Inequity Aversion, Overconfidence, and Group Performance

Inauguraldissertation

zur

Erlangung des Doktorgrades

der

Wirtschafts- und Sozialwissenschaftlichen Fakultät

der

Universität zu Köln

2012

vorgelegt von

Diplom-Volkswirt Nannan Zhou

aus

Shanghai

Referent:	Professor Dr. Dirk Sliwka
Korreferent:	Professor Dr. Oliver Gürtler
Tag der Promotion:	03.02.2012

Für meine Eltern und für Anne

Danksagung

Viele Menschen haben zur Entstehung dieser Arbeit beigetragen. Mein erster und besonderer Dank geht an meinen Doktorvater Dirk Sliwka, der mich von Anfang an mit außergewöhnlichem Engagement unterstützt und mir bei meinen unzähligen Fragen stets mit hilfreichen Ratschlägen und Ideen zur Seite gestanden hat. Er ist für mich in der zurückliegenden Zeit weit mehr als ein Mentor gewesen. Außerdem danke ich Oliver Gürtler für die freundliche Unterstützung und die Bereitschaft meine Doktorarbeit als Zweitgutachter zu bewerten.

Ich danke auch meinen Koautoren Christian Feige, Chong-Dae Kim, Felix Kölle, Petra Nieken und Abdolkarim Sadrieh für die vielen interessanten Diskussionen und hilfreichen Ratschläge.

Nicht minder möchte ich mich bei meinen aktuellen und ehemaligen Kolleginnen und Kollegen Johannes Berger, Torsten Biemann, Kathrin Breuer, Anastasia Danilov, Christine Harbring, Claus Herbertz, Patrick Kampkötter und Tom McKenzie für die sehr gute Zusammenarbeit und die freundschaftliche Atmosphäre am Lehrstuhl bedanken. Darüber hinaus danke ich unserer Sekretärin Beate Ommer sowie unseren wissenschaftlichen und studentischen Hilfskräften Philip Aminski, Isabella Cetnarowski, Paul Cibu-Deaconu, Alexander Creson, Behnud Djawadi, Tobias Hinze, David Hirschfeld, Thorsten Hormesch, Elmar Janssen, Naum Kocherovskiy, Katharina Laske, Christian Ruppert, Christiane Schäfer, Ursula Schuh, Julia Sohn, Andreas Staffeld und Timo Vogelsang für ihre Unterstützung und Hilfsbereitschaft.

Ein ganz besonderer Dank geht an meine Eltern, die mich immer bedingungslos unterstützt und in all den zurückliegenden Jahren alles Mögliche für mich getan haben. Ich danke meinem Vater, dass er vor 25 Jahren den großen Schritt gewagt hat, in ein fremdes Land zu gehen. Meiner Mutter danke ich, dass sie vor 16 Jahren ihre geliebte Heimat hinter sich gelassen hat, um mich in ein neues Leben zu begleiten.

Ein großer Dank geht auch an meine Schwiegereltern, die mich bei der Entscheidung zur Promotion ermutigt haben. Ohne sie hätte ich diese Arbeit vermutlich nie begonnen.

Schließlich gilt mein letzter und wichtigster Dank meiner Frau Anne, ohne deren Geduld, Unterstützung und Liebe diese Arbeit nicht möglich gewesen wäre.

Contents

List of Figures

Chapter 1

Introduction

Perhaps the most well-known result of the economic research on group performance is that the free-rider problem arises when a common output is fully shared by all group members, while the cost of contribution is solely borne by each contributor himself. Indeed, the free-rider problem usually occurs in the team production environment and can also be observed in the private provision of public goods by groups (see. e.g. Olson (1965), Hardin (1968), Alchian and Demsetz (1972)). In particular, economic theory predicts that the free-rider problem will reduce individual contributions, leading to inefficient production outcome and lower level of social welfare. While mutual monitoring, peer pressure, and punishment may help to discourage free-riding behavior, incentive-compatible reward schemes are often considered as the most powerful instrument against free-riding and moral hazard (see e.g. Groves (1973), Holmström (1982), Kandel and Lazear (1992), Prendergast (1999)). Unfortunately, the economically efficient behavior is not always contractible. Nevertheless, evidences for altruistic and cooperative behavior in teams and voluntary contribution in public goods are often reported both in the lab and in the field.[1]

In this thesis, the impact of inequity aversion and overconfidence on group performance is studied. The following three chapters theoretically investigate how other-regarding preferences and overly optimistic self-perception affect

[1] An overwiew of the related literature will be presented at the beginning of each chapter.

individuals' contribution to a common output that is equally shared by all group members. The studies presented here concentrate on two main topics: First, we investigate the private provision of public goods when agents are motivated by fairness concerns in terms of inequity aversion. Second, we study the consequences of overly optimistic self-perception for the allocation of tasks and the incentives for cooperation in teams.

All chapters are common regarding three features: First, they all focus on situations in which economically efficient effort choices and contribution levels are not contracted such that incentives for free-riding behavior may exist. Second, both in the public goods and in the team production settings the group outcome is always fully and equally shared by all group members. Finally, both inequity aversion and overconfidence can help to mitigate the negative impact of the free-rider problem leading to more efficient outcomes, even without having to implement the optimal incentive contracts.

In the following, we will briefly present the main research questions and the key results of each chapter. This thesis can be divided into two parts. In the first part (chapter 2), we investigate the effects of inequality in wealth on the incentives to contribute to a public good when agents are inequity averse and differ in their abilities. The research questions we address in this chapter include: How does (un)equal distribution of wealth affect agents' contribution to a public good when they also care for fairness? And what is the optimal wealth distribution that maximizes the social welfare? The results of our formal analysis show that it is worthwhile to introduce ex-ante inequality in wealth when agents of different abilities are inequity averse. The reason is that inequality in favor of a more able agent can motivate this agent to exert higher efforts and helps the group to coordinate on equilibria with higher contributions and less free-riding. In particular, the stronger the agents' inequity aversion, the greater is also this incentive effect of inequality and the larger should be the difference in initial wealth. In contrast, treating heterogenous agents equally may lead to a reduction of public good provision below levels generated by purely selfish agents.

In the second part (chapters 3 and 4), we analyze the effects of biased self-perception in terms of overconfidence on team performance. In chapter

3, we study the consequences of manager overconfidence for organizational performance in a setting in which the manager chooses the allocation of two tasks with different impact. The research questions we study include: What are the effects of manager overconfidence on the task allocation in firms? Is managers' biased self-perception generally harmful or beneficial for firm performance? And can manager overconfidence persist in an organization? In this regard, we show that overconfident managers may exhibit responsibility hoarding behavior and carry out the critical task that has greater impact on firm outcome more often themselves than fully rational managers would. Hence, manager overconfidence may counteract shirking, causing managers to take up more responsibility and reducing inefficient job distributions. While responsibility hoarding is individually suboptimal for the overconfident manager, it can raise the firm output and the total welfare of all involved parties compared to the case with a fully rational manager, when the overconfident manager's true ability is sufficiently high relative to his self-perception bias. Hence, our results imply that firms will not generally avoid overconfident and responsibility hoarding managers, but may even prefer them to fully rational ones. Moreover, an overconfident manager's biased self-perception and his responsibility hoarding behavior can persist in an organization, as long as the manager can rationalize his biased overestimation of the own ability by underestimating the ability of the agent.

Chapter 4 investigates the effect of agent overconfidence on the incentives for helping and cooperation in teams. The research questions we focus on include: Are managers more likely to help fully rational or overconfident agents? And how does agents' biased self-perception affect the team performance? In a setting with complementary production technology, we show that overconfident agents generally tend to exert higher effort to the team production, even though it is individually suboptimal as they suffer from higher effort costs. However, managers may provide more helping to overconfident agents than to fully rational ones due to the synergy of efforts. Interestingly, both individual utility and total welfare of all involved parties can be higher when agents are moderately overconfident. However, the positive effect of agent overconfidence crucially depends on the manager's awareness of the agent's biased self-perception. In particular, our results imply that firms should not generally avoid hiring overconfident agents. But, to exploit the benefits from the employment of overconfident agents, well-founded knowledge on the agents' self-perception bias is necessary. Finally, agent overconfidence and higher level of cooperation among the team members can be sustained in an organization, when the stochastic feedback structure obscures feedback.

We now discuss the content of the following chapters in more detail. The first part of chapter 2 presents the formal model and the results of the equilibrium analysis. We first introduce a simple setting with two agents who are inequity averse as formalized in Fehr and Schmidt (1999). Both agents may differ in their abilities and decide simultaneously on their contributions to a public good which is increasing in each agent's contribution. As both agents may have different abilities, the marginal effect of their contributions may be different. For simplicity, we assume that both agents benefit to the same extent from the public good. As the more able agent provides higher inputs and, in turn, bears higher costs, equality in initial wealth may then lead to inequity.[2] Our equilibrium analysis shows that there are two possible types of equilibrium due to the form of the agents' utility functions. While one agent attains a higher utility than the other one in an *inequitable equilibrium*, both agents are equally well off in an *equitable equilibrium*. In particular, the initial wealth differential is crucial to determine how inequity averse agents choose their effort, and thus, which type of equilibrium will eventually be established, i.e. for high levels of inequality in initial wealth there is a unique inequitable equilibrium, and for intermediate values of initial wealth differential equitable equilibria exist. In this regard, we also show that there are always multiple equitable equilibria. The reason is that inequity averse agents have some interest to adapt their own effort according to their group member's effort in order to avoid the disutility from inequity. This leads to a coordination problem as the reaction functions are upward sloping.

[2] In the following, we use the term inequality describing inequality in initial wealth, and the term inequity describing inequality in wealth after agents have contributed and received their benefits from the public good.

Moreover, the set of equitable equilibria is the larger, the higher the agents' degree of inequity aversion: The more the agents care for equity, the larger is their willingness to adapt their efforts to reduce inequity which may either be triggered by inequality in initial wealth or the group member's effort level. Likewise, the stronger the agents' aversion against inequity, the larger may be the maximal initial wealth inequality the agents are willing to offset by adapting their contributions ending up in an equitable equilibrium.

Based on the results of the first part of this chapter, we compare the contribution level of inequity averse agents and that of purely selfish agents and then examine possible effects of redistribution policy. In this context, our results indicate that for intermediate levels of initial wealth inequality, inequity aversion indeed helps to reduce the free-rider problem as both agents exert higher efforts relative to the levels chosen by purely selfish agents maximizing their material payoffs. However, for larger initial wealth differential, inequity aversion leads to an asymmetric reaction as the favored agent chooses a higher effort than the level maximizing her material payoff and the disadvantaged contributes less than would be optimal from a payoff maximizing perspective. Regarding the effects of redistribution policy, we first consider the agents' individual preferences for redistribution of a given amount of total initial wealth. We show that a less able agent can be better off ex-post by transferring parts of her initial wealth to her more able colleague when both agents are inequity averse. Second, from the perspective of a social planner who is either egalitarian or utilitarian it can be optimal to give the agents different initial wealth exactly because agents are inequity averse, where the more productive agent is provided with a higher amount of initial wealth. Most strikingly, the stronger the agents' inequity aversion, the larger should be the difference in initial wealth. The basic mechanism is the following: If agents are purely selfish and maximize their own payoff, more productive agents will provide more effort as their marginal return is higher. As all agents receive the same share of the public good, the more productive agent gets a smaller total payoff than the less able if both agents are endowed with the same initial wealth. In particular, the less able agent receives a higher payoff than the more able by free-riding. Inequity averse agents with high ability dislike this, and consequently, reduce their effort, although it is optimal for them to exert higher effort than the less able. However, if agents receive different initial wealth they might choose high level of contributions to the public good. The reason is that under certain conditions inequity averse agents coordinate on an equitable equilibrium with high contribution when the distribution of initial wealth is aligned to the difference in their abilities. In these equilibria, both agents have an incentive to match their colleagues' contribution and, in turn, the free-rider problem can be substantially reduced when the agents coordinate on the Pareto-dominant equilibrium. Therefore, a less able agent may benefit when her high ability colleague's income is increased because, in turn, this colleague is willing to contribute more. Finally, we demonstrate that under the optimal distribution of wealth, total contributions are independent of the group composition, i.e. homogeneous and heterogeneous groups provide the same amount of public good and attain the identical levels of social welfare.

In chapter 3, the effects of manager overconfidence on task allocation and firm performance are being investigated. In the first part of this chapter, we introduce a joint production setting, in which a manager and an agent can exert effort into a team production that consists of two distinct tasks with unequal impact on the total output. Either player can be assigned to perform either task, but both tasks must be allocated and each must be assigned to a different player. The allocation of the tasks is determined by the manager, who decides whether to perform the *critical task* (i.e. the task with the higher impact on the output) himself or to delegate it to the agent. Furthermore, both manager and agent are risk neutral and receive the same share of the total output. The outcome of each task is endogenous and depends on the true ability and the effort choice of the task owner. However, the effort levels as well as the output of each task cannot be observed, and only the total output is observable for all parties. In the first step, we derive the efficient task allocation in the first-best case and the utility-maximizing one chosen by a fully rational manager. The results of our analysis show that the fully rational manager assigns the critical task more often to the agent than is efficient. The reason is that in equilibrium the critical task requires higher level of effort due to its higher impact on total output, and thus, causes higher effort costs than the other non-critical task. As long as the agent's ability is not too low, delegating the critical task reduces the manager's effort costs more than it reduces the expected outcome. Hence, the manager may benefit from delegation. Consequently, this may lead to inefficient job distributions and lower total welfare in equilibrium. In the second step, we derive the task allocation in equilibrium chosen by an *overconfident* manager who has an overly optimistic perception of his own ability. We show that an overconfident manager tends to *hoard responsibility*, i.e. to carry out the critical task more often himself than a fully rational manager would. The intuition is that the overconfident manager overestimates his own ability, and consequently, expects a higher outcome when doing the critical task himself. In particular, this kind of responsibility hoarding behavior is more likely to be observed the larger the overconfident manager's self-perception bias is. Although responsibility hoarding is individually suboptimal for the overconfident manager as he suffers from higher costs of effort by performing the critical task, it increases the total welfare of the involved parties, as long as the overconfident manager's self-perception bias is not too large. Moreover, the task allocations chosen by overconfident managers may be closer to the efficient allocation than those of fully rational managers. Therefore, overconfidence can be considered as a commitment device for managers to take up more responsibilities and increase the efficiency of the job distribution, positively affecting the total welfare.

In the remainder of chapter 3, the persistence of manager overconfidence is being analyzed and discussed. Under the assumption that only the total output is observable we show that manager overconfidence and responsibility hoarding behavior can persist in an organization, as long as the manager can rationalize his biased overestimation of the own ability by underestimating the ability of the agent. In this regard, our model provides three interesting implications. First, manager overconfidence is less robust, if the manager carries out the critical task himself. Because of the greater impact of the critical task it is easier for the overconfident manager to rationalize his overly optimistic self-perception by underestimating the agent's contribution to the total outcome when the critical task is carried out by the agent. Second, manager overconfidence has better chance to survive if the tasks are relatively similar with respect to their impact on total outcome. The reason is that the negative output effect of inefficient task allocation causing by the overconfident manager's responsibility hoarding behavior is lower the more similar the tasks are. Finally, both the manager's and the agent's abilities must be relatively high to enable a persistent responsibility hoarding. Intuitively, the more productive the agent is, the more room there is for underestimation. When the agent's ability is sufficiently high, responsibility hoarding only occurs if the manager's ability is also relatively close to the level of the agent. Hence, persistent manager overconfidence and responsibility hoarding are more likely to be observed in firms with high-ability workers.

While the third chapter deals with the effects of manager overconfidence on task allocation and team performance, chapter 4 investigates the impact of agent overconfidence on manager's incentives for helping when there is a complementarity between the agent's productive effort and the manager's helping effort in the production technology. In the first part of this chapter, we introduce a model with a manager and an agent who can both exert effort into a team production. Both individuals are risk neutral and benefit to the same extent from the team output. In addition to the productive effort for her own task, the manager can provide helping to the agent. Furthermore, all effort choices are chosen simultaneously and only the total output is observable for all involved parties. Similar to the previous chapter, we characterize an *overconfident* agent as someone who systematically overestimates his own ability. We start our analysis by deriving the first-best effort choices and comparing them with the optimal effort choices chosen by a manager and a fully rational agent. Our results demonstrate the typical free-riding behavior in a team production setting that fully rational individuals generally exert lower productive and helping efforts than is efficient. In a next step, we derive the effort choices in equilibrium with a fully rational manager and an overconfident agent where we differentiate between two possible cases, i.e. whether the manager is aware of the agent's biased self-perception or not. Regardless of the information setting, overconfident agents generally tend to exert higher effort to the team production than fully rational ones, even though it is individually suboptimal as they suffer from higher effort costs. When managers are aware of the agents' biased self-perception, they will extend their helping effort, leading to higher level of cooperation and better team outcome. Consequently, both productive and helping efforts may be closer to the first-best level than with fully rational agents.

In the second part of this chapter, we analyze the effects of the agent's biased self-perception on the individuals' utilities and the total welfare. Although overconfidence causes individually suboptimal effort choices of the agent, it is not always harmful with respect to his utility. In particular, agent overconfidence can be worthwhile for the total welfare of all involved parties if the agent's self-perception bias is on a moderate level relative to his true ability. However, the manager's awareness of the agent's self-perception bias is crucial for the positive welfare effect of overconfidence. Like in the previous chapter, we also consider the persistence of agent overconfidence when only the total team output is observable. Following the same mechanism described above, agent overconfidence can be sustained as long as the contribution of the manager can be underestimated. Particularly, agent overconfidence is more robust if the manager has perfect information on the agent's self-perception bias. The intuition is that when the manager anticipates higher productive effort of the overconfident agent, he will adapt the level of helping effort additionally enhancing the total output, providing the overconfident agent more room for rationalizing his biased self-perception.

Chapter 2

Inequality, Inequity Aversion, and the Provision of Public Goods[1]

The doctrine of equality! There is no more poisonous poison anywhere: for it seems to be preached by justice itself, whereas it really is the termination of justice. "Equal to the equal, unequal to the unequal" - that would be the true slogan of justice; and also its corollary: "Never make equal what is unequal." (Friedrich Nietzsche)

2.1 Introduction

There is now a broad number of studies indicating that many people tend to dislike inequity. Formal models of inequity aversion such as those by Fehr and Schmidt (1999) or Bolton and Ockenfels (2000) have been quite successful in explaining patterns of behavior observed in laboratory experiments and in the field.[2] In this chapter, we analyze the effect of ex-ante inequality

[1] This chapter is based upon Kölle et al. (2011).

[2] For experimental evidence see for example Roth and Kagel (1995), Camerer (2003) and Engelmann and Strobel (2004). Using a more general notion of fairness, field evidence is given by e.g. Blinder and Choi (1990), Agell and Lundborg (1995), Campbell and Kamlani

in wealth on the motivation of heterogeneous and inequity averse agents to contribute to a public good. While a straightforward conjecture would be that inequity aversion should lead to the optimality of a more egalitarian wealth distribution, we show that the optimal degree of wealth inequality may actually increase with the importance of inequity aversion in the agents' preferences.

We consider a simple setting in which two agents who are inequity averse simultaneously decide on their contributions to a public good. The joined output is increasing in each agent's contribution but both agents may have different abilities which determine the marginal effect of their contributions. When both agents benefit to the same extent from the public good, equality in initial wealth may then lead to inequity as the more able agent provides higher inputs and, in turn, has higher costs.[3] We show that this inequity is endogenously offset to some degree as the agents adapt their contributions. Treating agents of different abilities equally may then have detrimental effects for the provision of the public good. But allocating a higher wealth to the more able agent may motivate the latter to increase her contribution. When the distribution of initial wealth is aligned to the difference in the agents' abilities, there will be multiple equilibria in which the agents attain the same utility even though their initial wealth differs. In these equilibria both agents have an incentive to match their group members' contribution and, in turn, the free-rider problem can be substantially reduced when the agents coordinate on the Pareto-dominant equilibrium. In particular, for intermediate levels of wealth inequality both agents exert higher efforts relative to the efforts maximizing their material payoffs.

