance of earning money if you indicate your actual assessment. The mechanism works as follows: You indicate your estimated winning probability. In addition, the computer draws a number X between 0 and 100. Every number between 0 and 100 is drawn with the same probability.

A comparison of these two values determines according to which of the following criteria you are paid for your estimation.

- Option A: you receive 200 EP if you solve more calculations correctly than a randomly chosen participant of the reference group
- Option B: you receive 200 EP with probability X

The option according to which you are paid is selected as follows: If X is higher than your indicated winning probability, option B is chosen and you receive 200 EP with probability X. If X is lower than your indicated winning probability, option A is selected and you receive 200 EP if your performance in this round is better than the one of a randomly chosen subject of the reference group.

This means that you maximize your chance of winning the price of 200 EP by indicating your actual estimation of how likely you think it is to be better than a randomly chosen subject of the reference group.

An example: You think that you can solve more calculation tasks correctly that a randomly chosen person of the reference group with a probability of 62%.

- If you truly state this estimation, you are paid according to option A if X is lower than 62, which gives you a higher winning probability (namely 62%) than option B (X%). If X is higher than 62, you are paid according to option B, which gives you a higher chance of winning (namely X%) that option A (62%). Regardless of which number X is drawn randomly, you are always paid according to the rule that gives you a higher chance of receiving 200 EP.
- If you state another probability, for example 10%, you are paid according to option A if X is smaller than 10, which gives you a higher winning probability (namely 62%) than option B. If X is higher than 10, you are paid according to option B, which possibly gives you a lower winning probability than option A. If X is between 10 and 62, you are paid according to option B, although with option A you would have had a higher winning probability of 62%. In this case, stating a false probability reduces your chance to receive 200 EP.

The logic of this example applies to all other probabilities as well. Whatever you think how likely it is that you are going to solve more calculations correctly than a randomly chosen subject from the reference group, you maximize the chance of receiving 200 EP by giving your best possible estimation.

Payment Descriptions

Round 1 You receive 50 EP for each correct calculation.

Round 2 The performance of an individual from the reference group in this round is chosen randomly. Remember: the reference group consist of participants of a past date who solved the same tasks you are solving now. If you solve more calculations correctly than this person, you receive 100 EP per correct calculation; otherwise, you do not get any payoff for this round. If you solve the same number of calculations correctly as the randomly selected reference person, the computer will randomly choose a winner.

Round 3 In this round, you can choose how you want to be paid.

- *Option A*: As in round 1, you receive 50 EP for each correct calculation.
- *Option B*: As in round 2, the performance of an individual from the reference group in this round is chosen randomly. If you solve more calculations correctly than this person, you receive 100 EP per correct calculation; otherwise, you do not get any payoff for this round. If you solve the same number of calculations correctly as the randomly selected reference person, the computer will randomly choose a winner.

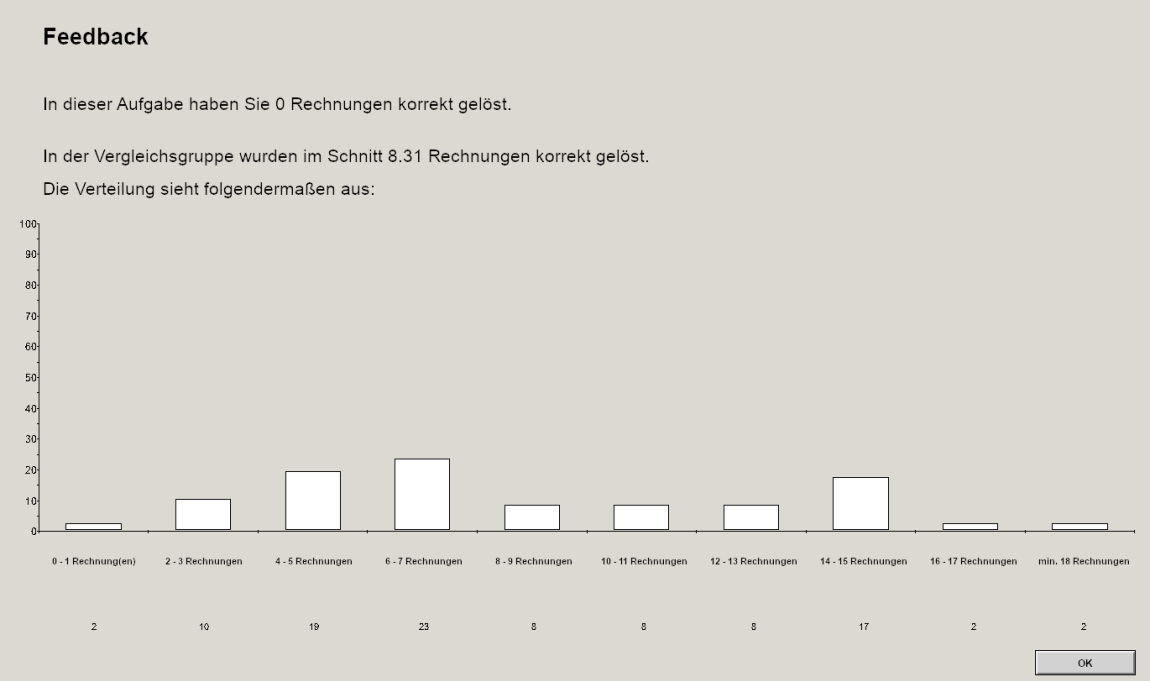
If the competitor was drawn from the selected competition group, the following sentence was added to Option B: In contrast to round 2, the person is selected only among the other participants in the reference group who have chosen option B.

Round 4 In this round, the number of correct calculations will again determine your payoff. However, in contrast to the previous rounds, you are not going to work again on the task. Instead, your performance in round 2 determines your earnings. You can choose how you want to be paid.

- *Option A*: As in round 1, you receive 50 EP for each correct calculation.
- *Option B*: As in round 2, the performance of an individual from the reference group in this round is chosen randomly. If you solve more calculations correctly than this person, you receive 100 EP per correct calculation; otherwise, you do not get any payoff for this round. If you solve the same number of calculations correctly as the randomly selected reference person, the computer will randomly choose a winner.

If the competitor was drawn from the selected competition group, the following sentence was added to Option B: In contrast to round 2, the person is selected only among the other participants in the reference group who have chosen option B.

Feedback



This translates to: "In this task, you solved X sum correctly.

In the reference group, an average number of Z calculations were solved correctly. The distribution of performances looks as follows"

Bibliography

Abele, S., Stasser, G., and Chartier, C. (2010). Conflict and coordination in the provision of public goods: A conceptual analysis of continuous and step-level games. *Personality and Social Psychology Review*, 14(4):385–401.

Abowd, J. M., Haltiwanger, J., Jarmin, R., Lane, J., Lengermann, P., McCue, K., McKinney, K., and Sandusky, K. (2005). The relation among human capital, productivity, and market value: Building up from micro evidence. In Corrado, C., Haltiwanger, J., and Sichel, D., editors, *Measuring Capital in the New Economy*, chapter 5, pages 153–204. University of Chicago Press.

Abreu, D. (1988). On the theory of infinitely repeated games with discounting. *Econometrica*, 56(2):383–396.