We further analyze the optimal degree of initial inequality for two simple settings. In the first setting, we analyze the agents' individual preferences for redistribution of a given amount of total initial wealth. Here, we show that the less able agent may even benefit from initial wealth inequality to

(1997), Bewley (1999) and Carpenter and Seki (2006). For a summary of the empirical evidence on social preferences see for instance Fehr and Schmidt (2002) and Sobel (2005).

[3]In the following we use the term inequality describing inequality in initial wealth, and the term inequity describing inequality in wealth after agents have contributed and received their benefits from the public good.

her *disadvantage*. The reason is that the increased incentives of the more able agent to contribute to the public good can outweigh the loss in initial wealth. In the second setting, we show that not only a utilitarian but also an egalitarian social planner will choose an unequal wealth distribution favoring the more productive agent. Most strikingly, the stronger the agents' inequity aversion, the *larger* should be the difference in initial wealth. Moreover, we show that an egalitarian wealth distribution can only be optimal when all agents have the same ability. On the contrary, in the case of heterogeneous agents such a policy always leads to a stronger underprovision of the public good causing welfare losses. Finally, we demonstrate that under the optimal distribution of wealth, total contributions are independent of the group composition, i.e. homogeneous and heterogeneous groups provide the same amount of the public good and identical levels of social welfare are attained.

In the existing public good literature, a well established result is that the private provision of a public good is unaffected by any reallocation of income amongst contributing agents. This result has first been shown by Warr (1983) and later been extended by Bergstrom et al. (1986). However, the latter also shows that an income redistribution which increases inequality by transferring wealth from non-contributing individuals to contributing individuals can have positive welfare effects (see also Itaya et al. (1997)). In a similar vein, Andreoni (1990) argues that public good provision can be enhanced by redistributing wealth from less altruistic to more altruistic people. We add to this literature by showing that redistribution can be beneficial even for the case of symmetric preferences and even if the set of contributors is left unchanged. While the reason for inequality in our model stems from the heterogeneity in the agents' characteristics, the agents' fairness concerns appear to be an important factor influencing the optimal degree of inequality.

In recent years, there also has been a couple of (predominantly experimental) studies investigating the effects of wealth heterogeneity on public good provision. However, empirical results from these studies are not clear-cut. While some papers find that inequality leads to lower contributions (e.g. Ostrom et al. (1994), Van Dijk et al. (2002), Cherry et al. (2005) and Anderson et al. (2008)), other studies report a neutral or even positive effect of wealth inequality (e.g. Chan et al. (1999), Buckley and Croson (2006)).[4] One reason for the non-conclusive evidence might be that these studies investigated inequality only in the income dimension. Yet, the claim of our study is that there is an interplay of inequality in the income dimension and heterogeneity in the agents' characteristics that affect "psychological" inequity costs which might hamper the cooperation in social dilemmas.

In this regard, our analysis also contributes to the literature on the interplay of equity and equality in social exchanges (e.g. Homans (1958), Adams (1965), Konow (2000), Cappelen et al. (2007) or Konow et al. (2009)). Psychological equity theory (Adams (1965)) for instance argues that individuals do not strive to receive equal benefits or make equal contributions as long as the ratio between benefits and contributions is similar. Analogously, we show that if agents are sufficiently heterogenous, i.e. if the difference in abilities (and hence their inputs) is large, equity between agents is only feasible when initial wealth levels are unequal suggesting that (in)equality does not necessarily imply (in)equity and vice versa.

Applied to a team production context within firms, our study provides insights on the question whether equal wages are always the best wage policy. While it has often been argued that unequal reward schemes provoke morale problems among co-workers leading to lower performances (e.g. Akerlof and Yellen (1990), Bewley (1999)), some other studies questioned whether equal payment, realized by wage compression, does eliminate all these problems.[5] Winter (2004), for instance, shows that it might be even optimal to reward identical agents differently as coordination can be improved which has recently be confirmed in an experiment by Goerg et al. (2010). In another experiment, Abeler et al. (2010) find that paying equal wages after an unequal performance may lead to inequity and, in turn, to substantially lower efforts and a decline in efficiency over time. But while these papers argue for inequality in ex-post performance rewards, our results show that it may even be optimal to introduce ex-ante inequality in the non-performance contin-

[4] Chan et al. (1996) find evidence which in line with the model of Bergstrom et al. (1986) on an aggregate level but not an individual level.

[5] See e.g. Lazear (1989) who argues that "... it is far from obvious that pay equality has these effects".

gent wage components. Furthermore, our analysis also adds to the literature on behavioral contract theory studying the effects of inequity aversion on incentives.[6] However, while in most of the studies inequity aversion leads to more equal payment structures, our model shows that inequity aversion may be a reason to introduce ex-ante inequality.

The remainder of this chapter is structured as follows. The model is described in section 2.2. Section 2.3 presents the equilibrium analysis. In section 2.4, we compare the effort levels chosen by inequity averse and purely selfish agents. Section 2.5 analyzes preferences for redistribution and examines the effects of distribution policies and group composition on the public good provision and social welfare. Section 2.6 concludes.

2.2 The model

Two agents i and j can both contribute to a public good. An agent's contribution depends on her effort e_i and her ability a_i. Individual effort costs are linear in the exerted effort and equal to $c \cdot e_i$, $c \in \mathbb{R}^+$. The group output is determined by the sum of both agents' contribution:

$$a_i\sqrt{e_i} + a_j\sqrt{e_j}.$$

The agents directly benefit from a higher group output. Each agent receives a share η of the group output indicating her individual valuation of the public good (marginal per capital return). Furthermore, each agent i is provided with an initial endowment w_i.[7] Let $\Delta w_i = w_i - w_j$ be the difference in initial endowments. Both agents are inequity averse with a Fehr and Schmidt (1999)

[6] For a theoretical investigation of this topic see for instance Itoh (2004), Grund and Sliwka (2005), Huck and Rey-Biel (2006), Demougin et al. (2006), Fehr et al. (2007), Rey-Biel (2008), Dur and Glazer (2008), Mohnen et al. (2008), Kragl and Schmid (2009), Neilson and Stowe (2010), Bartling and von Siemens (2010) and Englmaier and Wambach (2010).

[7] In a team context, η represents e.g. the degree of team identification or the intrinsic benefit of the work output and w_i represents the wage.

type utility function. An agent's utility is[8]

$$U_i = w_i - c \cdot e_i + \eta \cdot \left(a_i\sqrt{e_i} + a_j\sqrt{e_j}\right) - v\left(w_i - c \cdot e_i - w_j + c \cdot e_j\right)$$

with

$$v(\Delta) = \begin{cases} -\alpha \cdot \Delta & if \quad \Delta < 0 \\ \beta \cdot \Delta & if \quad \Delta > 0 \end{cases}$$

where α measures the "psychological costs" of disadvantageous inequity and β that of advantageous inequity. Following Fehr and Schmidt (1999) we assume that $\alpha \geq \beta \geq 0$. Additionally, we assume that $\beta \leq \frac{1}{2}$.[9]

2.3 Equilibrium analysis

Each agent i maximizes

$$\max_{e_i} w_i + \eta \cdot \left(a_i\sqrt{e_i} + a_j\sqrt{e_j}\right) - c \cdot e_i - v\left(w_i - c \cdot e_i - w_j + c \cdot e_j\right).$$

The function is continuous but not continuously differentiable as it has a kink at $e_i = \frac{\Delta w_i}{c} + e_j$ where i attains the same utility as j. Off the kink, the second derivative with respect to e_i is $-\frac{\eta a_i \sqrt{e_i}}{4e_i^2} < 0$. As the right-sided derivative at the kink is strictly smaller than the left-sided derivative, the function is strictly concave.

We have to consider two possible equilibrium types depending on whether there is inequity in equilibrium or whether both agents are equally well off. In an *inequitable equilibrium* one agent i is better off given the chosen effort levels, i.e. $w_i - ce_i > w_j - ce_j$. Suppose that such an equilibrium exists. When

[8] Hence, we allow that the disutility from inequity $v(\Delta)$ depends on the difference of the agents' net-wealth (rewards minus costs of effort).

[9] Note that $\beta > \frac{1}{2}$ connotes a very strong form of inequity aversion implying that ex-post, agents would be willing to donate parts of their wealth to less wealthy group members up to the point where wealth levels are completely equalized (compare Rey-Biel (2008)). We discuss implications of this assumption at the end of section 2.3.

agent i is better off, the following two conditions must hold in equilibrium

$$\begin{aligned}\frac{\partial U_i}{\partial e_i} &= -c + \frac{\eta a_i}{2\sqrt{e_i}} + \beta c = 0, \\ \frac{\partial U_j}{\partial e_j} &= -c + \frac{\eta a_j}{2\sqrt{e_j}} - \alpha c = 0.\end{aligned}$$

The respective equilibrium efforts are therefore

$$e_i^* = \frac{\eta^2 a_i^2}{4\left(1-\beta\right)^2 c^2} \text{ and } e_j^* = \frac{\eta^2 a_j^2}{4\left(1+\alpha\right)^2 c^2}. \tag{2.1}$$

Such an equilibrium exists if at these effort levels we indeed have that $w_i - ce_i > w_j - ce_j$ or

$$w_i - c \cdot \left(\frac{\eta^2 a_i^2}{4\left(1-\beta\right)^2 c^2}\right) > w_j - c \cdot \left(\frac{\eta^2 a_j^2}{4\left(1+\alpha\right)^2 c^2}\right).$$

This directly leads to the following result:

Proposition 1 *If the difference in initial wealth* $\Delta w_i \equiv w_i - w_j > \frac{\eta^2}{4c}\left(\frac{a_i^2}{(1-\beta)^2} - \frac{a_j^2}{(1+\alpha)^2}\right)$, *there exists a unique inequitable equilibrium. In this equilibrium, agent i is strictly better off than agent j; the equilibrium effort levels satisfy:*

$$e_i^* = \frac{\eta^2 a_i^2}{4\left(1-\beta\right)^2 c^2} \textit{ and } e_j^* = \frac{\eta^2 a_j^2}{4\left(1+\alpha\right)^2 c^2}.$$

Note that both agents adapt their efforts as the contribution of the favored agent i increases in the degree of "compassion" β and that of her disadvantaged counterpart j decreases in the degree of "envy" α. Still, they here end up a situation which is inequitable ex-post. But as the result shows this is only the case when the initial inequality in wealth is sufficiently large.

We now have to check whether there are also *equitable equilibria* in which both agents attain the same payoff. In that case $w_i - ce_i = w_j - ce_j$ and both agents choose their effort levels at the kink of the respective utility function. An effort tuple $\left(\overline{e}_i^*, \overline{e}_j^*\right)$ can be sustained in such an equitable Nash equilibrium if no agent has an incentive to deviate. As the function is strictly concave, necessary and sufficient conditions for the existence of the equilibrium are that for both agents the left hand side derivative of the utility function must be positive at $\left(\overline{e}_i^*, \overline{e}_j^*\right)$, the right hand side derivative negative and $w_i - c\overline{e}_i^* = w_j - c\overline{e}_j^*$. Hence, in an *equitable equilibrium*, the following five conditions must be met:

$$\left.\frac{\partial_- U_i}{\partial e_i}\right|_{e_i=\overline{e}_i^*} = -c + \frac{\eta a_i}{2\sqrt{\overline{e}_i^*}} + \beta c \geq 0 \Leftrightarrow \overline{e}_i^* \leq \frac{\eta^2 a_i^2}{4(1-\beta)^2 c^2} \tag{2.2}$$

$$\left.\frac{\partial_+ U_i}{\partial e_i}\right|_{e_i=\overline{e}_i^*} = -c + \frac{\eta a_i}{2\sqrt{\overline{e}_i^*}} - \alpha c \leq 0 \Leftrightarrow \overline{e}_i^* \geq \frac{\eta^2 a_i^2}{4(1+\alpha)^2 c^2} \tag{2.3}$$

$$\left.\frac{\partial_- U_j}{\partial e_j}\right|_{e_j=\overline{e}_j^*} = -c + \frac{\eta a_j}{2\sqrt{\overline{e}_j^*}} + \beta c \geq 0 \Leftrightarrow \overline{e}_j^* \leq \frac{\eta^2 a_j^2}{4(1-\beta)^2 c^2} \tag{2.4}$$

$$\left.\frac{\partial_+ U_j}{\partial e_j}\right|_{e_j=\overline{e}_j^*} = -c + \frac{\eta a_j}{2\sqrt{\overline{e}_j^*}} - \alpha c \leq 0 \Leftrightarrow \overline{e}_j^* \geq \frac{\eta^2 a_j^2}{4(1+\alpha)^2 c^2} \tag{2.5}$$

$$\overline{e}_j^* = \overline{e}_i^* - \frac{\Delta w_i}{c} \tag{2.6}$$

From these conditions the following result can be derived:

Proposition 2 *If* $\frac{\eta^2}{4c}\left(\frac{a_i^2}{(1+\alpha)^2} - \frac{a_j^2}{(1-\beta)^2}\right) \leq \Delta w_i \leq \frac{\eta^2}{4c}\left(\frac{a_i^2}{(1-\beta)^2} - \frac{a_j^2}{(1+\alpha)^2}\right)$, *there exists a continuum of equitable equilibria. Specifically, any pair* $\left(\overline{e}_i^*, \overline{e}_j^*\right)$ *of effort levels such that*

$$\max\left\{\frac{\eta^2 a_i^2}{4(1+\alpha)^2 c^2}; \frac{\eta^2 a_j^2}{4(1+\alpha)^2 c^2} + \frac{\Delta w_i}{c}\right\} \leq \overline{e}_i^* \leq \min\left\{\frac{\eta^2 a_i^2}{4(1-\beta)^2 c^2}; \frac{\eta^2 a_j^2}{4(1-\beta)^2 c^2} + \frac{\Delta w_i}{c}\right\} \tag{2.7}$$

and $\overline{e}_j^* = \overline{e}_i^* - \frac{\Delta w_i}{c}$ *is an equitable equilibrium.*

Proof.
Inserting the equity condition (2.6) in conditions (2.4) and (2.5), we can conclude that an effort level $\overline{e}_i^*$ can be sustained if and only if

$$\max\left\{\frac{\eta^2 a_i^2}{4(1+\alpha)^2 c^2}; \frac{\eta^2 a_j^2}{4(1+\alpha)^2 c^2} + \frac{\Delta w_i}{c}\right\} \leq \overline{e}_i^* \leq \min\left\{\frac{\eta^2 a_i^2}{4(1-\beta)^2 c^2}; \frac{\eta^2 a_j^2}{4(1-\beta)^2 c^2} + \frac{\Delta w_i}{c}\right\}.$$

Note that $\frac{\eta^2 a_i^2}{4(1+\alpha)^2 c^2} < \frac{\eta^2 a_i^2}{4(1-\beta)^2 c^2}$ and $\frac{\eta^2 a_j^2}{4(1+\alpha)^2 c^2} + \frac{\Delta w_i}{c} < \frac{\eta^2 a_j^2}{4(1-\beta)^2 c^2} + \frac{\Delta w_i}{c}$. Hence, the set is non-empty for certain values of Δw_i if

$$\frac{\eta^2 a_i^2}{4(1-\beta)^2 c^2} \geq \frac{\eta^2 a_j^2}{4(1+\alpha)^2 c^2} + \frac{\Delta w_i}{c}$$

and

$$\frac{\eta^2 a_j^2}{4(1-\beta)^2 c^2} + \frac{\Delta w_i}{c} \geq \frac{\eta^2 a_i^2}{4(1+\alpha)^2 c^2}$$

which is the case when

$$\frac{\eta^2 a_i^2}{4(1+\alpha)^2 c} - \frac{\eta^2 a_j^2}{4(1-\beta)^2 c} \leq \Delta w_i \leq \frac{\eta^2 a_i^2}{4(1-\beta)^2 c} - \frac{\eta^2 a_j^2}{4(1+\alpha)^2 c}. \blacksquare \tag{2.8}$$

This result has several interesting implications. First, note that there are always multiple equitable equilibria. The reason is that inequity averse agents have some interest to adapt their own effort according to the group member's effort in order to avoid the disutility from inequity. This leads to a coordination problem as the reaction functions are upward sloping.

Second, the set of equitable equilibria defined by (2.7) is the larger, the higher the agents' degree of inequity aversion: The more the agents care for equity, the larger is their willingness to adapt their efforts to reduce inequity which may either be triggered by inequality in initial wealth or the group member's effort level. The lower boundary of the equilibrium set is decreasing in α as more "envious" agents are willing to reduce their efforts to avoid being worse off than their group member. Analogously, the upper boundary is increasing in β as more "compassionate" agents are more willing to raise their efforts to reduce a group member's disadvantage. Likewise, the set defined by (2.8) is also increasing in the agents' inequity aversion implying that the stronger the agents' aversion against inequity, the larger may be the maximal initial wealth inequality the agents are willing to offset by adapting their contributions ending up in an equitable equilibrium.

Finally, note that the lower boundary for Δw_i as defined by condition (2.8) exceeds zero (or the upper boundary is smaller than zero) when the abilities differ strongly and inequity aversion is not too strong. In these cases, equitable equilibria never exist when $\Delta w_i = 0$ and, hence, equity cannot be attained when wealth is distributed equally. The reason is that due to the higher marginal productivity of effort, the more productive agent will have a higher incentive to exert more effort than her less productive fellow agent and, in turn, bears higher costs. But as both agents benefit equally from the public good the more able agent is worse off when both have the same initial wealth.[10]

Figure 2.1 shows the sustainable equilibrium effort levels of both agents i and j as a correspondence of Δw_i.[11] There are two cut-off values for Δw_i. For small values of Δw_i $(= -\Delta w_j)$ below $\frac{\eta^2}{4c}\left(\frac{a_i^2}{(1+\alpha)^2} - \frac{a_j^2}{(1-\beta)^2}\right)$ there is a unique inequitable equilibrium with $e_i^* = \frac{\eta^2 a_i^2}{4(1+\alpha)^2 c^2}$ and $e_j^* = \frac{\eta^2 a_j^2}{4(1-\beta)^2 c^2}$. For large values of Δw_i above $\frac{\eta^2}{4c}\left(\frac{a_i^2}{(1-\beta)^2} - \frac{a_j^2}{(1+\alpha)^2}\right)$ there is a unique inequitable equilibrium with $e_i^* = \frac{\eta^2 a_i^2}{4(1-\beta)^2 c^2}$ and $e_j^* = \frac{\eta^2 a_j^2}{4(1+\alpha)^2 c^2}$. For intermediate values of Δw_i equitable equilibria exist.