Akerlof, G. A. (1982). Labor contracts as partial gift exchange. *Quarterly Journal of Economics*, 97(4):543–569.

Akerlof, G. A. and Kranton, R. E. (2000). Economics and identity. *Quarterly Journal of Economics*, 115(3):715–753.

Akerlof, G. A. and Kranton, R. E. (2005). Identity and the economics of organizations. *Journal of Economic Perspectives*, 19(1):9–32.

Almlund, M., Duckworth, A. L., Heckman, J. J., and Kautz, T. D. (2011). Personality psychology and economics. In Hanushek, E. A., Machin, S., and Woessmann, L., editors, *Handbook of the Economics of Education*, volume 4, chapter 1, pages 1–181. Elsevier.

Anderson, L. R. and Holt, C. A. (1997). Information cascades in the laboratory. *American Economic Review*, 87(5):847–862.

Andreoni, J. and Miller, J. H. (1993). Rational cooperation in the finitely repeated prisoner's dilemma: Experimental evidence. *Economic Journal*, 103(418):570–585.

Ariely, D., Kamenica, E., and Prelec, D. (2008). Man's search for meaning: The case of Legos. *Journal of Economic Behavior & Organization*, 67(3):671–677.

Arkes, H. R., Christensen, C., Lai, C., and Blumer, C. (1987). Two methods of reducing overconfidence. *Organizational Behavior and Human Decision Processes*, 39(1):133–144.

Ashforth, B. E. and Mael, F. (1989). Social identity theory and the organization. *Academy of Management Review*, 14(1):20–39.

Ashraf, N., Bandiera, O., and Lee, S. S. (2014). Awards unbundled: Evidence from a natural field experiment. *Journal of Economic Behavior & Organization*, 100:44–63.

Ashton, M. and Lee, K. (2007). Empirical, theoretical, and practical advantages of the HEXACO model of personality structure. *Personality and Social Psychology Review*, 11(2):150–166.

Ashton, M. and Lee, K. (2009). The HEXACO-60: A short measure of the major dimensions of personality. *Journal of Personality Assessment*, 91(4):340–345.

Ashton, M., Lee, K., and de Vries, R. E. (2014). The HEXACO honesty-humility, agreeableness, and emotionality factors. *Personality and Social Psychology Review*, 18(2):139–152.

Athey, S. and Stern, S. (2002). The impact of information technology on emergency health care reforms. *RAND Journal of Economics*, 33(3):399–432.

Au, W. T. and Budescu, D. V. (1999). Sequential effects in give-some and take-some social dilemmas. In Foddy, M., Smithson, M., Schneider, S., and Hogg, M., editors, *Resolving Social Dilemmas: Dynamic, Structural, and Intergroup Aspects*, chapter 6, pages 87–99. Psychology Press New York.

Auriol, E. and Renault, R. (2008). Status and incentives. *RAND Journal of Economics*, 39(1):305–326.

Autor, D. H., Levy, F., and Murnane, R. J. (2002). Upstairs, downstairs: Computers and skills on two floors of a large bank. *Industrial and Labor Relations Review*, 55(3):432–447.

Azmat, G. and Iriberri, N. (2010). The importance of relative performance feedback information: Evidence from a natural experiment using high school students. *Journal of Public Economics*, 94(7):435–452.

Balafoutas, L., Kerschbamer, R., and Sutter, M. (2012). Distributional preferences and competitive behavior. *Journal of Economic Behavior & Organization*, 83(1):125–135.

Balliet, D., Parks, C., and Joireman, J. (2009). Social value orientation and cooperation in social dilemmas: A meta-analysis. *Group Processes & Intergroup Relations*, 12(4):533–547.

Bandiera, O., Barankay, I., and Rasul, I. (2007). Incentives for managers and inequality among workers: Evidence from a firm-level experiment. *Quarterly Journal of Economics*, 122(2):729–773.

Bandiera, O., Barankay, I., and Rasul, I. (2009). Social connections and incentives in the workplace: Evidence from personnel data. *Econometrica*, 77(4):1047–109.

Bandiera, O., Larcinese, V., and Rasul, I. (2015a). Blissful ignorance? A natural experiment on the effect of feedback on students' performance. *Labour Economics*, 34:13–25.

Bandiera, O., Prat, A., Guiso, L., and Sadun, R. (2015b). Matching firms, managers and incentives. *Journal of Labor Economics*, 33(3):623–681.

Barankay, I. (2011a). Gender differences in productivity responses to performance rankings: Evidence from a randomized workplace experiment. Working Paper, Wharton School of Business, University of Pennsylvania.

Barankay, I. (2011b). Rankings and social tournaments: Evidence from a crowdsourcing experiment. Working Paper, Wharton School of Business, University of Pennsylvania.

Baron, J. N. and Kreps, D. M. (1999). *Strategic Human Resources: Frameworks For General Managers*. Wiley New York.

Barrick, M. R. and Mount, M. K. (1991). The Big Five personality dimensions and job performance: A meta-analysis. *Personnel Psychology*, 44(1):1–26.

Barrick, M. R. and Mount, M. K. (2012). Nature and use of personality in selection. In Schmitt, N., editor, *The Oxford Handbook of Personnel Assessment and Selection*, chapter 11, pages 225–251. Oxford University Press.

Bartling, B., Fehr, E., Maréchal, M. A., and Schunk, D. (2009). Egalitarianism and competitiveness. *American Economic Review*, 99(2):93–98.

Bartling, B., Fehr, E., and Schmidt, K. M. (2012). Screening, competition, and job design: Economic origins of good jobs. *American Economic Review*, 102(2):834–864.

Beaman, L. and Magruder, J. (2012). Who gets the job referral? Evidence from a social networks experiment. *American Economic Review*, 102(7):3574–3593.

Becker, A., Deckers, T., Dohmen, T., Falk, A., and Kosse, F. (2012). The relationship between economic preferences and psychological personality measures. *Annual Review of Economics*, 4:453–478.

Becker, B. and Gerhart, B. (1996). The impact of human resource management on organizational performance: Progress and prospects. *Academy of Management Journal*, 39(4):779–801.

Becker, B. and Huselid, M. (2006). Strategic human resources management: Where do we go from here? *Journal of Management*, 32(6):898–925.

Bellemare, C. and Shearer, B. (2009). Gift giving and worker productivity: Evidence from a firm-level experiment. *Games and Economic Behavior*, 67(1):233–244.

Bellemare, C. and Shearer, B. (2011). On the relevance and composition of gifts within the firm: Evidence from field experiments. *International Economic Review*, 52(3):855–882.

Benkard, C. L. (2000). Learning and forgetting: The dynamics of aircraft production. *American Economic Review*, 90(4):1034–1054.

Benoît, J.-P. and Dubra, J. (2011). Apparent overconfidence. *Econometrica*, 79(5):1591–1625.

Bentler, P. M. and Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, 88(3):588–606.

Berg, N. and Fast, N. (1975). The Lincoln Electric Company. Case 376-028, Harvard Business School.

Berlin, N. and Dargnies, M.-P. (2016). Gender differences in reactions to feedback and willingness to compete. *Journal of Economic Behavior & Organization*, 130:320–336.

Bernheim, B. D. (1994). A theory of conformity. *Journal of Political Economy*, 102(5):841–877.

Berninghaus, S. K. and Ehrhart, K.-M. (2001). Coordination and information: Recent experimental evidence. *Economics Letters*, 73(3):345 – 351.

Besley, T. and Ghatak, M. (2005). Competition and incentives with motivated agents. *American Economic Review*, 95(3):616–636.

Besley, T. and Ghatak, M. (2008). Status incentives. *American Economic Review*, 98(2):206–211.

Bikhchandani, S., Hirshleifer, D., and Welch, I. (1998). Learning from the behavior of others: Conformity, fads, and informational cascades. *Journal of Economic Perspectives*, 12(3):151–170.

Blanes i Vidal, J. and Nossol, M. (2011). Tournaments without prizes: Evidence from personnel records. *Management Science*, 57(10):1721–1736.

Bloom, N., Eifert, B., Mahajan, A., McKenzie, D., and Roberts, J. (2013). Does management matter? Evidence from India. *Quarterly Journal of Economics*, 128(1):1–51.

Bloom, N., Lemos, R., Sadun, R., Scur, D., and Van Reenen, J. (2014). The new empirical economics of management. *Journal of the European Economic Association*, 12(4):835–876.

Bloom, N., Sadun, R., and Reenen, J. V. (2012). Americans do IT better: US multinationals and the productivity miracle. *American Economic Review*, 102(1):167–201.

Bloom, N. and Van Reenen, J. (2007). Measuring and explaining management practices across firms and countries. *Quarterly Journal of Economics*, 122(4):1351–1408.

Bogaert, S., Boone, C., and Declerck, C. (2008). Social value orientation and cooperation in social dilemmas: A review and conceptual model. *British Journal of Social Psychology*, 47(3):453–480.

Bolger, F., Pulford, B. D., and Colman, A. M. (2008). Market entry decisions: Effects of absolute and relative confidence. *Experimental Psychology*, 55(2):113–120.

Borghans, L., Duckworth, A. L., Heckman, J. J., and Ter Weel, B. (2008). The economics and psychology of personality traits. *Journal of Human Resources*, 43(4):972–1059.

Bourdage, J. S., Lee, K., Lee, J.-H., and Shin, K.-H. (2012). Motives for organizational citizenship behavior: Personality correlates and coworker ratings of OCB. *Human Performance*, 25(3):179–200.

Bowers, K. S. (1973). Situationism in psychology: An analysis and a critique. *Psychological Review*, 80(5):307–336.

Bowles, S. (1998). Endogenous preferences: The cultural consequences of markets and other economic institutions. *Journal of Economic Literature*, 36(1):75–111.

Bradler, C., Dur, R., Neckermann, S., and Non, A. (2013). Employee recognition and performance: A field experiment. Discussion Paper 13-017, ZEW-Centre for European Economic Research.

Braun, N. (1994). The threshold model revisited. Working Paper, University of Berne.

Braun, N. (1995). Individual thresholds and social diffusion. *Rationality and Society*, 7(2):167–182.

Bridoux, F., Coeurderoy, R., and Durand, R. (2011). Heterogeneous motives and the collective creation of value. *Academy of Management Review*, 36(4):711–730.

Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. Guilford Press, New York.

Bruhin, A., Fehr-Duda, H., and Epper, T. (2010). Risk and rationality: Uncovering heterogeneity in probability distortion. *Econometrica*, 78(4):1375–1412.

Brynjolfsson, E. and Milgrom, P. (2013). Complementarity in organizations. In Gibbons, R. and Roberts, J., editors, *The Handbook of Organizational Economics*, pages 11 – 55. Princeton University Press.

Bull, C. (1987). The existence of self-enforcing implicit contracts. *Quarterly Journal of Economics*, 102(1):147–159.

Burks, S., Cowgill, B., Hoffman, M., and Housman, M. (2015). The value of hiring through employee referrals. *Quarterly Journal of Economics*, 130(2):805–839.

Buser, T., Gerhards, L., and Van Der Weele, J. J. (2016). Measuring responsiveness to feedback as a personal trait. Working Paper.

Buser, T., Niederle, M., and Oosterbeek, H. (2014). Gender, competitiveness, and career choices. *Quarterly Journal of Economics*, 129(3):1409–1447.

Cabral, L., Ozbay, E. Y., and Schotter, A. (2014). Intrinsic and instrumental reciprocity: An experimental study. *Games and Economic Behavior*, 87:100–121.

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(5):2261–2278.

Cain, D. M., Moore, D. A., and Haran, U. (2015). Making sense of overconfidence in market entry. *Strategic Management Journal*, 13(1):1–18.

Caliendo, M., Fossen, F., and Kritikos, A. (2012). Trust, positive reciprocity, and negative reciprocity: Do these traits impact entrepreneurial dynamics? *Journal of Economic Psychology*, 33(2):394–409.

Caliendo, M., Fossen, F., and Kritikos, A. S. (2014). Personality characteristics and the decisions to become and stay self-employed. *Small Business Economics*, 42(4):787–814.

Camerer, C. (2004). Behavioral economics: Past, present, future. In Camerer, C., Loewenstein, G., and Rabin, M., editors, *Advances in Behavioral Economics*. Princeton: University Press.

Camerer, C. and Lovallo, D. (1999). Overconfidence and excess entry: An experimental approach. *American Economic Review*, 89(1):306–318.

Camerer, C. and Weber, R. (2013). Experimental organizational economics. In Gibbons, R. and Roberts, J., editors, *The Handbook of Organizational Economics*, pages 213–262. Princeton University Press.

Cappelli, P. and Neumark, D. (2001). Do “high-performance" work practices improve establishment-level outcomes? *Industrial and Labor Relations Review*, 54(4):737–775.

Cassar, L. (2014). Optimal contracting with endogenous project mission. Working Paper 150, University of Zurich, Department of Economics.

Centola, D. and Macy, M. (2007). Complex contagions and the weakness of long ties. *American Journal of Sociology*, 113(3):702–734.

Charness, G. and Kuhn, P. (2011). Lab labor: What can labor economists learn from the lab? In Ashenfelter, O. and Card, D., editors, *Handbook of Labor Economics*, volume 4, pages 229–330. Elsevier.

Charness, G., Masclet, D., and Villeval, M. C. (2014). The dark side of competition for status. *Management Science*, 60(1):38–55.

Charness, G. and Rabin, M. (2002). Understanding social preferences with simple tests. *Quarterly Journal of Economics*, 117(3):817–869.

Chaudhuri, A. (2011). Sustaining cooperation in laboratory public goods experiments: A selective survey of the literature. *Experimental Economics*, 14(1):47–83.

Chen, Y., Harper, F. M., Konstan, J., and Li, S. X. (2010). Social comparisons and contributions to online communities: A field experiment on MovieLens. *American Economic Review*, 100(4):1358–1398.

Chiang, Y.-S. (2007). Birds of moderately different feathers: Bandwagon dynamics and the threshold heterogeneity of network neighbors. *Journal of Mathematical Sociology*, 31(1):47–69.

Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*, 6(4):284–290.

Clark, C. B., Thorne, C. B., Vann, J., and Cropsey, K. L. (2014). Five-factor personality and cooperative behavior. *North American Journal of Psychology*, 16(3):481–493.