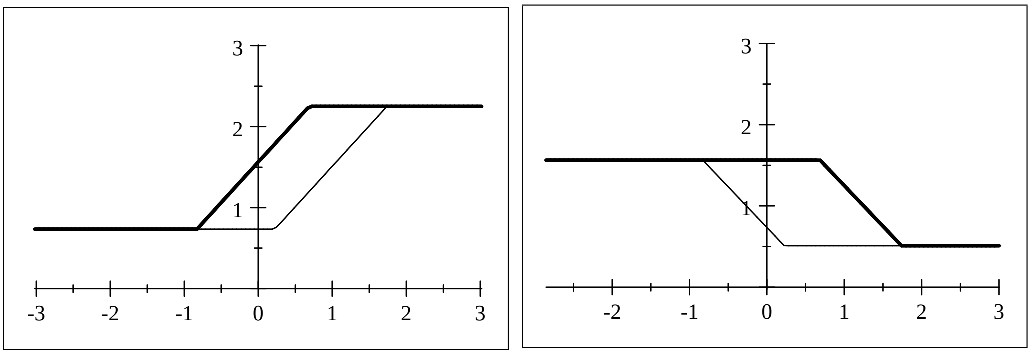

Figure 2.1: Effort Choice of agent i (left) and j (right) depending on Δw_i

Note that as both agents attain identical payoffs in an equitable equilibrium, they prefer the same one. Consequently, it is important to compare the different feasible equitable equilibria with respect to the agents' utility which leads to the following result:

[10] Note that this is always the case when the agents are purely selfish (i.e. $\alpha = \beta = 0$).

[11] The figure shows a setting in which $a_i = 12$, $a_j = 10$, $\alpha = 0.4$, $\beta = 0.2$, $\eta = 0.2$, and $c = 1$.

Corollary 1 *As long as $\beta \leq \frac{1}{2}$ the equitable equilibrium in which the agents' utility is highest is always Pareto optimal within the set of Nash equilibria.*

Proof. See the appendix.

To understand this result note that there is a free-rider problem which is particularly strong when agents are selfish. Inequity aversion helps to overcome this free-rider problem as it allows agents to coordinate on higher effort levels which come closer to the first best. As long as β does not exceed $\frac{1}{2}$ the highest feasible equilibrium is still lower than the first-best and therefore is preferred by the agents.[12] With a β larger than $\frac{1}{2}$, however, inequity aversion becomes so strong that an agent even would have an incentive to match an inefficiently high effort level chosen by her group member even though both would be better off with a lower effort.

Hence, both agents benefit from playing the equitable equilibrium with the highest sustainable effort level when they are not extremely "compassionate". This effort level is equal to $\min\left\{\frac{\eta^2 a_i^2}{4(1-\beta)^2 c^2}; \frac{\eta^2 a_j^2}{4(1-\beta)^2 c^2} + \frac{\Delta w_i}{c}\right\}$ and, hence, strictly increasing in the degree of advantageous inequity aversion β.

2.4 Do inequity averse agents contribute more?

We now compare the attained effort levels with those chosen by purely selfish agents to study the effects of inequity aversion on the motivation to contribute to the public good. From Propositions 1 and 2 as well as Corollary 1 (assuming that the agents play the Pareto best equitable equilibrium)[13] we know that the equilibrium effort levels of inequity averse agents $\left(e_i^*, e_j^*\right)$ are

[12] The agents' first-best efforts can be derived by maximizing $w_i + w_j - c \cdot e_i - c \cdot e_j + 2\eta \cdot \left(a_i\sqrt{e_i} + a_j\sqrt{e_j}\right)$ and are given by $e_i^{FB} = \frac{\eta^2 a_i^2}{c^2}$ and $e_j^{FB} = \frac{\eta^2 a_j^2}{c^2}$.

[13] Cooper et al. (1992), Blume and Ortmann (2007) for instance find experimentally that simple ex-ante cheap talk communication indeed very frequently leads to the choice of the pareto efficient Nash equilibrium in coordination games. See Demichelis and Weibull (2008) for a theoretical argument based on lexicographic preferences for honesty.

given by

$$\begin{cases} \left(\frac{\eta^2 a_i^2}{4(1+\alpha)^2 c^2}, \frac{\eta^2 a_j^2}{4(1-\beta)^2 c^2}\right) & if\ \Delta w_i < \frac{\eta^2}{4c}\left(\frac{a_i^2}{(1+\alpha)^2} - \frac{a_j^2}{(1-\beta)^2}\right) \\ \left(\frac{\eta^2 a_j^2}{4(1-\beta)^2 c^2} + \frac{\Delta w_i}{c}, \frac{\eta^2 a_j^2}{4(1-\beta)^2 c^2}\right) & if\ \Delta w_i \in \left[\frac{\eta^2}{4c}\left(\frac{a_i^2}{(1+\alpha)^2} - \frac{a_j^2}{(1-\beta)^2}\right), \frac{\eta^2}{4c}\left(\frac{a_i^2 - a_j^2}{(1-\beta)^2}\right)\right] \\ \left(\frac{\eta^2 a_i^2}{4(1-\beta)^2 c^2}, \frac{\eta^2 a_i^2}{4(1-\beta)^2 c^2} - \frac{\Delta w_i}{c}\right) & if\ \Delta w_i \in \left[\frac{\eta^2}{4c}\left(\frac{a_i^2 - a_j^2}{(1-\beta)^2}\right), \frac{\eta^2}{4c}\left(\frac{a_i^2}{(1-\beta)^2} - \frac{a_j^2}{(1+\alpha)^2}\right)\right] \\ \left(\frac{\eta^2 a_i^2}{4(1-\beta)^2 c^2}, \frac{\eta^2 a_j^2}{4(1+\alpha)^2 c^2}\right) & if\ \Delta w_i > \frac{\eta^2}{4c}\left(\frac{a_i^2}{(1-\beta)^2} - \frac{a_j^2}{(1+\alpha)^2}\right) \end{cases} \tag{2.9}$$

as depicted by the solid upper boundary of the graphs in Figure 2.1. Note that both functions are continuous and weakly monotonic.

Suppose, w.l.o.g., that i is the more able agent i.e. $a_i \geq a_j$. Purely selfish agents' effort choices are not affected by initial wealth inequality as they consider only their marginal returns when choosing their efforts. Hence, efforts are given by[14]

$$e_i^{selfish} = \frac{\eta^2 a_i^2}{4c^2} \text{ and } e_j^{selfish} = \frac{\eta^2 a_j^2}{4c^2}. \tag{2.10}$$

By comparing these effort levels of selfish agents with those of inequity averse agents as given by (2.9) we obtain the following result:

Proposition 3 *If* $\frac{\eta^2 a_i^2}{4c} - \frac{\eta^2 a_j^2}{4(1-\beta)^2 c} < \Delta w_i < \frac{\eta^2 a_i^2}{4(1-\beta)^2 c} - \frac{\eta^2 a_j^2}{4c}$*, both agents contribute more when they are inequity averse (i.e.* $e_i^* > e_i^{selfish}$ *and* $e_j^* > e_j^{selfish}$*). If* $\Delta w_i \geq \frac{\eta^2 a_i^2}{4(1-\beta)^2 c} - \frac{\eta^2 a_j^2}{4c}$*, inequity aversion motivates agent i to exert higher efforts but de-motivates agent j (i.e.* $e_i^* > e_i^{selfish}$ *and* $e_j^* < e_j^{selfish}$*). The opposite holds if* $\Delta w_i \leq \frac{\eta^2 a_i^2}{4c} - \frac{\eta^2 a_j^2}{4(1-\beta)^2 c}$.

Proof. By comparing (2.9) with (2.10) it is straightforward to see that $e_i^* > e_i^{selfish}$ if $\Delta w_i \geq \frac{\eta^2}{4c}\left(\frac{a_i^2 - a_j^2}{(1-\beta)^2}\right)$ and $e_i^* < e_i^{selfish}$ if $\Delta w_i \leq \frac{\eta^2}{4c}\left(\frac{a_i^2}{(1+\alpha)^2} - \frac{a_j^2}{(1-\beta)^2}\right)$. We only have to check the case in which $\Delta w_i \in \left(\frac{\eta^2}{4c}\left(\frac{a_i^2}{(1+\alpha)^2} - \frac{a_j^2}{(1-\beta)^2}\right), \frac{\eta^2}{4c}\left(\frac{a_i^2 - a_j^2}{(1-\beta)^2}\right)\right)$. In this case $e_i^* > e_i^{selfish}$ if

$$\frac{\eta^2 a_j^2}{4(1-\beta)^2 c^2} + \frac{\Delta w_i}{c} > \frac{\eta^2 a_i^2}{4c^2} \Leftrightarrow \Delta w_i > \frac{\eta^2 a_i^2}{4c} - \frac{\eta^2 a_j^2}{4(1-\beta)^2 c}.$$

[14] To see that, just replace $\alpha = \beta = 0$ in the equilibrium efforts given by (2.1).

Hence, we can conclude that $e_i^* > e_i^{selfish}$ if Δw_i exceeds this cut-off.[15] Analogously, $e_j^* > e_j^{selfish}$ if $\Delta w_j = -\Delta w_i > \frac{\eta^2 a_j^2}{4c} - \frac{\eta^2 a_i^2}{4(1-\beta)^2 c}$ which gives us the upper boundary. It is straightforward to check that the interval in which both $e_i^* > e_i^{selfish}$ and $e_j^* > e_j^{selfish}$ is non-empty. ■

Hence, the initial wealth differential Δw_i is crucial to determine how inequity averse agents adapt their effort choices relative to the efforts maximizing their material payoffs. For intermediate levels of initial wealth inequality, inequity aversion indeed helps to reduce the free-rider problem as both agents contribute more when coordinating on the Pareto-superior equilibrium.

But if initial wealth inequality becomes stronger, inequity aversion leads to an asymmetric reaction as the favored agent chooses a higher effort than the level maximizing her material payoff and the disadvantaged contributes less than would be optimal from a payoff maximizing perspective.

But it is important to note that the latter demotivating effect may arise for the more able agent even when she is richer than her less able colleague: The lower boundary for Δw_i in Proposition 3 is larger than zero if

$$\tfrac{\eta^2 a_i^2}{4c} - \tfrac{\eta^2 a_j^2}{4(1-\beta)^2 c} > 0 \Leftrightarrow a_i > \frac{a_j}{1-\beta}.$$

Hence, when a_i is much larger than a_j or when β is sufficiently small, the more able agent reduces her effort below $e_i^{selfish}$ unless Δw_i exceeds a *strictly positive* cut-off value. Or, in other words, she has to be paid sufficiently more than her colleague or otherwise will reduce her effort below the selfishly optimal level. To understand the reason for this effect, note again that the payoff maximizing effort is always larger for the more able agent as her marginal returns to effort are higher. As both equally benefit from the public good, she is worse off than her less able colleague when both have the same initial wealth. But when being inequity averse she suffers from this disadvantage which is the higher the larger a_i relative to a_j. If β is high, the more able agent will still choose an equilibrium effort level above $e_i^{selfish}$ as also her less able but "compassionate" counterpart puts in a sufficiently high effort and they can coordinate to a superior equilibrium. But when β is small, she can

[15] Note that this cut-off is indeed always in the interior of the relevant interval.

only reduce inequity by lowering her effort. Hence, not awarding the more able agent more money up front leads to an unfair distribution of payoffs and, in turn, to lower efforts.

2.5 Social welfare, redistribution, and group composition

We proceed by analyzing redistribution preferences of a) the agents and b) a social planner who can allocate a fixed budget. We further investigate the welfare consequences of a policy implementing an egalitarian wealth distribution irrespective of the distribution of the agents' abilities. Finally, we examine the effect of group composition under the optimal distribution of the initial wealth.

2.5.1 Individual preferences for redistribution

We first study the agents' ex-ante preferences on the initial wealth differential Δw_i when they take into account their equilibrium effort choices. These considerations will be a useful starting point for welfare analysis. To do that, it is instructive to consider a situation in which a certain budget $W = w_i + w_j$ can be distributed between the two agents. By inserting the equilibrium effort choices (2.9) into the agents' utility functions we can describe their utility as a function of the initial wealth differential Δw_i. Analyzing the shape of the indirect utility functions we obtain the following result:

Proposition 4 *The agents' utility function is continuous in* Δw_i. *If* $\Delta w_i < \frac{\eta^2}{4c}\left(\frac{a_i^2 - a_j^2}{(1-\beta)^2}\right)$ *or* $\Delta w_i > \frac{\eta^2}{4c}\left(\frac{a_i^2}{(1-\beta)^2} - \frac{a_j^2}{(1+\alpha)^2}\right)$ *an agent i's utility is strictly increasing in* Δw_i. *But between these two cut-off values it is strictly decreasing. Both agents' utility functions attain a local maximum at* $\Delta w_i^* = \frac{\eta^2}{4c}\left(\frac{a_i^2 - a_j^2}{(1-\beta)^2}\right)$.

Proof. See the appendix.

This result is illustrated in Figure 2.2. The solid line shows agent i's utility and the dashed line agent j's utility both as a function of Δw_i.[16] For extreme values of Δw_i each agent benefits from a redistribution in her favor and there is a straightforward conflict of interest between both agents. But in the interval between $\frac{\eta^2}{4c}\left(\frac{a_j^2}{(1-\beta)^2}-\frac{a_i^2}{(1+\alpha)^2}\right)$ and $\frac{\eta^2}{4c}\left(\frac{a_i^2}{(1-\beta)^2}-\frac{a_j^2}{(1+\alpha)^2}\right)$ both agents' interests are fully aligned. The reason is that within this interval only equitable equilibria exist, and hence, any ex-ante inequality in wealth will be offset by adapted effort levels. Moreover, all values of Δw_i within this interval are Pareto-dominated by a initial wealth differential of $\Delta w_i^* = \frac{\eta^2}{4c}\left(\frac{a_i^2-a_j^2}{(1-\beta)^2}\right)$ as at this point, agents can coordinate on an equilibrium leading to the highest contributions.

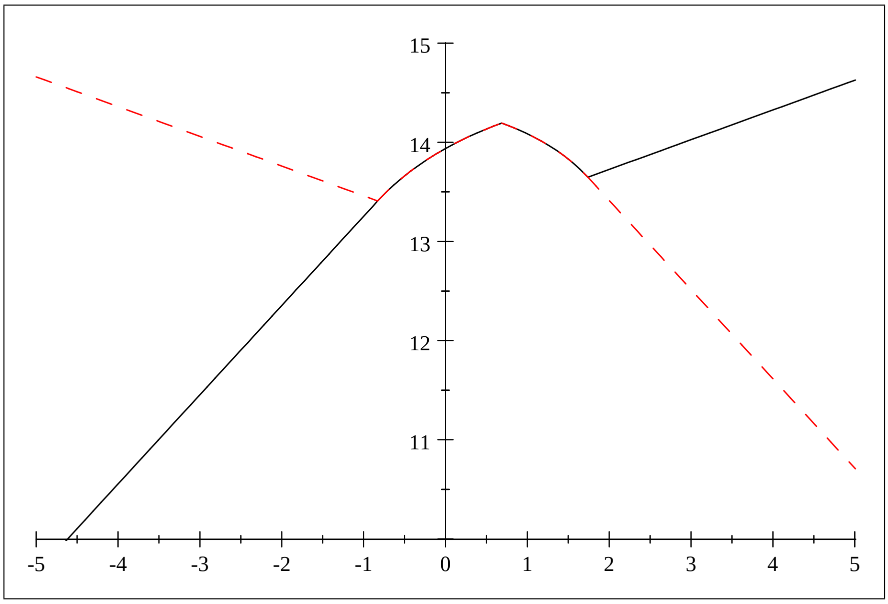

Figure 2.2: Agent i's and agent j's utilities in equilibrium depending on Δw_i

Proposition 4 has several interesting implications. Consider the situation of an individual agent who can (re-)distribute a given wealth allocation. Interestingly, an individual may benefit from ex-ante redistribution at her own expense as the following result shows:

[16] The figure shows a setting in which $a_i = 12$, $a_j = 10$, $\alpha = 0.4$, $\beta = 0.2$, $\eta = 0.2$, and $c = 1$.

Corollary 2 *If both agents receive the same initial wealth (i.e.* $\Delta w_i = 0$*) the less able agent* j *can be made better off by reducing her own initial wealth by* $\frac{\eta^2}{8c}\left(\frac{a_i^2-a_j^2}{(1-\beta)^2}\right)$ *and transferring this money to the more able colleague* i *if*

$$\alpha > \tfrac{1-\beta}{3-4\beta}\sqrt{\left(2\left(6\beta^2-7\beta+2\right)\right)}+\tfrac{2(1-\beta)^2}{(3-4\beta)}-1.$$

If α *is smaller than this cut-off, such a transfer is still beneficial for agent* j *when her ability is not to small, i.e. if*

$$\frac{a_j}{a_i} > \min\left\{\tfrac{1-\beta}{1+\alpha},\sqrt{\tfrac{1}{(1-2\beta)}\left(\tfrac{4(1-\beta)^2}{(1+\alpha)}+1-4\left(1-\beta\right)-\tfrac{2\beta(1-\beta)^2}{(1+\alpha)^2}\right)}\right\}.$$

Proof. See the appendix.

Hence, a less able agent can be better off ex-post when sacrificing parts of her initial wealth which are then transferred to a more able individual. She then benefits from this colleague's higher willingness to contribute to the public good and this helps to reduce the free-rider problem. Interestingly, this is always the case irrespective of the difference in abilities if α is sufficiently large. Moreover, note that the cut-off value for α is equal to $\frac{1}{3}$ when $\beta = 0$ and strictly decreasing in β. Hence, this condition holds for moderate values of α even when the agents only suffer from disadvantageous inequity. The reason is that a more able agent resents being worse off than her less able colleague when exerting a higher effort due to her higher productivity. But she is willing to exert higher efforts when she earns more. Therefore, a less able agent may benefit when her colleague's income is increased because, in turn, this colleague is willing to contribute more.

If α is rather small, the result still holds if the less able agent's productivity is not too small relative to her more able colleague's productivity. If, however, her ability is much smaller the transfer necessary to implement a performance maximizing equitable equilibrium is too large such that agent j prefers to stick with the case in which both receive the same initial wealth although this leads to a lower group output.

2.5.2 Social welfare

We now study a situation in which an external authority can decide on the distribution of wealth. To do so, we consider, a social planner who has a social welfare function which is either egalitarian (i.e. who wants to maximize the utility of the least well-off) or utilitarian (i.e. wants to maximize the sum of both agents' utility). It directly follows from Proposition 4 that such a social planner always has a dominant choice:

Corollary 3 *A social planner who is either utilitarian or egalitarian will set* $\Delta w_i^* = \frac{\eta^2}{4c}\left(\frac{a_i^2-a_j^2}{(1-\beta)^2}\right)$.

Proof. It is straightforward to see that within the set of initial wealth differentials inducing equitable equilibria both egalitarian and utilitarian planners will always choose $\Delta w_i^* = \frac{\eta^2}{4c}\left(\frac{a_i^2-a_j^2}{(1-\beta)^2}\right)$ as at this spread, both the sum and the minimum of the agents' utilities are maximized. Moreover, for an egalitarian social planner any wealth differential which is not inducing an equitable equilibrium is always dominated by this choice as the utility of the least well off agent is always lower in an inequitable equilibrium. To see this formally note that when, w.l.o.g., j is favored $\frac{\partial u_i}{\partial \Delta w_i} = \frac{1}{2} + \alpha > 0$ for all Δw_i inducing an inequitable equilibrium (see (2.12) in the proof of Proposition 4) and as the utility function is continuous, i's utility is always larger in an equitable equilibrium.

A utilitarian social planner will neither choose a wealth distribution inducing an inequitable equilibrium, as in an inequitable equilibrium which, w.l.o.g., favors agent j, the marginal gain from transferring money to agent i $\frac{\partial u_i}{\partial \Delta w_i} = \frac{1}{2} + \alpha$ is always larger than j's marginal loss which is equal to $\frac{1}{2} - \beta$ (see again (2.12)). ■

Hence, even an egalitarian social planner who only considers the utility of the least well off individual should allow for inequality in initial wealth. The reason is that it is precisely this inequality in initial wealth induces an equilibrium in which equity is attained ex-post and in which the more able agent is willing to contribute more. This observation bears some resemblance to the result by Andreoni (1990) who argues that redistribution of income will increase the total contribution if it benefits the more altruistic individuals.[17] It directly follows that the implementation of an egalitarian wealth distribution policy has detrimental effects if the group considered is not entirely homogenous in terms of abilities.