Cobb-Clark, D. A. and Tan, M. (2011). Noncognitive skills, occupational attainment, and relative wages. *Labour Economics*, 18(1):1–13.

Cohn, A., Fehr, E., and Goette, L. (2015). Fair wages and effort: Evidence from a field experiment. *Management Science*, 61(8):1777–1794.

Colquitt, J. A. (2008). From the editors publishing laboratory research in AMJ: A question of when, not if. *Academy of Management Journal*, 51(4):616–620.

Contreras, O. F. and Zanarone, G. (2017). Managing social comparison costs in organizations. Working Paper.

Croson, R. and Marks, M. (1998). Identifiability of individual contributions in a threshold public goods experiment. *Journal of Mathematical Psychology*, 42(2–3):167–190.

Croson, R. T. A. and Marks, M. B. (2000). Step returns in threshold public goods: A meta- and experimental analysis. *Experimental Economics*, 2(3):239–259.

Cyert, R. M. and March, J. G. (1963). *A Behavioral Theory of the Firm*, volume 2. Prentice-Hall.

Dalal, R. S., Meyer, R. D., Bradshaw, R. P., Green, J. P., Kelly, E. D., and Zhu, M. (2015). Personality strength and situational influences on behavior: A conceptual review and research agenda. *Journal of Management*, 41(1):261–287.

Dannenberg, A. (2015). Leading by example versus leading by words in voluntary contribution experiments. *Social Choice and Welfare*, 44(1):71–85.

Dawes, R. M. (1980). Social dilemmas. *Annual Review of Psychology*, 31(1):169–193.

Dawes, R. M. and Messick, D. M. (2000). Social dilemmas. *International Journal of Psychology*, 35(2):111–116.

Dechenaux, E., Kovenock, D., and Sheremeta, R. M. (2015). A survey of experimental research on contests, all-pay auctions and tournaments. *Experimental Economics*, 18(4):609–669.

Delery, J. E. (1998). Issues of fit in strategic human resource management: Implications for research. *Human Resource Management Review*, 8(3):289–309.

Delery, J. E. and Doty, D. H. (1996). Modes of theorizing in strategic human resource management: Tests of universalistic, contingency, and configurational performance predictions. *Academy of Management Journal*, 39(4):802–835.

Delfgaauw, J. and Dur, R. (2007). Signaling and screening of workers' motivation. *Journal of Economic Behavior & Organization*, 62(4):605–624.

DellaVigna, S. (2009). Psychology and economics: Evidence from the field. *Journal of Economic Literature*, 47(2):315–372.

Dickinson, D. and Villeval, M.-C. (2008). Does monitoring decrease work effort? The complementarity between agency and crowding-out theories. *Games and Economic Behavior*, 63(1):56–76.

Diekmann, A. and Przepiorka, W. (2016). "Take one for the team!" Individual heterogeneity and the emergence of latent norms in a volunteer's dilemma. *Social Forces*, 94(3):1309–1333.

Dijkstra, J. (2013). Put your money where your mouth is: Reciprocity, social preferences, trust and contributions to public goods. *Rationality and Society*, 25(3):290–334.

Dohmen, T. and Falk, A. (2011). Performance pay and multidimensional sorting: Productivity, preferences, and gender. *American Economic Review*, 101(2):556–590.

Dohmen, T., Falk, A., Huffman, D., and Sunde, U. (2008). Representative trust and reciprocity: Prevalence and determinants. *Economic Inquiry*, 46(1):84–90.

Dohmen, T., Falk, A., Huffman, D., and Sunde, U. (2009). Homo reciprocans: Survey evidence on behavioural outcomes. *Economic Journal*, 119(536):592–612.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., and Wagner, G. G. (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association*, 9(3):522–550.

Dorsey, R. E. (1992). The voluntary contributions mechanism with real time revisions. *Public Choice*, 73(3):261–282.

Dreber, A., Fudenberg, D., and Rand, D. G. (2014). Who cooperates in repeated games: The role of altruism, inequity aversion, and demographics. *Journal of Economic Behavior & Organization*, 98:41–55.

Ducharme, W. M. and Donnell, M. L. (1973). Intrasubject comparison of four response modes for "subjective probability" assessment. *Organizational Behavior and Human Performance*, 10(1):108–117.

Dufwenberg, M. and Kirchsteiger, G. (2004). A theory of sequential reciprocity. *Games and Economic Behavior*, 47(2):268–298.

Dur, R., Non, A., and Roelfsema, H. (2010). Reciprocity and incentive pay in the workplace. *Journal of Economic Psychology*, 31(4):676–686.

Dur, R. and Tichem, J. (2015). Altruism and relational incentives in the workplace. *Journal of Economics & Management Strategy*, 24(3):485–500.

Eckel, C. C. and Grossman, P. J. (2008). Men, women and risk aversion: Experimental evidence. In Plott, C. R. and Smith, V. L., editors, *Handbook of Experimental Economics Results*, volume 1, pages 1061–1073. Elsevier.

Ellingsen, T. and Johannesson, M. (2007). Paying respect. *Journal of Economic Perspectives*, 21(4):135–150.

Englmaier, F., Fahn, M., and Schwarz, M. A. (2015). Long-term employment relations when agents are present-biased. Working Paper, LMU Munich.

Englmaier, F., Kolaska, T., and Leider, S. (2016). Reciprocity in organizations: Evidence from the UK. *CESifo Economic Studies*, 62(3):522–546.

Englmaier, F. and Leider, S. (2012a). Contractual and organizational structure with reciprocal agents. *American Economic Journal: Microeconomics*, 4(2):146–183.

Englmaier, F. and Leider, S. (2012b). Gift exchange in the lab – it is not (only) how much you give... CESifo Working Paper 2944.

Englmaier, F. and Leider, S. G. (2012c). Managerial payoff and gift exchange in the field. CESifo Working Paper 3707.

Englmaier, F. and Schüßler, K. (2016). Complementarities of human-resource management practices: A case for a behavioral-economics perspective. *Journal of Institutional and Theoretical Economics*, 172(2):312–341.

Englmaier, F., Strasser, S., and Winter, J. (2014). Worker characteristics and wage differentials: Evidence from a gift-exchange experiment. *Journal of Economic Behavior & Organization*, 97:185–203.

Enke, B. and Zimmermann, F. (2017). Correlation neglect in belief formation. Working Paper.

Eriksson, T., Teyssier, S., and Villeval, M.-C. (2009). Self-selection and the efficiency of tournaments. *Economic Inquiry*, 47(3):530–548.

Esponda, I. (2008). Behavioral equilibrium in economies with adverse selection. *American Economic Review*, 98(4):1269–1291.

Evans, J. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology*, 59:255–278.

Evans, J. and Stanovich, K. (2013). Dual-process theories of higher cognition: Advancing the debate. *Perspectives on Psychological Science*, 8(3):223–241.

Eyster, E. and Rabin, M. (2005). Cursed equilibrium. *Econometrica*, 73(5):1623–1672.

Falk, A., Fehr, E., and Fischbacher, U. (2008). Testing theories of fairness – intentions matter. *Games and Economic Behavior*, 62(1):287–303.

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

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

Falk, A. and Kosfeld, M. (2006). The hidden costs of control. *American Economic Review*, 96(5):1611–1630.