2.5.3 Optimal group composition

So far, we only considered how wealth should be distributed treating the composition of agents within a group as exogenously given. However, it is also interesting to study the case in which the formation of groups can be determined as well. A straightforward conjecture is that group composition matters for the willingness to contribute if the agents are inequity averse towards their fellow group members. To investigate this, we consider a simple situation in which there are four agents, two of high ability and two of low ability, that can be assigned into two groups of two. By comparing total contributions, we can derive the following result:

Proposition 5 *If all agents have the same initial wealth, total contributions are always higher with homogenous than with heterogenous groups. But when wealth can be adapted optimally, total contributions are independent of the group composition.*

Proof. Let $a_H > a_L$ be the ability of the high and low productive agent and let w_H and w_L denote the initial wealth levels of the two agents, respectively. Given the same initial wealth ($\Delta w = w_H - w_L = 0$) the total contribution with two homogenous groups is equal to

$$\frac{2\eta^2 a_H^2}{4(1-\beta)^2 c^2} + \frac{2\eta^2 a_L^2}{4(1-\beta)^2 c^2} = \frac{\eta^2\left(a_H^2 + a_L^2\right)}{2(1-\beta)^2 c^2}. \tag{2.11}$$

With heterogenous groups it is given by

$$\begin{array}{lll} \frac{\eta^2 a_H^2}{2(1+\alpha)^2 c^2} + \frac{\eta^2 a_L^2}{2(1-\beta)^2 c^2} & if & \Delta w < \frac{\eta^2}{4c}\left(\frac{a_H^2}{(1+\alpha)^2} - \frac{a_L^2}{(1-\beta)^2}\right) \\ \frac{\eta^2\left(a_L^2 + a_L^2\right)}{2(1-\beta)^2 c^2} & if & \Delta w \in \left[\frac{\eta^2}{4c}\left(\frac{a_H^2}{(1+\alpha)^2} - \frac{a_L^2}{(1-\beta)^2}\right), \frac{\eta^2}{4c}\left(\frac{a_H^2 - a_L^2}{(1-\beta)^2}\right)\right] \end{array} .$$

[17] Similarly, with respect to social welfare, Thurow (1971) argues that some redistribution of income is necessary in order to achieve a Pareto optimum.

In both cases, the expression is strictly smaller than (2.11).

If, however, the distribution of the initial wealth can be optimally adapted, i.e. $\Delta w^* = \frac{\eta^2}{4c}\left(\frac{a_H^2 - a_L^2}{(1-\beta)^2}\right)$, the total contribution of two heterogenous groups is

$$\begin{aligned} 2\left(\frac{\eta^2 a_H^2}{4(1-\beta)^2 c^2} + \frac{\eta^2 a_H^2}{4(1-\beta)^2 c^2} - \frac{\Delta w^*}{c}\right) &= \frac{2\eta^2 a_H^2}{4(1-\beta)^2 c^2} + \frac{2\eta^2 a_H^2}{4(1-\beta)^2 c^2} - \frac{2\eta^2}{4c^2}\left(\frac{a_H^2 - a_L^2}{(1-\beta)^2}\right) \\ &= \frac{2\eta^2\left(a_H^2 + a_L^2\right)}{4(1-\beta)^2 c^2}. \end{aligned}$$

But this is equal to the total contribution of the homogenous groups which is again given by (2.11) as $\Delta w^* = 0$ is optimal in this case. ■

Hence, when the wealth level is fixed and equally distributed it is beneficial to have homogenous groups. The reason is straightforward from the analysis above: Heterogeneity in abilities leads to a de-motivation of the more qualified agent when wealth is equally distributed. By matching agents into homogenous teams, this de-motivational effect can be avoided and group homogeneity helps the agents to coordinate on more favorable equilibria.

It is, however, interesting to note that group composition is irrelevant for total contributions when the wealth level can be optimally adapted. In this case, the disadvantage of the more able agent can be entirely offset and, in turn, motivation to contribute is restored to the levels attainable in homogenous groups.

2.6 Conclusion

We analyzed the effects of wealth inequality on the incentives to contribute to a public good when agents are inequity averse. We have shown that it is optimal to introduce ex-ante inequality in wealth if agents differ in their abilities. The reason is that inequality in favor of a more able agent can motivate this agent to exert higher efforts. In particular, the stronger the agents' inequity aversion, the stronger is also this incentive effect of inequality and the larger should be the difference in initial wealth. Furthermore, we have shown that compared to the case when agents are purely self-interested, contributions are higher when agents are inequity averse as inequity aversion helps to reduce the free-rider problem and agents can coordinate on higher efforts.

Our results have several interesting implications. First of all, they cast doubt on simple statements sometimes heard in practice claiming that inequality among the members of a group is demotivating when people care for fairness. While this is indeed true for very large wealth differentials in our model, the opposite can also be the case, when wealth differentials are too small. Allocating agents of different abilities the same initial wealth can lead to highly inequitable situations. The reason is that in a public good setting, all agents equally benefit from the group output, but more able agents exert higher efforts as their marginal returns to effort are higher and, in turn, they incur higher costs. When agents are inequity averse this can demotivate the more able agents which is bad for the overall performance as their contributions are more valuable.

The results also may cast some light on the discussion about distributional politics (Alesina and Angeletos (2005), Durante and Putterman (2009)) and the effects on citizens' willingness to voluntary donate to a common good. Some previous studies (e.g. Warr (1983) and Bergstrom et al. (1986)) have argued that the total provision of a public good is independent of the distribution of wealth. In contrast, our results indicate that equality in wealth may crowd-out the motivation to contribute. But introducing inequality may have positive effects on the citizens' willingness to work for the common good. However, our model also shows that this is the case only if the higher wealth is in the hands of those who can provide the most valuable contributions.

2.7 Appendix

2.7.1 Proof of Corollary 1:

The value of $e_i^{\max}$ directly follows from the upper boundary given by (2.7). Let

$$v_i^E(e_i) = w_i - ce_i + \eta \left(a_i\sqrt{e_i} + a_j\sqrt{e_i - \frac{\Delta w_i}{c}} \right)$$

be agent i's utility which is equal to agent j's utility in any equitable equilibrium. To compare the equilibria in the set defined by (2.7) we have to check which value of e_i maximizes this utility. Note that

$$\begin{aligned} \frac{\partial v_i^E(e_i)}{\partial e_i} &= -c + \eta \left(\frac{a_i}{2\sqrt{e_i}} + \frac{a_j}{2\sqrt{e_i - \frac{\Delta w_i}{c}}} \right) \text{ and} \\ \frac{\partial^2 v_i^E(e_i)}{\partial e_i^2} &= \eta \left(-\frac{a_i}{4} e_i^{-\frac{3}{2}} - \frac{a_j}{4} \left(e_i - \frac{\Delta w_i}{c} \right)^{-\frac{3}{2}} \right) < 0. \end{aligned}$$

As $v_i^E(e_i)$ is strictly concave, $\left.\frac{\partial v_i^E(e_i)}{\partial e_i}\right|_{e_i = e_i^{\max}} \geq 0$ is a necessary and sufficient condition for $e_i^{\max}$ to be Pareto optimal. If $\Delta w_i < \frac{\eta^2}{4c}\left(\frac{a_i^2}{(1-\beta)^2} - \frac{a_j^2}{(1-\beta)^2}\right)$, $e_i^{\max}$ is equal to $\frac{\eta^2 a_j^2}{4(1-\beta)^2 c^2} + \frac{\Delta w_i}{c}$ and the condition is equivalent to

$$-c + \eta \left(a_i \frac{1}{2\sqrt{\frac{\eta^2 a_j^2}{4(1-\beta)^2 c^2} + \frac{\Delta w_i}{c}}} + a_j \frac{1}{2\sqrt{\frac{\eta^2 a_j^2}{4(1-\beta)^2 c^2} + \frac{\Delta w_i}{c} - \frac{\Delta w_i}{c}}} \right) \geq 0 \Leftrightarrow$$

$$\Delta w_i \leq \frac{\eta^2}{4c} \left(\frac{a_i^2}{\beta^2} - \frac{a_j^2}{(1-\beta)^2} \right).$$

But $\frac{\eta^2}{4c}\left(\frac{a_i^2}{\beta^2} - \frac{a_j^2}{(1-\beta)^2}\right) \geq \frac{\eta^2}{4c}\left(\frac{a_i^2}{(1-\beta)^2} - \frac{a_j^2}{(1-\beta)^2}\right)$ as long as $\beta \leq \frac{1}{2}$. Hence, both agent's utility is maximal at $e_i^{\max}$ in this case. If, however, $\Delta w_i \geq \frac{\eta^2}{4c}\left(\frac{a_i^2}{(1-\beta)^2} - \frac{a_j^2}{(1-\beta)^2}\right)$, $e_i^{\max}$ is equal to $\frac{\eta^2 a_i^2}{4(1-\beta)^2 c^2}$ and the condition is equivalent to

$$-c+\eta\left(a_i\frac{1}{2\sqrt{\frac{\eta^2a_i^2}{4(1-\beta)^2c^2}}}+a_j\frac{1}{2\sqrt{\frac{\eta^2a_i^2}{4(1-\beta)^2c^2}-\frac{\Delta w_i}{c}}}\right)\geq 0\Leftrightarrow$$

$$\Delta w_i\geq\frac{\eta^2}{4c}\left(\frac{a_i^2}{(1-\beta)^2}-\frac{a_j^2}{\beta^2}\right)$$

But $\frac{\eta^2}{4c}\left(\frac{a_i^2}{(1-\beta)^2}-\frac{a_j^2}{\beta^2}\right)\leq\frac{\eta^2}{4c}\left(\frac{a_i^2}{(1-\beta)^2}-\frac{a_j^2}{(1-\beta)^2}\right)$ is again equivalent to $\beta\leq\frac{1}{2}$. ■

2.7.2 Proof of Proposition 4:

By substituting the equilibrium efforts (2.9) into agent i's utility function we obtain:

$$u_i=\begin{cases}\frac{W+\Delta w_i}{2}+\eta\left(a_i\sqrt{\frac{\eta^2a_i^2}{4(1+\alpha)^2c^2}}+a_j\sqrt{\frac{\eta^2a_j^2}{4(1-\beta)^2c^2}}\right)-\frac{\eta^2a_i^2}{4(1+\alpha)^2c}+\alpha\left(\Delta w_i-\frac{\eta^2a_i^2}{4(1+\alpha)^2c}+\frac{\eta^2a_j^2}{4(1-\beta)^2c}\right)\\ if\ \Delta w_i<\frac{\eta^2}{4c}\left(\frac{a_i^2}{(1+\alpha)^2}-\frac{a_j^2}{(1-\beta)^2}\right)\\ \frac{W+\Delta w_i}{2}+\eta\left(a_i\sqrt{\frac{\eta^2a_j^2}{4(1-\beta)^2c^2}+\frac{\Delta w_i}{c}}+a_j\sqrt{\frac{\eta^2a_j^2}{4(1-\beta)^2c^2}}\right)-\left(\frac{\eta^2a_j^2}{4(1-\beta)^2c}+\Delta w_i\right)\\ if\ \Delta w_i\in\left[\frac{\eta^2}{4c}\left(\frac{a_i^2}{(1+\alpha)^2}-\frac{a_j^2}{(1-\beta)^2}\right),\frac{\eta^2}{4c}\left(\frac{a_i^2-a_j^2}{(1-\beta)^2}\right)\right]\\ \frac{W+\Delta w_i}{2}+\eta\left(a_i\sqrt{\frac{\eta^2a_i^2}{4(1-\beta)^2c^2}}+a_j\sqrt{\frac{\eta^2a_i^2}{4(1-\beta)^2c^2}-\frac{\Delta w_i}{c}}\right)-\frac{\eta^2a_i^2}{4(1-\beta)^2c}\\ if\ \Delta w_i\in\left[\frac{\eta^2}{4c}\left(\frac{a_i^2-a_j^2}{(1-\beta)^2}\right),\frac{\eta^2}{4c}\left(\frac{a_i^2}{(1-\beta)^2}-\frac{a_j^2}{(1+\alpha)^2}\right)\right]\\ \frac{W+\Delta w_i}{2}+\eta\left(a_i\sqrt{\frac{\eta^2a_i^2}{4(1-\beta)^2c^2}}+a_j\sqrt{\frac{\eta^2a_j^2}{4(1+\alpha)^2c^2}}\right)-\frac{\eta^2a_i^2}{4(1-\beta)^2c}-\beta\left(\Delta w_i-\frac{\eta^2a_i^2}{4(1-\beta)^2c}+\frac{\eta^2a_j^2}{4(1+\alpha)^2c}\right)\\ if\ \Delta w_i>\frac{\eta^2}{4c}\left(\frac{a_i^2}{(1-\beta)^2}-\frac{a_j^2}{(1+\alpha)^2}\right)\end{cases}.$$

The first derivative of this function is

$$\frac{\partial u_i}{\partial \Delta w_i}=\begin{cases} \frac{1}{2}+\alpha & if\ \Delta w_i < \frac{\eta^2}{4c}\left(\frac{a_i^2}{(1+\alpha)^2}-\frac{a_j^2}{(1-\beta)^2}\right) \\ -\frac{1}{2}+\frac{\eta\cdot a_i}{2c\sqrt{\frac{\eta^2 a_j^2}{4(1-\beta)^2c^2}+\frac{\Delta w_i}{c}}} & if\ \Delta w_i \in \left[\frac{\eta^2}{4c}\left(\frac{a_i^2}{(1+\alpha)^2}-\frac{a_j^2}{(1-\beta)^2}\right), \frac{\eta^2}{4c}\left(\frac{a_i^2-a_j^2}{(1-\beta)^2}\right)\right] \\ \frac{1}{2}-\frac{\eta\cdot a_j}{2c\sqrt{\frac{\eta^2 a_i^2}{4(1-\beta)^2c^2}-\frac{\Delta w_i}{c}}} & if\ \Delta w_i \in \left[\frac{\eta^2}{4c}\left(\frac{a_i^2-a_j^2}{(1-\beta)^2}\right), \frac{\eta^2}{4c}\left(\frac{a_i^2}{(1-\beta)^2}-\frac{a_j^2}{(1+\alpha)^2}\right)\right] \\ \frac{1}{2}-\beta & if\ \Delta w_i > \frac{\eta^2}{4c}\left(\frac{a_i^2}{(1-\beta)^2}-\frac{a_j^2}{(1+\alpha)^2}\right) \end{cases}. \tag{2.12}$$

Note that the slope in the second interval is *strictly* positive if

$$\begin{aligned} -\frac{1}{2}+\frac{\eta\cdot a_i}{2c\sqrt{\frac{\eta^2 a_j^2}{4(1-\beta)^2c^2}+\frac{\Delta w_i}{c}}} &> 0 \Leftrightarrow \\ \frac{\eta^2}{4c}\left(4a_i^2-\frac{a_j^2}{(1-\beta)^2}\right) &> \Delta w_i \end{aligned}$$

which is always true for any $\Delta w_i \in \left[\frac{\eta^2}{4c}\left(\frac{a_i^2}{(1+\alpha)^2}-\frac{a_j^2}{(1-\beta)^2}\right), \frac{\eta^2}{4c}\left(\frac{a_i^2-a_j^2}{(1-\beta)^2}\right)\right[$ and $\beta \leq \frac{1}{2}$. Furthermore, it is always positive at $\Delta w_i = \frac{\eta^2}{4c}\left(\frac{a_i^2-a_j^2}{(1-\beta)^2}\right)$ for any $\beta \leq \frac{1}{2}$ and equal to zero if and only if $\beta = \frac{1}{2}$.

Similarly, the slope in third interval is *strictly* negative if

$$\begin{aligned} \frac{1}{2}+\eta\cdot a_j\frac{-\frac{1}{c}}{2\sqrt{\frac{\eta^2 a_i^2}{4(1-\beta)^2c^2}-\frac{\Delta w_i}{c}}} &< 0 \Leftrightarrow \\ \frac{\eta^2}{4c}\left(\frac{a_i^2}{(1-\beta)^2}-4a_j^2\right) &< \Delta w_i \end{aligned}$$

which is always true as well for any $\Delta w_i \in \left]\frac{\eta^2}{4c}\left(\frac{a_i^2-a_j^2}{(1-\beta)^2}\right), \frac{\eta^2}{4c}\left(\frac{a_i^2}{(1-\beta)^2}-\frac{a_j^2}{(1+\alpha)^2}\right)\right]$ and $\beta \leq \frac{1}{2}$. Furthermore, it is always negative at $\Delta w_i = \frac{\eta^2}{4c}\left(\frac{a_i^2-a_j^2}{(1-\beta)^2}\right)$ for any $\beta \leq \frac{1}{2}$ and equal to zero if and only if $\beta = \frac{1}{2}$. ■

2.7.3 Proof of Corollary 2:

The utility of agent j at $\Delta w_i = 0$ is always smaller as compared to $\Delta w_i = \frac{\eta^2}{4c}\left(\frac{a_i^2-a_j^2}{(1-\beta)^2}\right)$ if agent j's utility function is increasing at $\Delta w_i = 0$ which is the case when

$$0 > \frac{\eta^2}{4c}\left(\frac{a_i^2}{(1+\alpha)^2} - \frac{a_j^2}{(1-\beta)^2}\right) \Leftrightarrow \frac{a_j}{a_i} > \frac{1-\beta}{1+\alpha}.$$

If this is not the case we have to compare the utility of the less able agent j at the local maximum of both agents' utility function

$$\frac{W}{2} - \frac{\eta^2\left(a_i^2+a_j^2\right)}{8(1-\beta)^2 c} + \frac{\eta^2\left(a_i^2+a_j^2\right)}{2(1-\beta)c}$$

with her utility at $\Delta w_i = 0$, which for $\frac{a_j}{a_i} < \frac{1-\beta}{1+\alpha}$ is given by

$$\frac{W}{2} + \frac{\eta^2 a_j^2}{2(1-\beta)c} + \frac{\eta^2 a_i^2}{2(1+\alpha)c} - \frac{\eta^2 a_j^2}{4(1-\beta)^2 c} + \beta\left(\frac{\eta^2 a_j^2}{4(1-\beta)^2 c} - \frac{\eta^2 a_i^2}{4(1+\alpha)^2 c}\right).$$

Hence, agent j is better off with an unequal income when

$$\frac{a_j^2}{a_i^2} > \frac{1}{(1-2\beta)}\left(\frac{4(1-\beta)^2}{(1+\alpha)} + 1 - 4(1-\beta) - \frac{2\beta(1-\beta)^2}{(1+\alpha)^2}\right).$$

Note that if $\left(\frac{4(1-\beta)^2}{(1+\alpha)} + 1 - 4(1-\beta) - \frac{2\beta(1-\beta)^2}{(1+\alpha)^2}\right) < 0$ this holds for all values of a_j. As $4\beta - 3 < 0$ this condition is equivalent to

$$\left((1+\alpha) - \frac{2(1-\beta)^2}{(3-4\beta)}\right)^2 > \frac{2(1-\beta)^2}{(3-4\beta)^2}\left(6\beta^2 - 7\beta + 2\right).$$

Note that $6\beta^2 - 7\beta + 2 > 0$ as this function is $= 0$ at $\beta = \frac{1}{2}$ and decreasing for $0 < \beta < \frac{1}{2}$. Rearranging the equation gives

$$\alpha > \frac{1-\beta}{3-4\beta}\sqrt{2\left(6\beta^2 - 7\beta + 2\right)} + \frac{2(1-\beta)^2}{(3-4\beta)} - 1$$

which proves the first claim. If, however, $\left(\frac{4(1-\beta)^2}{(1+\alpha)} + 1 - 4(1-\beta) - \frac{2\beta(1-\beta)^2}{(1+\alpha)^2}\right) > 0$, the condition is equivalent to

$$\frac{a_j}{a_i} > \sqrt{\frac{1}{(1-2\beta)}\left(\frac{4(1-\beta)^2}{(1+\alpha)} + 1 - 4(1-\beta) - \frac{2\beta(1-\beta)^2}{(1+\alpha)^2}\right)}$$

which establishes the second claim. ■

Chapter 3

Overconfidence and Managers' Responsibility Hoarding[1]

3.1 Introduction

Overconfidence is a well-established behavioral pattern that involves overestimating the own capabilities, especially in tasks with a partially stochastic outcome (Svenson (1981), Lichtenstein et al. (1982), Russo and Schoemaker (1992), Soll (1996)). Although the degree of overconfidence may vary with the type of task (Grieco and Hogarth (2009)), it is generally found to persist when individuals assess the probability of their own success or the relative standing of their performance compared to others (Klayman et al. (1999)).

While the behavioral pattern of overconfidence and its effects on financial decision-making have been studied extensively (see for instance De Bondt and Thaler (1996)), the effects of overconfidence on organizational performance are not fully understood yet. In particular, the question how manager overconfidence affects organizational performance by biasing the manager's delegation and task distribution choices has not been studied so far.

To investigate these effects, we introduce a model, in which a manager and an agent can exert effort into a joint production that consists of two distinct tasks with unequal impact on the output. The allocation of the tasks is at the

[1] This chapter is based upon Nieken et al. (2011).

discretion of the manager, who decides whether to perform the "critical" task (i.e. the task with the higher impact on the output) himself or to delegate it to the agent.[2]

We show that an overconfident manager tends to *hoard responsibility*, i.e. to assign the critical task more often to himself than a fully rational manager would. Responsibility hoarding takes place, even though it is individually suboptimal for the manager, who suffers from a higher cost of effort by performing the critical task instead of the other task.

We also show that, despite adding to the overconfident manager's effort cost, responsibility hoarding may actually increase the total welfare of the involved parties. As long as the overconfident manager's self-perception bias is not too large, the total welfare effect can be positive, because the amount of effort exerted by the overconfident manager is closer to the efficient level than the amount provided by a fully rational manager, who chooses a payoff maximizing effort level, generally below the welfare maximizing level. Hence, by overestimating his own productivity and exerting a correspondingly greater amount of effort, the overconfident manager typically engages in less free-riding than his rational counterpart.

Finally, we show that responsibility hoarding can persist in an organization, as long as the overconfident manager can rationalize the overestimation of the own ability by underestimating the ability of the agent. The more leeway an overconfident manager has to rationalize the observed outcome without having to adapt his positively biased assessment of the own ability level, the more likely it is to observe persistent individually suboptimal but welfare improving delegation behavior in an organization.

Most of the existing literature on overconfidence in managerial settings is focused on the exaggerated investment risks taken by overconfident managers. While Barber and Odean (2000) and Deaves et al. (2008) report excessive trading by overconfident traders, Camerer and Lovallo (1999) observe excessive market entries in an experimental setting. Malmendier and Tate (2005)

[2] We use the male pronoun for the manager and the female pronoun for the agent, because males are generally found to exhibit a higher degree of overconfidence than females (see e.g. Barber and Odean (2001), Correll (2001), Bengtsson et al. (2005)).

argue that managerial overconfidence leads to distortive investment behavior and demonstrate that data on CEO investments in the own company are in line with their overconfidence model. Hackbarth (2008) shows that overconfident managers tend to bias the capital structure of the firm towards higher debt levels. Similar results can be found in Ben-David et al. (2007) who show that companies with overconfident CFOs have a significantly different capital structure than other firms. Furthermore, Malmendier and Tate (2008) find that overconfident managers also tend to overpay in mergers and even initiate value-destroying ones. Interestingly, this bias is sometimes advantageous for the firm value. For instance, Palomino and Sadrieh (2011) show that managerial overconfidence can be advantageous concerning financial decisions. They analyze a model in which overconfident portfolio managers, who share profits, may exhibit risk attitudes that are more in line with the investors' risk attitude than fully rational risk-averse managers. Analyzing data of large publicly traded firms from 1980 to 1994, Galasso and Simcoe (2011) present evidence that overconfident CEOs have a significantly higher probability to initiate corporate innovation.

In a team production setting, Gervais and Goldstein (2007) study a firm with complementary production technology and show that the presence of an overconfident agent can increase the firm output as it helps the agents to coordinate on a high effort level, and therefore, overcome the free-rider problem. Furthermore, Santos-Pinto (2008) shows that firms can benefit from using interdependent incentive schemes when workers exhibit mistaken beliefs about their coworkers' abilities. Regarding individual performance, Weinberg (2009) for instance shows that a moderate overestimation of own ability can also be advantageous relative to a realistic assessment as it lets overconfident individuals undertake more challenging tasks that might raise their expected output and utility. Recent experimental findings by Sautmann (2011) support the theory that overconfident agents accept lower wage offers, while Santos-Pinto (2010) shows that firms using tournaments as incentives can make higher profits if agents have a positive self image. Similarly, Ludwig et al. (2011b) find that moderate overconfidence can improve the agent's performance in a Tullock contest relative to an unbiased opponent resulting in an advantage for the overconfident agent. These results are supported by the recent experiment of Kinari et al. (2011) who report overconfidence to have a significant impact on increasing productivity in tournaments. Furthermore, Englmaier (2011) argues that firms should hire overoptimistic managers to ensure the implementation of certain investment strategies in R&D tournaments.

Regarding the literature dealing with the delegation of tasks in a principal-agent model, we are, to the best of our knowledge, the first to take overconfidence into account. Prendergast (1995) suggests a model where a manager has discretion over task assignments. In this setting, the manager may exhibit responsibility hoarding, i.e. does not delegate enough and carries out too many tasks himself. This is driven by the assumption that the manager can earn future rents from the on-the-job training that performing the additional tasks provides.[3] If the output of several tasks cannot be measured separately and the principal has to delegate at least one task, Itoh (1994) and Itoh (2001) find that the principal will execute some tasks himself or delegate all tasks to only one agent if the agents are risk adverse. Gürtler (2008) extends this model and compares partial delegation where the principal carries out one task and the other task is carried out by the agent with complete delegation with specialization of the agents on one certain task.

The remainder of this chapter is structured as follows. The basic model is described in section 3.2. In section 3.3, we present the first-best task allocation and the individually optimal task allocation that is chosen by the manager under perfect information on the agent's ability and analyze the effects of manager overconfidence on total welfare. In section 3.4, we consider the persistence of manager overconfidence and the underestimation of the agent's ability by an overconfident manager. In section 3.5, we discuss and outline the range of optimal and persistent manager overconfidence and responsibility-hoarding constellations. Section 3.6 concludes.

[3] This model of rational responsibility hoarding is especially useful when studying professions with extraordinary high rents for job experience, e.g. surgeons, pilots, or lawyers. Note, however, that even in such settings any degree of rational responsibility hoarding may be amplified by the manager's overconfidence.

3.2 The model

Consider a joint production setting, in which a manager and an agent can exert effort to generate a shared output. The total output Y is a function of the outcomes Y_1 and Y_2 of the tasks 1 and 2, correspondingly. A crucial assumption is that the two tasks differ in their impact on the total output, where w.l.o.g. we assume that task 1 is the critical task, i.e. it tends to have higher impact on total output than the non-critical task 2. Using an additive production function, we introduce the parameter λ that measures the relative impact of task 1 compared to task 2. Hence, the total output Y is defined as:

$$Y = \lambda Y_1 + (1 - \lambda) Y_2, \text{ with } \lambda \in \left(\frac{1}{2}, 1\right)$$

For simplicity, both manager and agent are risk neutral and benefit to the same extent from the total output, i.e. both individuals receive the same share of $\frac{1}{2}Y$. We assume that the effort levels as well as the output of both tasks is not observable by the firm. Only the total output Y is observable for all parties. Either player can be assigned to perform either task, where the task allocation is chosen by the manager at the outset of each period. The allocation must be complete and bijective, i.e. both tasks must be allocated and each must be allocated to a different player, because no player can perform both tasks in the same period. The outcome of each task j is endogenous, depending on the true ability a_i and the chosen effort e_i of the player i performing the task:

$$Y_j = a_i \cdot e_i, \text{ with } i = M[anager], A[gent] \text{ and } j = 1, 2.$$

Furthermore, the individual effort cost $C(e_i)$ is strictly convex:

$$C(e_i) = \frac{c}{2} e_i^2, \ c \in \mathbb{R}^+.$$

To simplify the discussion, we distinguish between those task allocation choices, in which the critical task 1 is performed by the manager (*non-delegation*), and those, in which the critical task 1 is allocated to the agent (*delegation*). More formally we define:

Definition 1 *A* ***delegation choice*** *is a task allocation in which the critical task 1 is allocated to the agent and the non-critical task 2 is allocated to the manager. In case of* ***non-delegation*** *the critical task 1 is allocated to the manager and the non-critical task 2 is carried out by the agent.*

Definition 2 *A task allocation is* ***individually optimal*** *if it maximizes the manager's utility.*

Definition 3 *A task allocation is* ***efficient*** *if it maximizes the total welfare of all involved parties.*

Following Gervais and Goldstein (2007) we characterize an overconfident manager as someone who systematically overestimates his own ability:

Definition 4 *An* ***overconfident*** *manager has an overly optimistic perception of his own ability, i.e.*

$$a_M^{OC} = a_M + b$$

where a_M denotes the manager's true ability and the parameter $b > 0$ his self-perception bias or the degree of his overconfidence.

Finally, we use the definitions above to characterize responsibility hoarding.

Definition 5 ***Responsibility hoarding*** *occurs when a manager performs the critical task (task 1) himself, even though a delegation choice is individually optimal for him.*

3.3 Perfect information on agent's ability

In the first step, we assume that the agent's ability is common knowledge, i.e. both the manager and the agent have perfect information on the agent's ability. We start by investigating the efficient task allocation choice (first-best case) and then proceed to the delegation choices of a fully rational manager and of an overconfident manager. Comparing the three results, we first show that fully rational managers delegate the critical task more often than is efficient. Next, we show that manager overconfidence always leads to less delegation compared to an equilibrium with fully rational managers. Finally, we prove that the manager's biased self-perception may increase efficiency, because responsibility hoarding can be beneficial for the total welfare of the involved parties, as long as the increase in the overconfident manager's contribution to firm output over-compensates the loss due to his individually suboptimal delegation and effort choices.

3.3.1 Efficient task allocation (first-best case)

Assume that the critical task 1 is carried out by the manager and task 2 is assigned to the agent. The outcomes of the two tasks are then given by

$$\begin{aligned} Y_1^{nd} &= a_M e_M \\ Y_2^{nd} &= a_A e_A \end{aligned} \tag{3.1}$$

where the index nd denotes the case of non-delegation.

In the first-best case the total welfare of the involved parties is maximized by:

$$\max_{e_M, e_A} W^{nd} = (\lambda a_M e_M + (1-\lambda) a_A e_A) - \frac{c}{2} e_M^2 - \frac{c}{2} e_A^2$$

which leads to the first-best effort levels given by

$$\begin{aligned} e_M^{ndFB} &= \frac{\lambda a_M}{c} \\ e_A^{ndFB} &= \frac{(1-\lambda) a_A}{c}. \end{aligned}$$

Hence, the total welfare in case of non-delegation is equal to

$$W^{ndFB} = \frac{\lambda^2 a_M^2}{2c} + \frac{(1-\lambda)^2 a_A^2}{2c}. \tag{3.2}$$

Next, assume that the critical task 1 is assigned to the agent and task 2 is carried out by the manager. Now, the outcomes of both tasks are given by

$$\begin{aligned} Y_1^d &= a_A e_A \\ Y_2^d &= a_M e_M \end{aligned} \tag{3.3}$$

where the index d denotes the case of delegation.

In this case, the welfare maximization problem becomes

$$\max_{e_M, e_A} W^d = \left((1-\lambda)\, a_M e_M + \lambda a_A e_A \right) - \frac{c}{2} e_M^2 - \frac{c}{2} e_A^2$$

which leads to the first-best effort levels described by

$$\begin{aligned} e_M^{dFB} &= \frac{(1-\lambda)\, a_M}{c} \\ e_A^{dFB} &= \frac{\lambda a_A}{c}. \end{aligned}$$

Hence, the total welfare in case of delegation is equal to

$$W^{dFB} = \frac{(1-\lambda)^2 a_M^2}{2c} + \frac{\lambda^2 a_A^2}{2c}. \tag{3.4}$$

Comparing (3.2) and (3.4), delegation is efficient if and only if

$$a_M \leq a_A.$$

Proposition 1 *In the efficient task allocation the critical task should be allocated to the agent if and only if her ability is at least as high as the manager's ability, i.e.* $a_M \leq a_A$*. Otherwise, the critical task should better be assigned to the manager.*

It is straightforward that maximizing the total welfare requires that the critical task (i.e. the task with a higher impact on the total output) to be carried out by the individual with the higher ability. Moreover, the positive welfare effect of delegation is the higher the higher the agent's ability is. However, it is not obvious that the welfare maximizing task allocation will generally be implemented when the task allocation is chosen by the manager.

3.3.2 Optimal delegation choice of a fully rational manager

Now, assume that the task allocation is chosen by a fully rational manager, maximizing his individual utility. Assume that the manager does not delegate, i.e. the critical task (task 1) is carried out by the manager and other task (task 2) is assigned to the agent. The outcomes of the two tasks are then given by (3.1).

In contrast to the first-best case, the manager's utility is now maximized with:

$$\max_{e_M} U_M^{nd} = \frac{1}{2}\left(\lambda a_M e_M + (1-\lambda)\, a_A e_A\right) - \frac{c}{2} e_M^2$$

which leads to his individually optimal effort level described by

$$e_M^{nd*} = \frac{\lambda a_M}{2c}.$$

Furthermore, the agent's optimization is given by

$$\max_{e_A} U_A^{nd} = \frac{1}{2}\left(\lambda a_M e_M + (1-\lambda)\, a_A e_A\right) - \frac{c}{2} e_A^2$$

and her individually optimal effort level is

$$e_A^{nd*} = \frac{(1-\lambda)\, a_A}{2c}.$$

Since we assume that the manager has perfect information about the agent's ability, his utility is equal to

$$U_M^{nd*} = \frac{\lambda^2 a_M^2}{8c} + \frac{(1-\lambda)^2 a_A^2}{4c}. \tag{3.5}$$

Next, assume that the manager delegates, i.e. the critical task 1 is assigned to the agent and task 2 is carried out by the manager. Now, the outcomes of both tasks are given by (3.3).

In this case, the manager's optimization problem is

$$\max_{e_M} U_M^d = \frac{1}{2}\left(\left(1-\lambda\right)a_M e_M + \lambda a_A e_A\right) - \frac{c}{2}e_M^2$$

with

$$e_M^{d*} = \frac{\left(1-\lambda\right)a_M}{2c}$$

as his individually optimal effort level.

For the agent, the optimization is given by

$$\max_{e_A} U_A^d = \frac{1}{2}\left(\left(1-\lambda\right)a_M e_M + \lambda a_A e_A\right) - \frac{c}{2}e_A^2$$

which leads to an individually optimal effort level of

$$e_A^{d*} = \frac{\lambda a_A}{2c}.$$

Hence, the manager's utility in case of delegation is equal to

$$U_M^{d*} = \frac{\left(1-\lambda\right)^2 a_M^2}{8c} + \frac{\lambda^2 a_A^2}{4c}. \tag{3.6}$$

Now, by comparing (3.5) and (3.6), the fully rational manager chooses delegation if and only if

$$a_M \leq \sqrt{2}a_A.$$

It is straightforward that the fully rational manager prefers to delegate the critical task as long as his own ability is smaller than the agent's ability, i.e. as long as $a_M \leq a_A$. Moreover, note that there is a range of values (i.e. $a_M \in \left(a_A; \sqrt{2}a_A\right]$) for which the manager also delegates the critical task to the agent, even though his ability is strictly higher than the agent's ability. This is due to the fact that in equilibrium the critical task 1 is performed with higher levels of effort and, thus, with a higher effort cost, than the other task. Hence, delegating the task may pay, because delegation reduces the manager's effort cost more than it reduces the expected outcome of the critical task when it is performed by the agent with the somewhat lower ability. As the fully rational manager cannot commit to the efficient task allocation this may lead to inefficient job distributions and lower total welfare in equilibrium. However, once the agent's ability falls below the threshold $\frac{\sqrt{2}}{2}a_M$, the manager prefers to perform the critical task himself, because the benefit from the own higher ability surpasses the higher effort cost.

3.3.3 Optimal delegation choice of an overconfident manager

In this section, we examine the task allocation choice of an overconfident manager, assuming that overconfidence leads to an overly optimistic perception of the own abilities. Recall that the self-perceived ability of an overconfident manager is given by

$$a_M^{OC} = a_M + b, \text{ with } b > 0.$$

Given this slight modification of the model, we derive the equilibrium choices of the overconfident manager and the agent and compare these to the case with a fully rational manager. By substituting a_M^{OC} for a_M and following the same procedure applied in the previous section, we derive the condition under which the overconfident manager chooses delegation:

$$a_M \leq \sqrt{2}a_A - b.$$

By decreasing the right-hand side of the inequality, any positive self-perception bias b lowers the threshold for non-delegation, reducing the range of values for which delegation is chosen by the overconfident manager. Hence, it is obvious that an overconfident manager is more likely to hoard responsibility than a fully rational manager of the same true ability. In particular, the higher the self-perception bias b, the larger the range of ability values for which a fully rational manager delegates the critical task, but an overconfident manager will not, i.e. the greater the range of ability values in which the critical task is carried out by the manager. We summarize our findings in the following proposition:

Proposition 2 *With perfect information on the agent's ability parameter*

a_A, any positive self-perception bias $b > 0$ leads to responsibility hoarding by the overconfident manager. In particular, the range of manager types choosing delegation is strictly decreasing in the managers' degree of overconfidence b.

As we have shown in the last section, fully rational managers cannot commit to the efficient task allocation as they have an incentive to lower their own effort cost by delegating the critical task to the agent as long as the agent's ability is sufficiently high. In contrast, overconfident managers overestimate their own ability, and therefore, allocate the critical task more often to themselves than fully rational managers. In particular, overconfident managers are more likely to hoard responsibility the larger their self-perception bias is. However, the task allocations chosen by overconfident managers may be closer to the efficient allocation than those of rational managers. Hence, overconfidence can be considered as a commitment device for managers to take more responsibilities and increase the efficiency of the job distribution, positively affecting the total welfare.