Farrell, J. and Rabin, M. (1996). Cheap talk. *Journal of Economic Perspectives*, 10(3):103–118.

Fehr, E., Brown, M., and Zehnder, C. (2009a). On reputation: A microfoundation of contract enforcement and price rigidity. *Economic Journal*, 119(536):333–353.

Fehr, E. and Falk, A. (2002). Psychological foundations of incentives. *European Economic Review*, 46(4):687–724.

Fehr, E., Goette, L., and Zehnder, C. (2009b). A behavioral account of the labor market: The role of fairness concerns. *Annual Review of Economics*, 1(1):355–384.

Fehr, E., Kirchsteiger, G., and Riedl, A. (1993). Does fairness prevent market clearing? An experimental investigation. *Quarterly Journal of Economics*, 108(2):437–459.

Fehr, E., Kirchsteiger, G., and Riedl, A. (1998). Gift exchange and reciprocity in competitive experimental markets. *European Economic Review*, 42(1):1–34.

Fehr, E. and Schmidt, K. (2006). The economics of fairness, reciprocity and altruism – experimental evidence and new theories. In Kolm, S. and Ythier, J., editors, *Handbook on the Economics of Giving, Reciprocity and Altruism*, pages 615–691. Amsterdam: Elsevier.

Fischbacher, U. (2007). z-tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, 10(2):171–178.

Flory, J. A., Leibbrandt, A., and List, J. A. (2015). Do competitive workplaces deter female workers? A large-scale natural field experiment on job-entry decisions. *Review of Economic Studies*, 82(1):122–155.

Fong, Y.-f. and Li, J. (2017). Relational contracts, limited liability, and employment dynamics. *Journal of Economic Theory*, 169:270–293.

Friedman, D. and Oprea, R. (2012). A continuous dilemma. *American Economic Review*, 102(1):337–363.

Fung, J. M. Y., Au, W. T., Hu, W., and Shi, K. (2012). Effect of risk orientation on cooperation and decision process in public goods dilemma. *Group Processes & Intergroup Relations*, 15(6):791–803.

Gächter, S. and Falk, A. (2002). Reputation and reciprocity: Consequences for the labour relation. *Scandinavian Journal of Economics*, 104(1):1–26.

Gerhart, B. (2007). Horizontal and vertical fit in human resource systems. In Ostroff, C. and Judge, T. A., editors, *Perspectives on Organizational Fit*, pages 317–348. Lawrence Erlbaum Associates, Taylor & Francis Group New York.

Gersick, C. J. G. (1988). Time and transition in work teams: Toward a new model of group development. *Academy of Management Journal*, 31(1):9–41.

Gibbons, R. and Henderson, R. (2012). Relational contracts and organizational capabilities. *Organization Science*, 23(5):1350–1364.

Gneezy, U., Güth, W., and Verboven, F. (2000). Presents or investments? An experimental analysis. *Journal of Economic Psychology*, 21(5):481–493.

Gneezy, U., Niederle, M., and Rustichini, A. (2003). Performance in competitive environments: Gender differences. *Quarterly Journal of Economics*, 118(3):1049–1074.

Goren, H., Kurzban, R., and Rapoport, A. (2003). Social loafing vs. social enhancement: Public goods provisioning in real-time with irrevocable commitments. *Organizational Behavior and Human Decision Processes*, 90(2):277–290.

Goren, H., Rapoport, A., and Kurzban, R. (2004). Revocable commitments to public goods provision under the real-time protocol of play. *Journal of Behavioral Decision Making*, 17(1):17–37.

Granovetter, M. (1978). Threshold models of collective behavior. *American Journal of Sociology*, 83(6):1420–1443.

Granovetter, M. and Soong, R. (1983). Threshold models of diffusion and collective behavior. *The Journal of Mathematical Sociology*, 9(3):165–179.

Grant, A. M. (2008). The significance of task significance: Job performance effects, relational mechanisms, and boundary conditions. *Journal of Applied Psychology*, 93(1):108–124.

Greiner, B. (2015). Subject pool recruitment procedures: Organizing experiments with ORSEE. *Journal of the Economic Science Association*, 1(1):114–125.

Grossman, S. J. and Hart, O. D. (1983). An analysis of the principal-agent problem. *Econometrica*, 51(1):7–45.

Grossman, Z. and Owens, D. (2012). An unlucky feeling: Overconfidence and noisy feedback. *Journal of Economic Behavior & Organization*, 84(2):510–524.

Gubler, T., Larkin, I., and Pierce, L. (2013). The dirty laundry of employee award programs: Evidence from the field. Harvard Business School NOM Unit Working Paper 13–069.

Hamilton, B. H., Nickerson, J. A., and Owan, H. (2003). Team incentives and worker heterogeneity: An empirical analysis of the impact of teams on productivity and participation. *Journal of Policital Economy*, 111(3):465–497.

Harrison, G. W., Martínez-Correa, J., Swarthout, J. T., and Ulm, E. R. (2015). Eliciting subjective probability distributions with binary lotteries. *Economics Letters*, 127:68–71.

Harrison, G. W., Martínez-Correa, J., Swarthout, J. T., and Ulm, E. R. (2017). Scoring rules for subjective probability distributions. *Journal of Economic Behavior & Organization*, 134:430–448.

Heineck, G. and Anger, S. (2010). The returns to cognitive abilities and personality traits in Germany. *Labour Economics*, 17(3):535–546.

Hennig-Schmidt, H., Rockenbach, B., and Sadrieh, A. (2010). In search of workers' real effort reciprocity – a field and a laboratory experiment. *Journal of the European Economic Association*, 8(4):817–837.

Highhouse, S. (2009). Designing experiments that generalize. *Organizational Research Methods*, 12(3):554–566.

Hilbig, B. E., Glöckner, A., and Zettler, I. (2014). Personality and prosocial behavior: Linking basic traits and social value orientations. *Journal of Personality and Social Psychology*, 107(3):529–539.

Hilbig, B. E. and Zettler, I. (2009). Pillars of cooperation: Honesty–humility, social value orientations, and economic behavior. *Journal of Research in Personality*, 43(3):516–519.

Hilbig, B. E., Zettler, I., and Heydasch, T. (2012). Personality, punishment and public goods: Strategic shifts towards cooperation as a matter of dispositional honesty–humility. *European Journal of Personality*, 26(3):245–254.

Hilbig, B. E., Zettler, I., Leist, F., and Heydasch, T. (2013a). It takes two: Honesty–humility and agreeableness differentially predict active versus reactive cooperation. *Personality and Individual Differences*, 54(5):598–603.

Hilbig, B. E., Zettler, I., Moshagen, M., and Heydasch, T. (2013b). Tracing the path from personality — via cooperativeness — to conservation. *European Journal of Personality*, 27(4):319–327.

Hölmstrom, B. (1979). Moral hazard and observability. *Bell Journal of Economics*, 10(1):74–91.

Hölmstrom, B. and Milgrom, P. (1994). The firm as an incentive system. *American Economic Review*, 84(4):972–991.

Holt, C. A. and Laury, S. K. (2002). Risk aversion and incentive effects. *American Economic Review*, 92(5):1644–1655.

Hoogveld, N. and Zubanov, N. (2014). The power of (no) recognition: Experimental evidence from the university classroom. IZA Working Paper 7953.

Hopwood, C. J. and Donnellan, M. B. (2010). How should the internal structure of personality inventories be evaluated? *Personality and Social Psychology Review*, 14(3):332–346.

Hough, L. M., Oswald, F. L., and Ock, J. (2015). Beyond the Big Five: New directions for personality research and practice in organizations. *Annual Review of Organizational Psychology and Organizational Behavior*, 2(1):183–209.

Hsieh, C.-T. and Klenow, P. (2009). Misallocation and manufacturing TFP in China and India. *Quarterly Journal of Economics*, 124(4):1403–1448.

Huang, F. and Cappelli, P. (2010). Applicant screening and performance-related outcomes. *American Economic Review*, 100(2):214–218.

Huselid, M. A. (1995). The impact of human resource management practices on turnover, productivity, and corporate financial performance. *Academy of Management Journal*, 38(3):635–672.

Ichniowski, C. and Shaw, K. (2003). Beyond incentive pay: Insiders' estimates of the value of complementary human resource management practices. *Journal of Economic Perspectives*, 17(1):155–180.

Ichniowski, C. and Shaw, K. (2013). Insider econometrics. In Gibbons, R. and Roberts, J., editors, *The Handbook of Organizational Economics*, pages 263 – 314. Princeton University Press.

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(3):291–313.

Innes, R. D. (1990). Limited liability and incentive contracting with ex-ante action choices. *Journal of Economic Theory*, 52(1):45–67.

Jackson, S. E., Schuler, R. S., and Jiang, K. (2014). An aspirational framework for strategic human resource management. *Academy of Management Annals*, 8(1):1–56.

Jehiel, P. (2016). Investment strategy and selection bias: An equilibrium perspective on overoptimism. Working Paper.

Jiang, K., Lepak, D. P., Han, K., Hong, Y., Kim, A., and Winkler, A.-L. (2012). Clarifying the construct of human resource systems: relating human resource management to employee performance. *Human Resource Management Review*, 22(2):73–85.

John, O. P., Naumann, L. P., and Soto, C. J. (2008). Paradigm shift to the integrative big five trait taxonomy: History, measurement, and conceptual issues. In John, O. P., Robbins, R. W., and Pervin, L. A., editors, *Handbook of Personality: Theory and Research*, pages 114–156. Guilford, New York.

Jones, D., Willness, C., and Madey, S. (2014). Why are job seekers attracted by corporate social performance? Experimental and field tests of three signal-based mechanisms. *Academy of Management Journal*, 57(2):383–404.

Judge, T. A. and Zapata, C. P. (2015). The person–situation debate revisited: Effect of situation strength and trait activation on the validity of the Big Five personality traits in predicting job performance. *Academy of Management Journal*, 58(4):1149–1179.

Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan.

Kamas, L. and Preston, A. (2012). The importance of being confident; gender, career choice, and willingness to compete. *Journal of Economic Behavior & Organization*, 83(1):82–97.

Kampkötter, P. and Sliwka, D. (2016). On the complementary use of experiments and field data to evaluate management practices - the case of subjective performance evaluations. *Journal of Institutional and Theoretical Economics*, 172(2):364–389.

Kanfer, R. and Ackerman, P. L. (1989). Motivation and cognitive abilities: An integrative/aptitude-treatment interaction approach to skill acquisition. *Journal of Applied Psychology*, 74(4):657–690.

Kelley, H. H. and Thibaut, J. W. (1978). *Interpersonal relations: A theory of interdependence*. John Wiley & Sons.

Kepes, S. and Delery, J. E. (2007). HRM systems and the problem of internal fit. In Boxall, P., Purcell, J., and Wright, P., editors, *Oxford Handbook of Human Resource Management*, pages 385–404. Oxford University Press.

Koellinger, P., Minniti, M., and Schade, C. (2007). "I think I can, I think I can": Overconfidence and entrepreneurial behavior. *Journal of Economic Psychology*, 28(4):502–527.

Kollock, P. (1998). Social dilemmas: The anatomy of cooperation. *Annual Review of Sociology*, 24(1):183–214.

Komorita, S. S. and Parks, C. D. (1995). Interpersonal relations: Mixed-motive interaction. *Annual Review of Psychology*, 46(1):183–207.

Kosfeld, M. and Neckermann, S. (2011). Getting more work for nothing? Symbolic awards and worker performance. *American Economic Journal: Microeconomics*, 3(3):86–99.

Kosfeld, M., Neckermann, S., and Yang, X. (2014). Knowing that you matter, matters! The interplay of meaning, monetary incentives, and worker recognition. Discussion Paper 14-097, ZEW-Centre for European Economic Research.

Kreps, D., Milgrom, P., Roberts, J., and Wilson, R. (1982). Rational cooperation in the finitely repeated prisoners' dilemma. *Journal of Economic Theory*, 27(2):245–252.

Kristof, A. L. (1996). Person-organization fit: An integrative review of its conceptualizations, measurement, and implications. *Personnel Psychology*, 49(1):1–49.

Kurzban, R., McCabe, K., Smith, V. L., and Wilson, B. J. (2001). Incremental commitment and reciprocity in a real-time public goods game. *Personality and Social Psychology Bulletin*, 27(12):1662–1673.

Lange, P. A. V., Joireman, J., Parks, C. D., and Dijk, E. V. (2013). The psychology of social dilemmas: A review. *Organizational Behavior and Human Decision Processes*, 120(2):125–141.

Larkin, I. and Leider, S. (2012). Incentive schemes, sorting, and behavioral biases of employees: experimental evidence. *American Economic Journal: Microeconomics*, 4(2):184–214.

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

Lazear, E. P. and Shaw, K. L. (2007). Personnel economics: The economist's view of human resources. *Journal of Economic Perspectives*, 21(4):91–114.

Lechner, C. and Kreutzer, M. (2011). Strategic initiatives: Past, present and future. In Mazzola, P. and Kellermanns, F.-W., editors, *Handbook of Research on Strategy Process*, pages 283–303. Edward Elgar, Cheltenham, UK.

Ledyard, J. O. (1995). Public goods: A survey of experimental research. In Kagel, J. H. and Roth, A. E., editors, *The Handbook of Experimental Economics*, page 111–194. Princeton university press Princeton, NJ.

Leng, A., Friesen, L., Kalaycı, K., and Man, P. (2016). A minimum effort coordination game experiment in continuous time. Working Paper.

Leuven, E., Oosterbeek, H., Sloof, R., and Van Klaveren, C. (2005). Worker reciprocity and employer investment in training. *Economica*, 72(285):137–149.

Levin, J. (2003). Relational incentive contracts. *American Economic Review*, 93(3):835–857.

Li, J., Matouschek, N., and Powell, M. (2017). Power dynamics in organizations. *American Economic Journal: Microeconomics*, 9(1):217–241.

List, J. A. (2011). Why economists should conduct field experiments and 14 tips for pulling one off. *The Journal of Economic Perspectives*, 25(3):3–15.

Lockett, A. and Thompson, S. (2001). The resource-based view and economics. *Journal of Management*, 27(6):723–754.