3.3.4 Is overconfidence beneficial or harmful?

As we have shown in the previous section, manager overconfidence may lead to less delegation and can, thus, improve the efficiency of the task allocations. Since an overconfident manager in general exerts more effort, the total output of the firm is often higher than with a rational manager. The higher effort level, however, also leads to a higher cost of effort for the overconfident manager than for the rational manager. Hence, it is not clear whether the manager's overconfidence is generally beneficial or harmful with regard to total welfare. In this section, we show that in many cases, including some in which the task allocation is not individually optimal for the manager, overconfidence is beneficial regarding the total welfare of the involved parties. Comparing the total welfare in equilibrium with a fully rational manager to that with an overconfident manager, we establish the following proposition:

Proposition 3 *If the manager's self-perception bias b is on a moderate level relative to his true ability (i.e. $b < 2a_M$) and his true ability is sufficiently* *high (i.e. $a_M > \sqrt{2}a_A$) or sufficiently low (i.e. $a_M < \sqrt{2}a_A - b$), the total welfare of the involved parties in equilibrium is strictly higher with an overconfident manager than with a fully rational manager. For any ability value of the manager between those two thresholds (i.e. $\sqrt{2}a_A - b \leq a_M \leq \sqrt{2}a_A$), this result still holds if the manager's true ability is at least as high as the agent's true ability (i.e. $a_M \geq a_A$).*

Proof. See the appendix.

This result has several interesting implications. First, note that the manager's overconfidence is not generally harmful and can even be beneficial for the total welfare, if it is not too strong. On the one hand, the overconfident manager overestimates his own ability, and therefore, exerts more effort than the fully rational manager, irrespective of the task allocation. On the other hand, the overconfident manager also expects a higher total outcome when carrying out the critical task himself, and thus, is more likely to allocate the critical task to himself than his fully rational counterpart. This type of responsibility hoarding behavior, in turn, may lead to higher efficiency of the job distribution if the manager is more able than the agent. Hence, overconfidence helps to reduce free-riding. Indeed, this positive incentive effect of manager overconfidence can even over-compensate the negative effect of individually suboptimal task allocation as long as the manager is at least as productive as the agent. Note that this finding is also in line with our result from the first-best case stating that the delegation of the critical task is only efficient if the agent is more productive than the manager. In particular, the total welfare in equilibrium with an overconfident manager is closer to the efficient allocation than with a fully rational manager of the same true ability. Hence, all involved parties may in fact benefit from a moderate level of manager overconfidence.

3.3.5 Optimal degree of manager overconfidence

As the manager's overconfidence can be beneficial for the total welfare, it is straightforward to proceed in our analysis with the determination of its optimal degree with respect to the total welfare. In this regard, we can show:

Proposition 4 *If the manager is overconfident with a positive self-perception bias $b > 0$, the total welfare is highest if the manager's self-perception bias (or degree of overconfidence) is equal to his true ability, i.e. $b^* = a_M$.*

Proof. See the appendix.

Note that the positive welfare effect of overconfidence is strictly increasing in the manager's true ability. Intuitively, the higher the manager's true ability is, the less harmful is his biased self-perception, the more likely responsibility hoarding may positively affect the total welfare. Moreover, it is also straightforward to see that the manager's effort choice exactly matches the efficient level, if the degree of his overconfidence is equal to his true ability. We summarize this result in the following corollary:

Corollary 1 *If a manager's degree of overconfidence is equal to the true value of his ability, i.e. $b = a_M$, his effort choice in equilibrium is exactly equal to the efficient effort level, both in case of delegation and non-delegation.*

Proof. The results follows directly by substituting a_M for b into the overconfident manager's incentive conditions. ■

3.4 Persistence of manager overconfidence

In the previous section, we have demonstrated that manager overconfidence can lead to less delegation, resulting in more efficient task allocations both in a perfect information setting. The question that remains to be answered is whether the managers' overconfidence and responsibility hoarding behavior can persist over time, given the feedback that managers receive from their previous decisions. If managers quickly learn to correct their overconfident assessment of the own ability, then overconfidence and responsibility hoarding will not persist. However, if the feedback from previous outcomes cannot be used to correct overconfidence, we can establish that responsibility hoarding can be a persistent phenomenon with a sustained effect on organizational performance.

In this section, we derive the conditions under which manager overconfidence (and responsibility hoarding) can persist, even though managers receive feedback on their previous decisions. We restrict our analysis to the case that the manager only receives feedback on the total output of the firm (or the organizational unit). Obviously, persistence of manager overconfidence with more exact information, e.g. on all ability and effort parameters, would not be feasible. In the more realistic situation that we analyze, we assume that the agent's ability parameter is her private information. More specifically, we assume that the manager uses an estimate of the agent's ability parameter denoted by $\widehat{a}_A$. Now, the feedback information is restricted to the total output, so that the overconfident manager faces one known parameter (his own effort level), two unknown parameters (the agent's ability and effort level), and one parameter that he believes to know, but actually does not (his own ability). Under these circumstances, we show that the overconfident manager may not be able to learn that his self-assessment is biased, because he can construct a consistent model that explains the observed total outcome with an overestimated own ability parameter and an underestimated ability parameter for the agent.[4] As long as the productivity of the agent can be underestimated sufficiently, the manager's overconfidence can persist.

We summarize this result in the following proposition:

Proposition 5 *If the agent's true ability a_A is sufficiently high, i.e. $a_A \geq \frac{\lambda}{(1-\lambda)}\sqrt{(a_M b + b^2)}$, the manager's overconfidence persists, because the manager rationalizes the observed outcome information by underestimating the agent's ability. The higher the manager's self-perception bias b is, the stronger the underestimation of the agent's ability will be.*

Proof. See the appendix.

A straightforward corollary to the proposition in this section is concerned with the limits of persistent overconfidence:

[4] Young (2002) shows that some games cannot be learned by rational players and demonstrates a class of learning environments in which convergence to equilibrium behavior fails to occur for any learning process, including the Bayesian updating of objectively correct priors.

Corollary 2 *Manager overconfidence is not persistent at any (positive) level if* $a_A = 0$*.*

The corollary simply points out that there is always some level of overconfidence that is persistent, as long as the agent's ability is not zero. Intuitively, it is clear that overconfidence can only persist, as long as the overconfident manager has the possibility to underestimate the agent's contribution to the observed total output, and thus, the ability of the agent. The range for the underestimation drops if the agent's ability is decreased, leaving less and less room for persistently overconfident managers. If the agent's ability is zero, she would not contribute at all to the total output and persistent overconfidence would no longer be possible. But, note that the extreme case of zero ability has no empirical relevance, because it describes a situation in which the agent cannot contribute to the output of the firm. Hence, the corollary shows that for any situation with empirical relevance, there is at least some level of persistent overconfidence, leading to some amount of persistent responsibility hoarding by overconfident managers.

3.5 Discussion

Using the results of the sections above, we discuss the range of optimal and persistent manager overconfidence and responsibility hoarding constellations in this section. The constellations are exhibited in Figure 3.1. It shows the four functions that determine the different outcome regions in the ability space. The manager's ability is plotted on the horizontal and the agent's (possibly estimated) ability is plotted on the vertical axis.[5]

The dashed bisecting line depicts the function $a_A = a_M$ which separates the area of efficient delegation choices above the line from the area of efficient non-delegation choices below the line. The solid line running through the origin depicts the function $a_A = \frac{\sqrt{2}}{2}a_M$ and separates the area of individually optimal delegation (i.e. rational delegation) above the line from the area of individually optimal non-delegation (i.e. rational non-delegation) below the

[5]We have fixed $\lambda = 0.55$ and $b = 4$ to make a two-dimensional plot possible.

line. The area of individually optimal delegation is larger than the area of individually optimal non-delegation, because - as we have seen in section 3.3 - the rational manager always prefers to avoid the high costs of effort for performing the critical task, as long as the agent's ability is not too low.

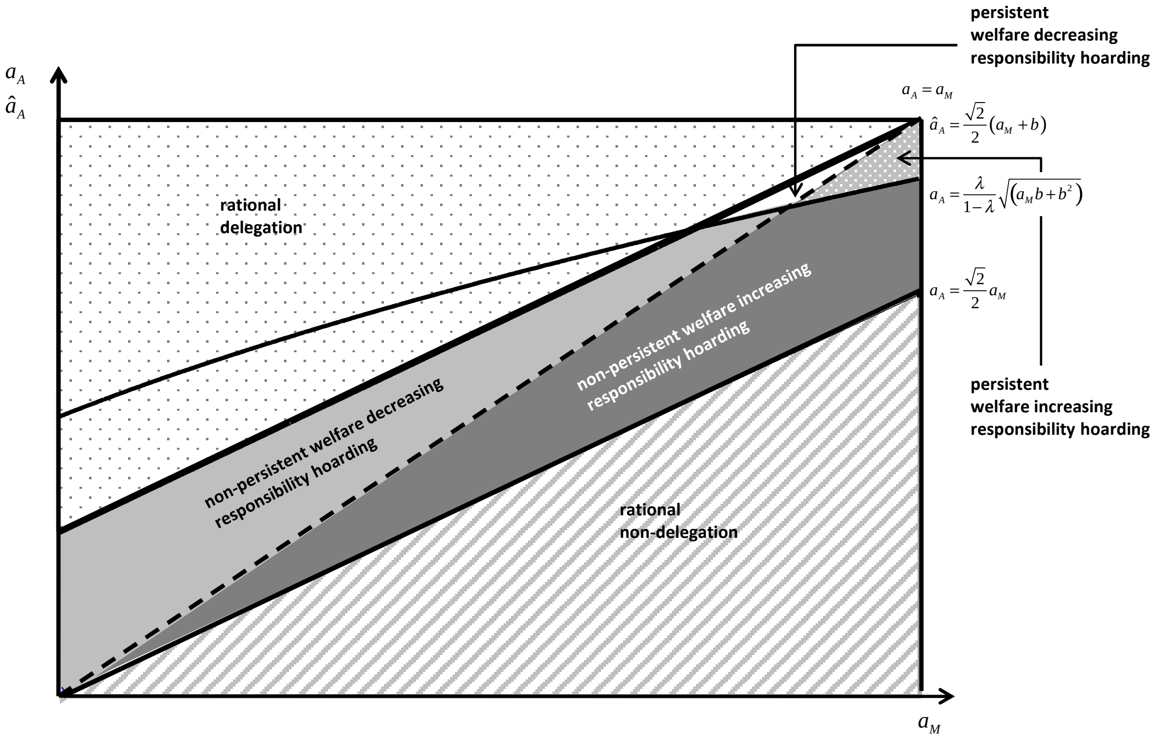

Figure 3.1: Range of optimal and persistent overconfidence and responsibility-hoarding constellations depending on manager's and agent's abilities

The solid line that intersects the vertical axis above the origin depicts the function $\widehat{a}_A = \frac{\sqrt{2}}{2}(a_M + b)$ and separates the area of delegation (the dotted area above the line) from the area of non-delegation (below the line) chosen by an overconfident manager.[6] Note that in the dotted area both the fully rational and the overconfident manager choose to delegate the critical task to the agent, while below the line running through the origin (the dashed area),

[6] $\widehat{a}_A$ denotes the estimated value of the agent's ability used by the manager.

both the fully rational and the overconfident manager choose to carry out the critical task themselves. The area enclosed by the functions $a_A = \frac{\sqrt{2}}{2}a_M$ and $\widehat{a}_A = \frac{\sqrt{2}}{2}(a_M + b)$ contains all ability constellations for responsibility hoarding, in which the overconfident manager still assigns the critical task to himself, but the rational manager does not.

The function $a_A = \frac{\lambda}{(1-\lambda)}\sqrt{(a_M b + b^2)}$ separates the area of persistent (above) from the area of non-persistent manager overconfidence (below). Note that this separation is only valid in the area where responsibility hoarding occurs, i.e. the area enclosed by the functions $a_A = \frac{\sqrt{2}}{2}a_M$ and $\widehat{a}_A = \frac{\sqrt{2}}{2}(a_M + b)$. Our graph shows a large area of non-persistent (the shaded area and the dark grey area) and a relatively small area of persistent manager overconfidence with responsibility hoarding by the overconfident manager (the white area and the white-dotted area). Intuitively, it seems clear that manager overconfidence has a lower chance to persist, if the manager carries out the critical task himself. The reason is that the overconfident manager always has more room to rationalize his overly optimistic self-perception by underestimating the agent's contribution when the critical task that has greater impact on the total outcome is carried out by the agent.

The difference between the tasks concerning their impact on the firm outcome, i.e. the value of the parameter λ, in fact, affects the location of the persistency curve (the function $a_A = \frac{\lambda}{(1-\lambda)}\sqrt{(a_M b + b^2)}$) and, thus, the size of persistent overconfidence areas in the graph (the white area and the white-dotted area). The more important the critical task is when compared to the other task (i.e. the higher λ), the smaller are the areas of persistent overconfidence and responsibility hoarding. The more asymmetric a task constellation is, the more difficult it is for the overconfident manager to find a feasible set of parameters, in which the overestimation of the own ability can be compensated by underestimating the ability of the agent.

A similar but more subtle effect exists concerning the self-perception bias b. As b increases, the area of responsibility hoarding obviously also increases. Note, however, that an increase in the level of overconfidence b also means that the persistency curve shifts upwards, reducing the area of persistent responsibility hoarding. Hence, more overconfident managers will tend to carry out the critical task more frequently, but are also more likely to receive feedback that lets them revise their self-assessment and reduce their overconfidence.

Another implication of our analysis is that both the manager's and the agent's abilities must be relatively high to enable a persistent manager overconfidence. This is because the agent's ability must be high enough to allow for the relatively high degree of underestimation that persistency of overconfidence requires. Since persistent overconfidence of the manager is more likely to occur, when the ability levels of the two players are rather close to each other, responsibility hoarding is most likely to be observed, when the overconfident manager's true ability is close the ability level of a high ability agent.

Finally, there are constellations of ability parameters for which manager overconfidence and responsibility hoarding have a sustained effect on the total welfare of all involved parties (the white area and the white-dotted area). However, persistent manager overconfidence and responsibility hoarding are welfare increasing only if the manager is indeed more able than the agent (the white-dotted area). This finding is in line with our results of the first-best case that the critical task should always be carried out by the more able individual due to its higher impact on total outcome.

3.6 Conclusions and managerial implications

We studied the consequences of manager overconfidence for organizational performance in a setting in which the manager chooses the allocation of tasks. We have proved that an overconfident manager may exhibit responsibility hoarding behavior, i.e. assign the critical task more often to himself than a rational manager would do. We have shown that while responsibility hoarding generally decreases the manager's individual utility, it tends to increase the firm output and the total welfare of the involved parties, when compared to the case of a fully rational manager. The reason for this seemingly counter-intuitive result is that overconfident managers generally exert higher levels of effort than rational managers, due to their overestimation of the own productivity. In this regard, overconfidence counterbalances shirking, causing managers to take up more responsibility and to reduce the inefficiency of effort minimizing task allocation.

Hence, our results imply that firms will not generally avoid overconfident and responsibility hoarding managers, but may even prefer them to fully rational managers. In a situation where the firm cannot establish a contract to enforce the efficient allocation of tasks, moderate overconfidence of a manager can mitigate the negative effects of free-riding. Then, the firm may prefer to hire a moderately overconfident manager to avoid the contract problem. In connection with the well-established evidence that men are generally more overconfident than women (see e.g. Barber and Odean (2001), Correll (2001), Bengtsson et al. (2005)), our result may also provide a further possible explanation why leadership positions are more often occupied by men than by women.[7]

Moreover, we have shown that an overconfident manager's biased self-perception and his responsibility hoarding behavior can persist, as long as the manager can rationalize observed outcomes, by underestimating the ability of the agent. Notably, the probability of persistent overconfidence does not only depend on the level of overconfidence, but also on the absolute level of the players' true abilities. The higher the ability levels in a workplace, the more likely it is to observe persistent overconfidence. This is due to the fact that high-ability agents can be underestimated to a higher extent than low-ability agents. The more an agent's ability can be underestimated, the easier it is for an overconfident manager to rationalize the observed output without having to adapt the overestimation of his own ability. Hence, responsibility hoarding is more likely to be widespread and persistent in workplaces with high ability workers and low accountability of work output than in settings with low ability workers or with a high traceability of exerted work effort.

Responsibility hoarding is also more likely to persist in situations, in which the asymmetry between tasks is relatively low. The more similar tasks are in their impact on total output, the easier it is for the overconfident manager to rationalize the observed total output by underestimating the

[7] For this point see also Palomino and Peyrache (2010).

contribution of the agent. If, in contrast, the task that the agent performs has very little impact on total output, the overconfident manager will find it difficult to rationalize observed low output levels without having to adapt the biased assessment of his own ability.

Note that if there is persistent responsibility hoarding at a workplace, the agent's work satisfaction will most probably decrease over time, due to the continued underestimation of her true ability. Hence, while the biased perception of the overconfident manager motivates him to exert more effort than a rational manager would, it may also cause lower satisfaction levels amongst the agents, leading to more tensions at the workplace and higher turnover rates. Interestingly, a high turnover rate amongst agents may even further support the persistence of the manager's overconfidence, because a constant input of new agents tends to reduce the power of the statistical evidence that would be needed for the overconfident manager to discover his self-perception bias.

Finally, our analysis also implies that allowing employees to choose their tasks may lead to lower degrees of free-riding than predicted, if the employees exhibit some degrees of overconfidence. Especially when the cost of a centrally planned task allocation is high, allowing overconfident employees to volunteer for high-effort tasks may be a cost efficient second-best solution.

3.7 Appendix

3.7.1 Proof of Proposition 3

First, we consider the case if $a_M > \sqrt{2}a_A$. In this case, both a fully rational manager and an overconfident manager will choose non-delegation. This is also the individually optimal task allocation for the manager. The total welfare with a fully rational manager is equal to the sum of the utilities of manager and agent which is given by

$$\begin{aligned} W^{nd*} &= U_M^{nd*} + U_A^{nd*} \\ &= \left(\lambda a_M e_M^{nd*} + (1-\lambda) a_A e_A^{nd*}\right) - \frac{c}{2} e_M^{nd*2} - \frac{c}{2} e_A^{nd*2} \\ &= \frac{3\lambda^2 a_M^2}{8c} + \frac{3(1-\lambda)^2 a_A^2}{8c}. \end{aligned} \tag{3.7}$$

With an overconfident manager it is equal to

$$\begin{aligned} W^{OCnd*} &= U_M^{OCnd*} + U_A^{OCnd*} \\ &= \left(\lambda a_M e_M^{OCnd*} + (1-\lambda) a_A e_A^{OCnd*}\right) - \frac{c}{2} e_M^{OCnd*2} - \frac{c}{2} e_A^{OCnd*2} \\ &= \frac{3\lambda^2 a_M^2}{8c} + \frac{3(1-\lambda)^2 a_A^2}{8c} + \frac{\lambda^2 (2a_M - b) b}{8c}. \end{aligned} \tag{3.8}$$

Comparing (3.7) and (3.8), it follows directly that the total welfare with an overconfident manager is strictly higher than with a fully rational manager if $b < 2a_M$.

Second, we consider the case if $a_M < \sqrt{2}a_A - b$. In this case, both a fully rational manager and an overconfident manager will choose delegation, which is also the individually optimal task allocation for the manager. The total welfare with a fully rational manager is given by

$$\begin{aligned} W^{d*} &= U_M^{d*} + U_A^{d*} \\ &= \left((1-\lambda) a_M e_M^{d*} + \lambda a_A e_A^{d*}\right) - \frac{c}{2} e_M^{d*2} - \frac{c}{2} e_A^{d*2} \\ &= \frac{3(1-\lambda)^2 a_M^2}{8c} + \frac{3\lambda^2 a_A^2}{8c}. \end{aligned} \tag{3.9}$$

With an overconfident manager total welfare is

$$\begin{aligned} W^{OCd*} &= U_M^{OCd*} + U_A^{OCd*} \\ &= \left(\left(1-\lambda\right) a_M e_M^{OCd*} + \lambda a_A e_A^{OCd*}\right) - \frac{c}{2} e_M^{OCd*2} - \frac{c}{2} e_A^{OCd*2} \\ &= \frac{3\left(1-\lambda\right)^2 a_M^2}{8c} + \frac{3\lambda^2 a_A^2}{8c} + \frac{\left(1-\lambda\right)^2 \left(2a_M - b\right) b}{8c}. \end{aligned} \tag{3.10}$$

Again by comparing (3.9) and (3.10), the total welfare with an overconfident manager is strictly higher than with a fully rational manager if $b < 2a_M$.

Finally, we consider the non-trivial case, in which $a_M \leq \sqrt{2}a_A \leq a_M + b$. In this case, a fully rational manager chooses to delegate the critical task to the agent, while an overconfident manager carries out the critical task himself.

Comparing (3.8) and (3.9), the total welfare is higher with an overconfident manager if

$$\left(2\lambda - 1\right)\left(a_M + a_A\right)\left(a_M - a_A\right) + \frac{\lambda^2}{3}\left(2a_M - b\right) b \geq 0.$$

Note that as long as the manager's ability is at least as high as the agent's, i.e. $a_M \geq a_A$, and the manager's degree of overconfidence is on a moderate level, i.e. $b < 2a_M$, the total welfare with an overconfident manager is higher than with a fully rational manager. ■

3.7.2 Proof of Proposition 4

First, we consider the case $a_M > \sqrt{2}a_A - b$, in which an overconfident manager chooses not to delegate the critical task. In this case, the total welfare is given by

$$U^{OCnd*} = \frac{3\lambda^2 a_M^2}{8c} + \frac{3\left(1-\lambda\right)^2 a_A^2}{8c} + \frac{\lambda^2 \left(2a_M - b\right) b}{8c}.$$

By solving the following optimization problem

$$\begin{aligned} &\max_b U^{OCnd*} \\ &s.t.\ b > \sqrt{2}a_A - a_M \end{aligned}$$

we obtain

$$\begin{aligned} & \frac{\partial U^{OCnd*}}{\partial b} \doteq 0 \\ \Leftrightarrow \quad & \frac{\lambda^2}{4c}\left(a_M - b\right) = 0 \\ \Leftrightarrow \quad & b^* = a_M . \end{aligned}$$

Note that the second-order condition is automatically satisfied as U^{OCnd*} is strictly concave in b. Furthermore, the constraint $b > \sqrt{2}a_A - a_M$ is also satisfied as long as $a_M > \frac{\sqrt{2}}{2}a_A$. Hence,

$$b^* = a_M \text{ if } a_M > \frac{\sqrt{2}}{2}a_A$$

Second, we consider the case $a_M \leq \sqrt{2}a_A - b$, in which an overconfident manager chooses to delegate the critical task. In this case, the total welfare is given by

$$U^{OCd*} = \frac{3\left(1-\lambda\right)^2 a_M^2}{8c} + \frac{3\lambda^2 a_A^2}{8c} + \frac{\left(1-\lambda\right)^2 \left(2a_M - b\right) b}{8c}.$$

Again, by solving the following optimization problem

$$\begin{aligned} & \max_b U^{OCd*} \\ & s.t.\ b \leq \sqrt{2}a_A - a_M \end{aligned}$$

we obtain

$$\begin{aligned} & \frac{\partial U^{OCd*}}{\partial b} \doteq 0 \\ \Leftrightarrow \quad & \frac{\left(1-\lambda\right)^2}{4c}\left(a_M - b\right) = 0 \\ \Leftrightarrow \quad & b^* = a_M . \end{aligned}$$

Note that the second-order condition is automatically satisfied as U^{OCd*} is strictly concave in b. Furthermore the constraint $b \leq \sqrt{2}a_A - a_M$ is also satisfied as long as $a_M \leq \frac{\sqrt{2}}{2} a_A$. Hence,

$$b^* = a_M \text{ if } a_M \leq \frac{\sqrt{2}}{2} a_A$$

and the optimal degree of overconfidence is given $b^* = a_M$. ■

3.7.3 Proof of Proposition 6

We prove the results by first analyzing the case of non-delegation and then the case of delegation. We derive the two conditions for sustained overconfidence. We then show that as long as the critical task contributes more to the total output than the other task, i.e. as long as $\frac{1}{2} < \lambda < 1$, the condition stated in the proposition is binding for both delegation and non-delegation situations. Finally, we show that the higher the manager's self-perception bias, the more the agent's ability is underestimated.

1. Persistence of overconfidence in the case of non-delegation

Our essential assumption is that the manager will not revise his assessment of the own ability as long as he observes outcomes that can be rationalized by varying the two unknown parameters, i.e. the agent's ability and effort level. As long as any observed outcome can be rationalized by the manager, overconfidence is persistent. In the following, we derive the sufficient condition for the persistence of overconfidence in case of non-delegation.

Recall that the total output in case of non-delegation is given by

$$Y^{OCnd*} = \lambda a_M e_M^{OCnd*} + (1 - \lambda) \, a_A e_A^{OCnd*}.$$

If an overconfident manager observes this total output, he overestimates his own contribution and underestimates the agent's contribution as follows:

$$Y_P^{OCnd} = \lambda \left(a_M + b \right) e_M^{OCnd*} + (1 - \lambda) \, \widehat{a}_A \widehat{e}_A^{OCnd*}$$

where $\widehat{a}_A$ denotes the estimated value of the agent's ability and $\widehat{e}_A$ the corresponding estimated value of the agent's effort used by the manager.

This biased model (i.e. the overconfident rationalization of the observed output) is only feasible as long as the following condition holds:

$$
\begin{aligned}
& Y^{OCnd*} \geq \lambda \left(a_M + b\right) e_M^{OCnd*} \\
\Leftrightarrow \quad & \lambda a_M e_M^{OCnd*} + \left(1 - \lambda\right) a_A e_A^{OCnd*} \geq \lambda \left(a_M + b\right) e_M^{OCnd*} \\
\Leftrightarrow \quad & a_A \geq \frac{\lambda}{(1-\lambda)} \sqrt{(a_M b + b^2)}.
\end{aligned}
$$

Next, we check for the degree of underestimation of the agent's contribution in case of non-delegation. Let $\tau \equiv a_A - \widehat{a}_A$ denote the underestimation of the agent's ability. Since $Y^{OCnd*} = Y_P^{OCnd}$ we can determine the level of underestimation by solving the following equation:

$$
\begin{aligned}
& Y^{OCnd*} = Y_P^{OCnd} \\
\Leftrightarrow \quad & \lambda a_M e_M^{OCnd*} + \left(1 - \lambda\right) a_A e_A^{OCnd*} = \lambda \left(a_M + b\right) e_M^{OCnd*} + \left(1 - \lambda\right) \widehat{a}_A \widehat{e}_A^{OCnd*} \\
\Leftrightarrow \quad & a_A - \widehat{a}_A = \frac{\lambda^2}{\left(1-\lambda\right)^2} \frac{a_M b + b^2}{\left(a_A + \widehat{a}_A\right)} \\
\Leftrightarrow \quad & \tau = \frac{\lambda^2}{\left(1-\lambda\right)^2} \frac{a_M b + b^2}{\left(a_A + \widehat{a}_A\right)} > 0.
\end{aligned}
$$

Since τ is greater than zero, we have established a positive underestimation of the agent's ability that increases in the manager's self-perception bias b.

2. Persistence of overconfidence in the case of delegation

Now, we derive the sufficient condition for the persistence of overconfidence in case the manager delegates the critical task to the agent. Recall that the total output in the case of delegation is

$$
Y^{OCd*} = \left(1 - \lambda\right) a_M e_M^{OCd*} + \lambda a_A e_A^{OCd*}.
$$

The overconfident manager rationalizes the observation of this output as follows

$$
Y_P^{OCd} = \left(1 - \lambda\right) \left(a_M + b\right) e_M^{OCd*} + \lambda \widehat{a}_A \widehat{e}_A^{OCd*}.
$$

The overestimation of the own contribution (underestimation of the agent's ability) is only feasible as long as the following condition holds:

$$\begin{aligned}
& Y^{OCd*} \geq (1-\lambda)\left(a_M+b\right) e_M^{OCd*} \\
\Leftrightarrow \quad & (1-\lambda)\, a_M e_M^{OCd*} + \lambda a_A e_A^{OCd*} \geq (1-\lambda)\left(a_M+b\right) e_M^{OCd*} \\
\Leftrightarrow \quad & a_A \geq \frac{(1-\lambda)}{\lambda} \sqrt{(a_M b + b^2)}.
\end{aligned}$$

Analogous to the non-delegation case, we check for the degree of underestimation of the agent's contribution in case of delegation. Let $\tau \equiv a_A - \widehat{a}_A$ denote the underestimation of the agent's ability. Since $Y^{d*} = Y_P^d$ we can determine the level of underestimation by solving the following equation:

$$\begin{aligned}
& Y^{OCd*} = Y_P^{OCd} \\
\Leftrightarrow \quad & (1-\lambda)\, a_M e_M^{OCd*} + \lambda a_A e_A^{OCd*} = (1-\lambda)\left(a_M+b\right) e_M^{OCd*} + \lambda \widehat{a}_A \widehat{e}_A^{OCd*} \\
\Leftrightarrow \quad & \tau = \frac{(1-\lambda)^2}{\lambda^2} \frac{a_M b + b^2}{(a_A + \widehat{a}_A)} > 0.
\end{aligned}$$

Again, we find that τ is greater than zero, i.e. the overconfident manager underestimates the agent's ability and the underestimation increases in the manager's self-perception bias b.

3. General conditions for both cases

Taking the results of the two parts together, we can show that for all cases in which the critical task contributes more to the total output than the other task, i.e. as long as $\frac{1}{2} < \lambda < 1$, the condition for persistence in the second case (delegation) is generally more restrictive than in the first case:

$$\frac{\lambda}{1-\lambda} > \frac{1-\lambda}{\lambda}, \text{ for any } \lambda \in \left(\frac{1}{2}, 1\right).$$

Hence, if $a_A \geq \frac{\lambda}{(1-\lambda)} \sqrt{(a_M b + b^2)}$ is true, then the condition $a_A \geq \frac{(1-\lambda)}{\lambda} \sqrt{(a_M b + b^2)}$ also holds, allowing us to use the former as a general condition in the proposition. ■

Chapter 4

Overconfidence, Helping Effort, and Team Performance[1]

4.1 Introduction

Overconfidence is one of the most well-studied cognitive biases in psychology and behavioral economics. It describes a behavioral pattern that involves the overestimation of the own capabilities, especially in tasks with a partially stochastic outcome (see e.g. Soll (1996)). The nature of this phenomenon is, on the one hand, that people usually tend to overestimate the reliability of their knowledge (see e.g. Lichtenstein et al. (1982), Russo and Schoemaker (1992)). On the other hand, people also tend to overestimate their own abilities. In this regard, a famous finding was reported by Svenson (1981) that most car drivers believe that they are safer and more skillful than the average driver.[2]

While the behavioral pattern of overconfidence and its effects on financial decision-making have been extensively studied, the effects of overconfidence on organizational performance are not fully understood yet. Most of the existing literature on overconfidence in managerial settings is focused on the excessive market entry or exaggerated investment risks taken by overconfident

[1] This chapter is based upon Zhou (2011).

[2] For a comprehensive overview of the psychological literature on overconfidence see e.g. Weinberg (2009).

managers (see e.g. Camerer and Lovallo (1999), Malmendier and Tate (2005), Malmendier and Tate (2008), Hackbarth (2008)). In contrast, Palomino and Sadrieh (2011) show that overconfidence doesn't have to be harmful at all as overconfident portfolio managers may exhibit risk attitudes that are more in line with the investors' risk attitudes compared to fully rational risk-averse managers. Analyzing Data of large publicly traded firms from 1980 to 1994, Galasso and Simcoe (2011) find a robust positive association between CEOs' overconfidence and their firms' innovative performance. The reason is that overconfident CEOs underestimate the probability of failure, and thus, are more likely to pursue innovation.

In this chapter, we investigate how agents' overly optimistic self-perception affects the incentives for helping and team performance. For this purpose, we introduce a model in which a manager and an agent can both exert effort into a joint production. Furthermore, the manager also can assist the agent by providing helping effort to his task.[3] We show that overconfident agents generally tend to exert higher effort to the team production than fully rational agents would, even though it is individually suboptimal as they suffer from higher effort costs. Surprisingly, this individually suboptimal behavior is not generally harmful for the agents' utility, and in contrast, can even be advantageous for all involved parties, as long as the agents' biased self-perception is not too strong. However, the positive effect of agent overconfidence crucially depends on the information setting, i.e. whether the managers are aware of the agents' overly optimistic self-assessment. If managers anticipate the agents' biased self-perception, and hence, expect higher effort levels of the agents, they will extend their helping effort respectively, leading to higher team outcome. As a result, both individual utility and total welfare of all involved parties will be higher. Intuitively, this effect results from the fact that overconfident agents overestimate their own productivity and exert higher effort than fully rational agents, due to their higher self-perceived marginal return of effort. This positive incentive effect of overconfidence is further

[3] In this chapter, we use the female pronoun for the manager and the male pronoun for the agent, because males are generally found to exhibit a higher degree of overconfidence than females (see e.g. Barber and Odean (2001), Correll (2001), Bengtsson et al. (2005)).

enhanced by higher levels of helping effort provided by managers due to the complementary production technology. However, if managers are not aware of agents' overconfidence, the latter effect doesn't occur.

Moreover, we also consider the persistence of agent overconfidence and show that it is sustainable, as long as the agent can rationalize his overly optimistic self-perception by underestimating the ability of the manager. The more leeway an overconfident agent has to rationalize the observed outcome without having to adapt his positively biased assessment of the own ability level, the more likely it is to observe persistent individually suboptimal but welfare improving contribution behavior and higher level of cooperation in an organization. However, the persistence of agent overconfidence also crucially depends on the information setting. In particular, agent overconfidence is more likely to persist if the manager is aware of the agent's biased self-perception.

In the recent years, there is a growing number of papers studying the effect of biased self-perception on firm performance. In a principal-agent context, Santos-Pinto (2008) investigates the effects of workers' mistaken beliefs about their abilities and shows that firms may have incentives to hire workers with mistaken beliefs when effort is observable. In particular, firms can take advantage from workers' mistaken beliefs about their coworkers' abilities by using interdependent incentive schemes instead of individualistic ones. In another theoretical framework, Santos-Pinto (2010) studies the effect of positive self-image on workers' productivity in firms where incentives are provided through tournaments and comes up to the conclusion that firms are usually better off if they hire workers which overestimate themselves in tournaments. Moreover, he also shows that a moderate level of overestimation by the workers can lead to higher welfare in the tournament. Similar results are also derived by Nieken et al. (2011) who show that a moderate level of manager overconfidence may lead to more efficient task allocations and improve firm performance as overconfident managers underestimate their effort costs, and hence, take up more responsibilities than fully rational managers would.

In a team production setting with complementary production technology, Gervais and Goldstein (2007) show that biased self-perception is not generally harmful for the team performance. Moreover, the presence of overconfident agents can even improve the coordination amongst the team members and helps to mitigate the free-rider problem. Similar results are derived by Ludwig et al. (2011a) who also show that overconfident agents are better off if they are not aware of other team members' biased self-perception. While our results also indicate the positive effect of agent overconfidence on team performance, we show that, in contrast, all involved parties may be better off if managers are perfectly informed on their subordinates' biased self-perception.

Regarding the incentives for cooperation and helping each other, a well-known result implies that incentive schemes purely based on individual performance may reduce the individuals' willingness to help each other as helping is usually costly and hinders them from working on their own tasks (see e.g. Drago and Turnbull (1988), Lazear (1989), Drago and Garvey (1998), Encinosa et al. (2007), Burks et al. (2009)). By considering a multi-tasking environment where agents can allocate their efforts to various independent tasks, Itoh (1991) shows that it can be beneficial to use reward scheme based on team performance when mutual support is useful. As agents usually align their own efforts to maximize the expected outcome of the task for which they are mainly responsible, and therefore, might provide insufficient helping effort to their co-workers.[4] This also implies that cooperation and helping are only meaningful if each agent increases his own effort responding to an increase in helping from the other agent, and vice versa. In our model, such a complementarity is imbedded between the agent's productive effort and the manager's helping effort.

[4]See also Holmström and Milgrom (1991), Itoh (1992), Drago and Garvey (1998) and Dur and Sol (2010). The empirical evidence on the effects of group incentives on helping on the job is so far rather miscellaneous. Drago and Garvey (1998), for instance, analyze data from the Australian manufacturing industry and find that profit sharing appear to have little positive effect on workers' helping efforts. Using employee data of the German Socio-Economic Panel, Heywood et al. (2005) show that profit sharing may lead workers to increase their coworkers' productivity through greater cooperation which is reflected in better relations among the workers. Berger et al. (2011) use data from an employer-employee matched survey of German companies and find a positive link between team-based compensation schemes and cooperation in teams. In contrast, neither incentives based on individual nor on firm performance affect cooperation among employees.

The remainder of this chapter is organized as follows. The basic model is described in section 4.2. Section 4.3 presents the first-best effort choices and the individually optimal effort choices chosen by manager and agent under different information settings. Section 4.4 analyzes the effect of agent overconfidence on individual utility and total welfare. Section 4.5 considers the persistence of agent overconfidence. Section 4.6 concludes.

4.2 The model

One manager and one agent ($i = M\,[anager]\,, A\,[gent]$) can both exert effort into a joint production. Both individuals are risk neutral. The total output is given by

$$Y = a_M \cdot e_M + a_A \cdot e_A + e_A \cdot h_M$$

where a_i denotes i's true ability, e_i her effort exerted to the team production, and h_M the manager's helping effort providing to the agent (e.g. support or mentoring functions).[5] All effort choices are chosen simultaneously. Furthermore, we assume that only the total output is observable for both manager and agent.

The individual effort costs are described by a convex cost function $C\,(e_i) = \frac{1}{2}e_i^2$. The manager's costs for helping effort (or opportunity costs of helping) are strictly increasing in the level of helping effort, i.e. $\theta\,(h_M) = \frac{c}{2}h_M^2$ with $c > 1$ indicating that helping effort is more costly than productive effort.

For simplicity, we assume that both manager and agent receive the same share of $\frac{1}{2}$ from the total output. Hence, the manager's utility is given by

$$U_M = \frac{1}{2}\left(a_M \cdot e_M + a_A \cdot e_A + e_A \cdot h_M\right) - \frac{1}{2}e_M^2 - \frac{c}{2}h_M^2$$

and the agent's utility is equal to

$$U_A = \frac{1}{2}\left(a_M \cdot e_M + a_A \cdot e_A + e_A \cdot h_M\right) - \frac{1}{2}e_A^2.$$

[5] In the following, we refer e_i as indidivual i's *productive effort.*

Finally, following Gervais and Goldstein (2007) we characterize an *overconfident* agent as someone who has an overly optimistic self-perception, and therefore, systematically overestimates his own ability, i.e.

$$a_A^{OC} = a_A + b$$

where a_A^{OC} denotes the overconfident agent's self-perceived ability and the parameter $b > 0$ his self-perception bias or his degree of overconfidence.