MacLeod, W. B. and Malcomson, J. M. (1989). Implicit contracts, incentive compatibility, and involuntary unemployment. *Econometrica*, 57(2):447–480.

Mailath, G. J. and Samuelson, L. (2006). *Repeated games and reputations: Long-run relationships*. Oxford University Press.

Malcomson, J. M. (2013). Relational incentive contracts. In Gibbons, R. and Roberts, J., editors, *The Handbook of Organizational Economics*, pages 1014 – 1065. Princeton University Press.

Marks, M. B. and Croson, R. T. (1999). The effect of incomplete information in a threshold public goods experiment. *Public Choice*, 99(1):103–118.

Mäs, M. and Nax, H. H. (2016). A behavioral study of "noise" in coordination games. *Journal of Economic Theory*, 162:195–208.

McAfee, A. (2002). The impact of enterprise information technology adoption on operational performance: An empirical investigation. *Production and Operations Management*, 11(1):33–53.

McClintock, C. G. and Allison, S. T. (1989). Social value orientation and helping behavior. *Journal of Applied Social Psychology*, 19(4):353–362.

McCrae, R. R. and Costa, P. T. (1987). Validation of a five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology*, 52(1):81–90.

McCrae, R. R., Zonderman, A. B., Costa, P. T., Bond, M. H., and Paunonen, S. V. (1996). Evaluating replicability of factors in the revised NEO personality inventory: Confirmatory factor analysis versus procrustes rotation. *Journal of Personality and Social Psychology*, 70(3):552–566.

Merkle, C. and Weber, M. (2011). True overconfidence: The inability of rational information processing to account for apparent overconfidence. *Organizational Behavior and Human Decision Processes*, 116(2):262–271.

Meyer, R. D., Dalal, R. S., and Hermida, R. (2010). A review and synthesis of situational strength in the organizational sciences. *Journal of Management*, 36(1):121–140.

Milgrom, P. and Roberts, J. (1992). *Economics, organization and management*. Prentice-Hall International.

Milgrom, P. and Roberts, J. (1995). Complementarities and fit strategy, structure, and organizational change in manufacturing. *Journal of Accounting and Economics*, 19(2):179–208.

Mischel, W. (1968). *Personality and assessment*. Wiley, New York.

Mischel, W. and Shoda, Y. (1995). A cognitive-affective system theory of personality: Reconceptualizing situations, dispositions, dynamics, and invariance in personality structure. *Psychological Review*, 102(2):246–268.

Mischel, W. and Shoda, Y. (2010). Toward a unified theory of personality: Integrating dispositions and processing dynamics within the cognitive-affective processing system (CAPS). In Oliver, J., Robinson, R., and Pervin, L., editors, *Handbook of Personality: Theory and Research*, pages 208–241. The Guilford Press, New York.

Möbius, M. M., Niederle, M., Niehaus, P., and Rosenblat, T. S. (2014). Managing self-confidence. Working Paper.

Moffitt, T. E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R. J., Harrington, H., Houts, R., Poulton, R., Roberts, B. W., Ross, S., Sears, M., Thomson, W., and Caspi, A. (2011). A gradient of childhood self-control predicts health, wealth, and public safety. *Proceedings of the National Academy of Sciences*, 108(7):2693–2698.

Moore, D. A. and Healy, P. J. (2008). The trouble with overconfidence. *Psychological Review*, 115(2):502–517.

Moore, D. A., Oesch, J. M., and Zietsma, C. (2007). What competition? Myopic self-focus in market-entry decisions. *Organization Science*, 18(3):440–454.

Moshagen, M., Hilbig, B. E., and Zettler, I. (2014). Faktorenstruktur, psychometrische Eigenschaften und Messinvarianz der deutschsprachigen Version des 60-Item HEXACO Persönlichkeitsinventars. *Diagnostica*, 60(2):86–97.

Mueller, G. and Plug, E. (2006). Estimating the effect of personality on male and female earnings. *ILR Review*, 60(1):3–22.

Murphy, R. O. and Ackermann, K. A. (2014). Social value orientation. *Personality and Social Psychology Review*, 18(1):13–41.

Murphy, R. O., Ackermann, K. A., and Handgraaf, M. J. (2011). Measuring social value orientation. *Judgment and Decision Making*, 6(8):771–781.

Netzer, N. and Schmutzler, A. (2014). Explaining gift-exchange - the limits of good intentions. *Journal of the European Economic Association*, 12(6):1586–1616.

Niederle, M. and Vesterlund, L. (2007). Do women shy away from competition? Do men compete too much? *Quarterly Journal of Economics*, 122(3):1067–1101.

Niederle, M. and Vesterlund, L. (2011). Gender and competition. *Annual Review of Economics*, 3(1):601–630.

O'Donoghue, T. and Rabin, M. (1999a). Doing it now or later. *American Economic Review*, 89(1):103–124.

O'Donoghue, T. and Rabin, M. (1999b). Incentives for procrastinators. *Quarterly Journal of Economics*, 114(3):769–816.

Ohlendorf, S. and Schmitz, P. W. (2012). Repeated moral hazard and contracts with memory: The case of risk-neutrality. *International Economic Review*, 53(2):433–452.

Okhuysen, G. A. and Bechky, B. A. (2009). 10 coordination in organizations: An integrative perspective. *Academy of Management Annals*, 3(1):463–502.

Oliver, P. E. (1993). Formal models of collective action. *Annual Review of Sociology*, 19(1):271–300.

Onyemah, V. (2008). Role ambiguity, role conflict, and performance: Empirical evidence of an inverted-u relationship. *Journal of Personal Selling & Sales Management*, 28(3):299–313.

Oprea, R., Charness, G., and Friedman, D. (2014). Continuous time and communication in a public-goods experiment. *Journal of Economic Behavior & Organization*, 108:212–223.

Organ, D. W. (1994). Personality and organizational citizenship behavior. *Journal of Management*, 20(2):465–478.

Ostrom, E. and Walker, J. (2003). *Trust and reciprocity: Interdisciplinary lessons for experimental research*. Russell Sage Foundation.

Oswald, F. L., Hough, L., and Ock, J. (2013). Theoretical and empirical structures of personality. In Christiansen, N. and Tett, R., editors, *Handbook of Personality at Work*, pages 11–29. Routledge.

Oyer, P. (2008). Salary of benefits? In Polachek, S. W. and Tatsiramos, K., editors, *Work, Earnings and Other Aspects of the Employment Relation (Research in Labor Economics, Volume 28)*, volume 28, pages 429–467. Emerald Group Publishing Limited.

Ozer, D. J. and Benet-Martînez, V. (2006). Personality and the prediction of consequential outcomes. *Annual Review of Psychology*, 57:401–421.

Pallais, A. and Sands, E. G. (2016). Why the referential treatment? Evidence from field experiments on referrals. *Journal of Political Economy*, 124(6):1793–1828.

Perugini, M., Gallucci, M., Presaghi, F., and Ercolani, A. P. (2003). The personal norm of reciprocity. *European Journal of Personality*, 17(4):251–283.

Pfeffer, J. (1994). *Competitive advantage through people: Unleashing the power of the work force*. Harvard Business School Press, Boston.

Pfeffer, J. (2007). Human resources from an organizational behavior perspective: Some paradoxes explained. *Journal of Economic Perspectives*, 21(4):115–134.

Potters, J., Sefton, M., and Vesterlund, L. (2005). After you—endogenous sequencing in voluntary contribution games. *Journal of Public Economics*, 89(8):1399–1419.

Potters, J., Sefton, M., and Vesterlund, L. (2007). Leading-by-example and signaling in voluntary contribution games: an experimental study. *Economic Theory*, 33(1):169–182.

Puranam, P., Alexy, O., and Reitzig, M. (2014). What's "new" about new forms of organizing? *Academy of Management Review*, 39(2):162–180.

Puranam, P., Raveendran, M., and Knudsen, T. (2012). Organization design: The epistemic interdependence perspective. *Academy of Management Review*, 37(3):419–440.

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

Rabin, M. (1998). Psychology and economics. *Journal of Economic Literature*, 36(1):11–46.

Raykov, T. (1998). On the use of confirmatory factor analysis in personality research. *Personality and Individual Differences*, 24(2):291–293.

Reuben, E. and Riedl, A. (2013). Enforcement of contribution norms in public good games with heterogeneous populations. *Games and Economic Behavior*, 77(1):122–137.

Reuben, E. and Suetens, S. (2012). Revisiting strategic versus non-strategic cooperation. *Experimental Economics*, 15(1):24–43.

Rivas, M. F. and Sutter, M. (2011). The benefits of voluntary leadership in experimental public goods games. *Economics Letters*, 112(2):176–178.

Rondeau, D., Poe, G. L., and Schulze, W. D. (2005). VCM or PPM? A comparison of the performance of two voluntary public goods mechanisms. *Journal of Public Economics*, 89(8):1581–1592.

Ross, S. R., Rausch, M. K., and Canada, K. E. (2003). Competition and cooperation in the five-factor model: Individual differences in achievement orientation. *Journal of Psychology*, 137(4):323–337.

Rustichini, A. (2009). Neuroeconomics: What have we found, and what should we search for. *Current Opinion in Neurobiology*, 19(6):672–677.

Rustichini, A., DeYoung, C. G., Anderson, J. C., and Burks, S. V. (2012). Toward the integration of personality theory and decision theory in the explanation of economic and health behavior. Working Paper.

Sautmann, A. (2013). Contracts for agents with biased beliefs: Some theory and an experiment. *American Economic Journal: Microeconomics*, 5(3):124–156.

Scheider, B. (1987). The people make the place. *Personnel Psychology*, 40(3):437–453.

Schein, E. H. (2015). Organizational psychology then and now: Some observations. *Annual Review of Organizational Psychology and Organizational Behavior*, 2(1):1–19.

Schlag, K. H., Tremewan, J., and Van der Weele, J. J. (2015). A penny for your thoughts: A survey of methods for eliciting beliefs. *Experimental Economics*, 18(3):457–490.

Schuler, R. S. (1989). Strategic human resource management and industrial relations. *Human Relations*, 42(2):157–184.

Schuler, R. S. and Jackson, S. E. (1987). Linking competitive strategies with human resource management practices. *Academy of Management Executive*, 1(3):207–219.

Segal, C. (2012). Working when no one is watching: Motivation, test scores, and economic success. *Management Science*, 58(8):1438–1457.

Shearer, B. (2004). Piece rates, fixed wages and incentives: Evidence from a field experiment. *Review of Economic Studies*, 71(2):513–534.

Simon, B., Trötschel, R., and Dähne, D. (2008). Identity affirmation and social movement support. *European Journal of Social Psychology*, 38(6):935–946.

Simon, L. K. and Stinchcombe, M. B. (1989). Extensive form games in continuous time: Pure strategies. *Econometrica*, 57(5):1171–1214.

Simonsohn, U. (2010). eBay's crowded evenings: Competition neglect in market entry decisions. *Management Science*, 56(7):1060–1073.

Simpson, B. (2006). Social identity and cooperation in social dilemmas. *Rationality and Society*, 18(4):443–470.

Sliwka, D. (2007). Trust as a signal of a social norm and the hidden costs of incentive schemes. *American Economic Review*, 97(3):999–1012.

Smelser, N. J. (1963). *Theory of collective behavior*. Free Press of Glencoe, New York.

Smith, C., Organ, D. W., and Near, J. P. (1983). Organizational citizenship behavior: Its nature and antecedents. *Journal of Applied Psychology*, 68(4):653.

Sobel, J. (2005). Interdependent preferences and reciprocity. *Journal of Economic Literature*, 43(2):392–436.

Stigler, G. J. and Becker, G. S. (1977). De gustibus non est disputandum. *American Economic Review*, 67(2):76–90.

Sutter, M., Haigner, S., and Kocher, M. G. (2010). Choosing the carrot or the stick? Endogenous institutional choice in social dilemma situations. *Review of Economic Studies*, 77(4):1540.

Syverson, C. (2004). Product substitutability and productivity dispersion. *Review of Economics and Statistics*, 86(2):534–550.

Syverson, C. (2011). What determines productivity? *Journal of Economic Literature*, 49(2):326–365.

Tan, H. H. and Tan, M. L. (2008). Organizational citizenship behavior and social loafing: The role of personality, motives, and contextual factors. *Journal of Psychology*, 142(1):89–108.

Teyssier, S. (2012). Inequity and risk aversion in sequential public good games. *Public Choice*, 151(1):91–119.

Thaler, R. (2008). A short course in behavioral economics. EDGE master class 2008. retrieved from *https://www.edge.org/3rd_culture/thaler_sendhil08/class6.html* on September 11, 2017.

Udéhn, L. (1993). Twenty-five years with "the logic of collective action". *Acta Sociologica*, 36(3):239–261.

van Veldhuizen, R. (2017). Gender differences in tournament choices: Risk preferences, overconfidence or competitiveness? Working Paper.

Vandegrift, D., Yavas, A., and Brown, P. M. (2007). Incentive effects and overcrowding in tournaments: An experimental analysis. *Experimental Economics*, 10(4):345–368.

Watts, D. J. and Dodds, P. (2011). Threshold models of social influence. In Hedström, P. and Bearman, P. S., editors, *The Oxford Handbook of Analytical Sociology*, pages 475–497. Oxford: Oxford University Press.

Wolf, E. and Zwick, T. (2008). Reassessing the productivity impact of employee involvement and financial incentives. *Schmalenbach Business Review*, 60(2):160–181.

Woods, S. A., Lievens, F., De Fruyt, F., and Wille, B. (2013). Personality across working life: The longitudinal and reciprocal influences of personality on work. *Journal of Organizational Behavior*, 34(S1):S7–S25.

Wozniak, D., Harbaugh, W. T., and Mayr, U. (2014). The menstrual cycle and performance feedback alter gender differences in competitive choices. *Journal of Labor Economics*, 32(1):161–198.

Zafar, B. (2011). An experimental investigation of why individuals conform. *European Economic Review*, 55(6):774–798.

Zelmer, J. (2003). Linear public goods experiments: A meta-analysis. *Experimental Economics*, 6(3):299–310.

Zhao, K. and Smillie, L. D. (2015). The role of interpersonal traits in social decision making. *Personality and Social Psychology Review*, 19(3):277–302.