4.3 Equilibrium analysis

In the first step, we assume that the agent's true ability is common knowledge. We start by investigating the first-best case and proceed to the equilibrium with a fully rational agent. In the second step, we derive the equilibrium with an overconfident agent where we differentiate between two possible cases, i.e. whether the manager is aware of the agent's overconfidence or not. Comparing those results, we then show that, at first, the manager generally provides less helping to the fully rational agent than is efficient. Second, if the agent is overconfident, and moreover, if the manager is also aware of the agent's overly optimistic self-perception, she always provides more helping effort relative to the case with an fully rational agent of the same true ability. Furthermore, we prove that the agent's biased self-perception is not generally harmful. From the individual perspective of the manager, it is always advantageous to work with an overconfident agent irrespective of whether she is aware of the agent's overconfidence. Surprisingly, overconfidence is not always harmful for the agent either, and can even be worthwhile if the manager is perfectly informed on the agent's overconfidence and the agent's self-perception bias is on a moderate level relative to his true ability. Moreover from the point of view of a social planner, overconfidence may indeed be beneficial for the total welfare of all involved parties, as long as the increased total output resulted from the overconfident agent's higher effort (and the manager's increased helping effort when she is aware of the agent's overconfidence) over-compensates the agent's loss due to individually suboptimal effort choices. Finally, we also derive the optimal degree of the agent's overconfidence that maximizes the total welfare of all involved parties.

4.3.1 First-best equilibrium

In the first-best case, the total welfare of all involved parties is to be maximized:

$$\max_{e_M, e_A, h_M} W = (a_M \cdot e_M + a_A \cdot e_A + e_A \cdot h_M) - \frac{1}{2}e_M^2 - \frac{c}{2}h_M^2 - \frac{1}{2}e_A^2$$

which leads to the first-best effort levels described by[6]

$$\begin{aligned} e_M^{FB} &= a_M \\ h_M^{FB} &= \frac{a_A}{c-1} \ . \\ e_A^{FB} &= \frac{a_A c}{c-1} \end{aligned} \tag{4.1}$$

Note that the manager's helping effort is strictly increasing in the agent's true ability as the agent's productive effort raises with his ability. Due to the complementarity between the manager's helping effort and the agent's productive effort, the manager is more likely to provide helping effort to a highly productive agent. Furthermore, it is also straightforward to see that the higher the cost of helping is, the lower is the manager's helping effort.

4.3.2 Second-best equilibrium with exogenous compensation contracts and fully rational agent

We now derive the individuals' optimal effort choices in the second-best case with exogenous compensation contracts and a fully rational agent. In difference to the first-best case, the individual's utility is now to be maximized. With perfect information on the agent's ability the maximization problems

[6] See the appendix for a detailed formal derivation.

are given by

$$\begin{aligned}\max_{e_M,h_M} U_M &= \frac{1}{2}(a_M \cdot e_M + a_A \cdot e_A + e_A \cdot h_M) - \frac{1}{2}e_M^2 - \frac{c}{2}h_M^2 \\ s.t.\ e_A &= \arg\max \frac{1}{2}(a_M \cdot e_M + a_A \cdot e_A + e_A \cdot h_M) - \frac{1}{2}e_A^2\end{aligned}$$

and

$$\begin{aligned}\max_{e_A} U_A &= \frac{1}{2}(a_M \cdot e_M + a_A \cdot e_A + e_A \cdot h_M) - \frac{1}{2}e_A^2 \\ s.t.\ h_m &= \arg\max \frac{1}{2}(a_M \cdot e_M + a_A \cdot e_A + e_A \cdot h_M) - \frac{1}{2}e_M^2 - \frac{c}{2}h_M^2\end{aligned}$$

which leads to the optimal effort choices described by[7]

$$\begin{aligned} e_M^{SB} &= \frac{a_M}{2} \\ h_M^{SB} &= \frac{a_A}{4c-1} \quad . \\ e_A^{SB} &= \frac{2a_A c}{4c-1} \end{aligned} \tag{4.2}$$

Comparing (4.1) and (4.2), it is straightforward to see that both the individuals' productive effort and the manager's helping effort are below the first-best levels. The reason is that fully rational individuals cannot commit on the first-best effort choices as they always have an incentive to lower their effort costs by choosing effort levels such that their marginal return of effort matches their marginal costs of effort. Hence, typical free-rider behavior in a team production environment can be observed here.

[7]The derivation of the second-best effort choices follows analogously to the first-best case, unless that in the second-best case the utility functions of the individuals are to maximize.

4.3.3 Third-best equilibrium with exogenous compensation contracts and overconfident agent

In this section, we derive the individuals' optimal effort choices in an equilibrium with an overconfident agent.[8] As the manager's decision on helping effort is crucially affected by the information on agent's overconfidence, we differentiate between two possible cases in the following analysis, i.e. whether the manager is aware of the agent's self-perception bias or not.

Perfect information on agent's self-perception bias

First, we consider the case when the manager has perfect information on all ability and overconfidence parameters. Furthermore, the agent is convinced that his biased self-perception is correct and also shared by the manager. Hence, the individuals' optimization problems are now given by

$$\begin{aligned} \max_{e_M, h_M} U_M &= \frac{1}{2}\left(a_M \cdot e_M + a_A \cdot e_A + e_A \cdot h_M\right) - \frac{1}{2}e_M^2 - \frac{c}{2}h_M^2 \\ s.t.\ e_A &= \arg\max \frac{1}{2}\left(a_M \cdot e_M + a_A^{OC} \cdot e_A + e_A \cdot h_M\right) - \frac{1}{2}e_A^2 \end{aligned}$$

and

$$\begin{aligned} \max_{e_A} U_A &= \frac{1}{2}\left(a_M \cdot e_M + a_A^{OC} \cdot e_A + e_A \cdot h_M\right) - \frac{1}{2}e_A^2 \\ s.t.\ h_m &= \arg\max \frac{1}{2}\left(a_M \cdot e_M + a_A^{OC} \cdot e_A + e_A \cdot h_M\right) - \frac{1}{2}e_M^2 - \frac{c}{2}h_M^2 \end{aligned}$$

[8] We refer an equilibrium in such a setting as a "third-best" equilibrium.

which leads to the optimal effort choices described by[9]

$$\begin{aligned} e_M^{TB} &= \frac{a_M}{2} \\ h_M^{TB} &= \frac{a_A + b}{4c-1} \\ e_A^{TB} &= \frac{2\,(a_A + b)\,c}{4c-1} \end{aligned} \quad . \tag{4.3}$$

Obviously, both the agent's productive effort and the manager's helping effort are strictly increasing in the agent's self-perception bias b, i.e. the more overconfident the agent is, the higher is his productive effort, and in turn, the more helping he receives from the manager. Comparing (4.2) and (4.3), it is also straightforward to see that for any $b > 0$ an overconfident agent's effort is always higher than the effort chosen by a fully rational agent of the same true ability. Therefore, due to the complementarity between the manager's helping effort and the agent's productive effort, the manager always provides more helping to the overconfident agent than to the fully rational one.

Without information on agent's self-perception bias

Now, we consider the case when the manager still has perfect information on all ability parameters, but doesn't know that the agent is overconfident. In contrast to the previous case, the manager believes that the agent is fully rational while the overconfident agent is still convinced that his overly optimistic self-assessment is correct. In particular, the agent also believes that his biased self-perception is shared by the manager (see e.g. Squintani (2006) or Santos-Pinto (2010)). Hence, the individuals' optimization problems are

[9] With perfect information on ability and overconfidence parameters, the equilibrium effort choices can directly derived by substituting a_A^{OC} for a_A in (4.2).

now given by

$$\begin{aligned}\max_{e_M,h_M} U_M &= \frac{1}{2}\left(a_M \cdot e_M + a_A \cdot e_A + e_A \cdot h_M\right) - \frac{1}{2}e_M^2 - \frac{c}{2}h_M^2 \\ s.t.\ e_A &= \arg\max \frac{1}{2}\left(a_M \cdot e_M + a_A \cdot e_A + e_A \cdot h_M\right) - \frac{1}{2}e_A^2\end{aligned}$$

and

$$\begin{aligned}\max_{e_A} U_A &= \frac{1}{2}\left(a_M \cdot e_M + a_A^{OC} \cdot e_A + e_A \cdot h_M\right) - \frac{1}{2}e_A^2 \\ s.t.\ h_m &= \arg\max \frac{1}{2}\left(a_M \cdot e_M + a_A^{OC} \cdot e_A + e_A \cdot h_M\right) - \frac{1}{2}e_M^2 - \frac{c}{2}h_M^2\end{aligned}$$

which leads to the optimal effort choices described by[10]

$$\begin{aligned}\widetilde{e}_M^{TB} &= \frac{a_M}{2} \\ \widetilde{h}_M^{TB} &= \frac{a_A}{4c-1} \\ \widetilde{e}_A^{TB} &= \frac{2\left(a_A + b\right)c}{4c-1}\end{aligned} \quad . \tag{4.4}$$

Like in the previous case, the agent's productive effort is strictly increasing in his self-perception bias b, i.e. the more overconfident he is, the higher his effort choice. Furthermore by comparing (4.2) and (4.4), it is also straightforward to see that overconfident agents always exert higher effort than fully rational agents of the same true ability. However, in contrast to the previous case, the overconfident agent now receives a lower level of helping. The reason is that, now, the manager believes that the agent is fully rational. Hence, she expects a lower level of productive effort of the agent, and in turn, adapts the amount of helping effort respectively. We summarize our findings in the following proposition:

Proposition 1 *For any positive self-perception bias $b > 0$ an overconfident agent always exerts higher productive effort than a fully rational agent of the*

[10] See the appendix for a detailed formal derivation.

same true ability. Furthermore, with perfect information on the agent's true ability a_i and his self-perception bias b the manager provides more helping to the overconfident agent than to the fully rational one. In particular, the higher the overconfident agent's self-perception bias is the more helping he receives. However, if the manager is not aware of the agent's biased self-perception, both overconfident and fully rational agents receive the same level of helping.

As we have shown in the last section, fully rational agents cannot commit on the first-best effort levels as they have an incentive to lower their effort costs by choosing lower effort levels than is efficient. In contrast, overconfident agents overestimate their own ability, and therefore, choose a higher effort level than fully rational agents of the same true ability. As a result, the effort choices chosen by overconfident agents may be closer to the efficient level than those of fully rational agents. Furthermore, due to the complementarity in the production technology the manager also exerts higher helping effort when she is aware of the agent's overconfidence. Hence, overconfidence can be considered as a commitment device for all individuals to exert higher efforts reducing free-riding, and thus, positively affects the total welfare of all involved parties. However, the positive output effect of agent overconfidence crucially depends on the manager's awareness of the agent's biased self-perception. Otherwise, the manager would not adapt her helping effort, and correspondingly, the higher total output only results from the agent's higher productive effort.

4.4 Is overconfidence beneficial or harmful?

As we have shown above, overconfidence can lead to higher productive effort and increase the agent's contribution to the total output. Moreover, due to the complementarity of productive and helping efforts the manager also provides more helping to the overconfident agent when she is aware of the agent's biased self-perception. As a result, the total output is higher with an overconfident agent than with a fully rational agent. As both manager and agent benefit to the same extent from the total output, their utilities may also be higher. However, as the overconfident agent also suffers from higher effort costs causing by higher productive effort, it is not clear whether the agent's overconfidence is generally beneficial or harmful with regard to the individual utilities and the total welfare of all involved parties. In this section, we conjecture and prove that in many cases, a moderate level of agent overconfidence is beneficial for all involved parties both individually and with respect to the total welfare, despite of the overconfident agent's individually suboptimal effort choice.

4.4.1 Effects of agent overconfidence on individual utility

We start our utility analysis by considering the manager's individual utility. By comparing the different values of the manager's utility in equilibrium with fully rational and overconfident agents we derive the following proposition:

Proposition 2 *The manager is always strictly better off when working with an overconfident agent than with a fully rational agent of the same true ability. In particular, the manager's utility in equilibrium with an overconfident agent is at highest if she has perfect information on the agent's self-perception bias.*

Proof. See the appendix.

This result is illustrated in Figure 4.1. The solid (dotted) line shows the manager's utility in the third-best equilibrium with (without) awareness of the agent's biased self-perception as a correspondence of b, and the dashed line the manager's utility in the second-best equilibrium.[11] It is straightforward to see that the manager's utility in equilibrium with an overconfident agent is strictly increasing in the agent's self-perception bias b, and in particular, always above the utility level in equilibrium with a fully rational agent of the same true ability. Furthermore, the utility in a third-best equilibrium is always strictly higher if the manager is perfectly informed on the agent's

[11] The figure shows a setting in which $a_M = 10$, $a_A = 5$, and $c = 1.2$.

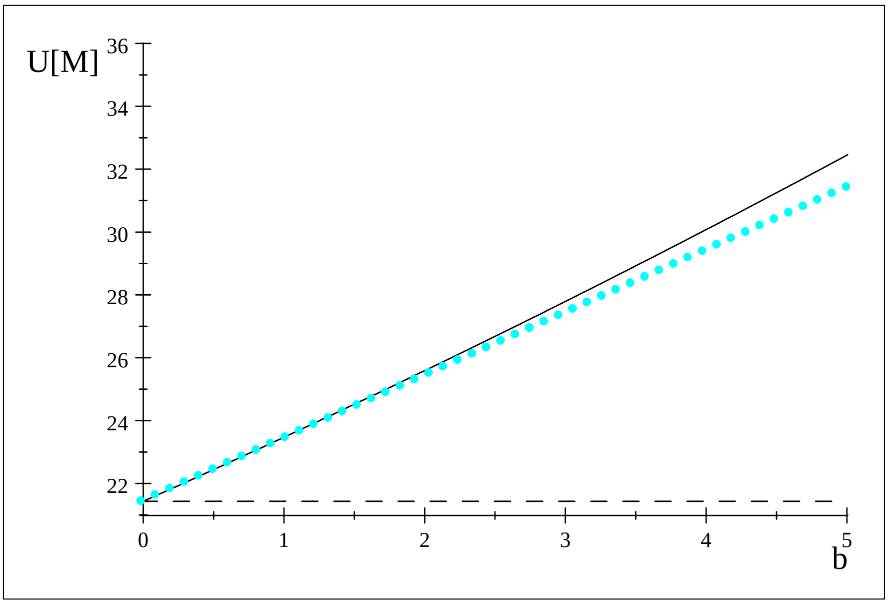

Figure 4.1: Manager's utility in equilibrium depending on b

overconfidence parameter. The reason for this result is that the overconfident agent overestimates his own ability, and hence, increases the total output by exerting higher effort to the team production. As the manager directly benefits from the higher total output, and furthermore, as the higher effort costs are exclusively born by the agent, the manager's utility increases regardless of whether she knows that the agent is overconfident. Moreover, in the case when the manager is aware of the agent's biased self-perception she can further adapt her helping effort, leading to higher total output.

Now, we proceed the utility analysis by comparing the agent's individual utility in equilibrium depending on whether he is fully rational or overconfident, and in the latter case, whether the manager is aware of his biased self-perception. The results concerning the agent's utility is somewhat miscellaneous. We summarize them in the following proposition:

Proposition 3 *With perfect information on ability and overconfidence parameters for the manager, the overconfident agent is always better off than the fully rational agent of the same true ability if his self-perception bias is on a moderate level, i.e. $b \leq \frac{a_A}{2c-1}$. If the manager is not aware of the agent's bi-*

ased self-perception, the overconfident agent is always worse off than the fully rational agent of the same true ability. Moreover irrespective of his degree of overconfidence, the overconfident agent is always worse off if the manager is not aware of his overconfidence.

Proof. See the appendix.

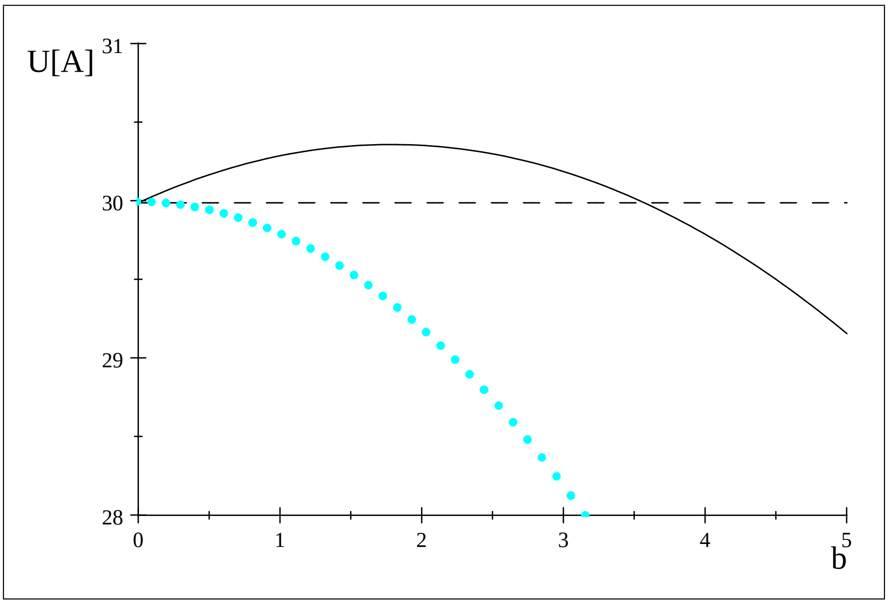

Figure 4.2: Agent's utility in equilibrium depending on b

Figure 4.2 depicts the agent's utility in equilibrium. The solid (dotted) line shows the agent's utility in third-best equilibrium if the manager is (not) aware of his overconfidence parameter as a correspondence of b and the dashed line the utility of a fully rational agent of the same true ability.[12] Note that the overconfident agent's utility without the manager's awareness of his biased self-perception is strictly declining in b, and in particular, always below the utility of the fully rational agent. In contrast, there is a range of overconfidence degrees in which the overconfident agent is strictly better off than the fully rational agent given the awareness of the manager. The intuition behind this result is as following: First, note that the overconfident agent always suffers from higher effort costs as he overestimates his

[12]The figure shows a setting in which $a_M = 10$, $a_A = 5$, and $c = 1.2$.

marginal return of effort, and therefore, exerts higher effort that is individually suboptimal irrespective of whether the manager is aware of his biased self-perception. Second, in case when the manager has perfect information on ability and overconfidence parameters she would increase her helping effort, leading to higher total output. As the agent benefits from higher total output, his higher effort costs can even be over-compensated. As long as the agent's self-perception bias is not too large, the positive output effect outweighs the negative cost effect. Furthermore, note that the upper bound for the agent's self-perception bias $\frac{a_A}{2c-1}$ is strictly increasing in the agent's true ability a_A and decreasing in the cost parameter for helping effort c, i.e. the higher (lower) the agent's true ability (the cost for helping) is, the more overconfident he might be without to be disadvantaged. However, if the manager is not aware of the agent's overconfidence, she would not adapt her helping effort, and hence, the overconfident agent will always be worse off compared to the fully rational agent of the same true ability.

4.4.2 Effects of agent overconfidence on total welfare

The previous analysis demonstrates that it is always advantageous for the manager to work with an overconfident agent. Furthermore, the agent can also benefit from the biased assessment of his own ability as long as his self-perception bias is not too strong, and moreover, the manager is also perfectly informed on this self-perception bias. In this section, we conjecture and show that a moderate level of overconfidence can also be beneficial with regard to the total welfare of all involved parties, even when the manager is not aware of the agent's overconfidence.

Analogously to the previous section, we compare the values of total welfare in equilibrium with fully rational and overconfident agents taking into account whether the manager is aware of the agent's overconfidence. The results are summarized in the following proposition:

Proposition 4 *With perfect information on ability and overconfidence parameters for the manager, the total welfare of all involved parties with an overconfident agent is higher than with a fully rational agent of the same true* *ability if the overconfident agent's self-perception bias is on a moderate level, i.e.* $b \leq \frac{8c+2}{4c-3}a_A$*. However, when the manager is not aware of the agent's biased self-perception the total welfare is still higher if the overconfident agent's degree of overconfidence is bounded by* $b \leq 2a_A$*. Moreover, the total welfare in equilibrium with an overconfident agent is always higher if the manager has perfect information on the agent's overconfidence parameter irrespective of his degree of overconfidence* b*.*

Proof. See the appendix.

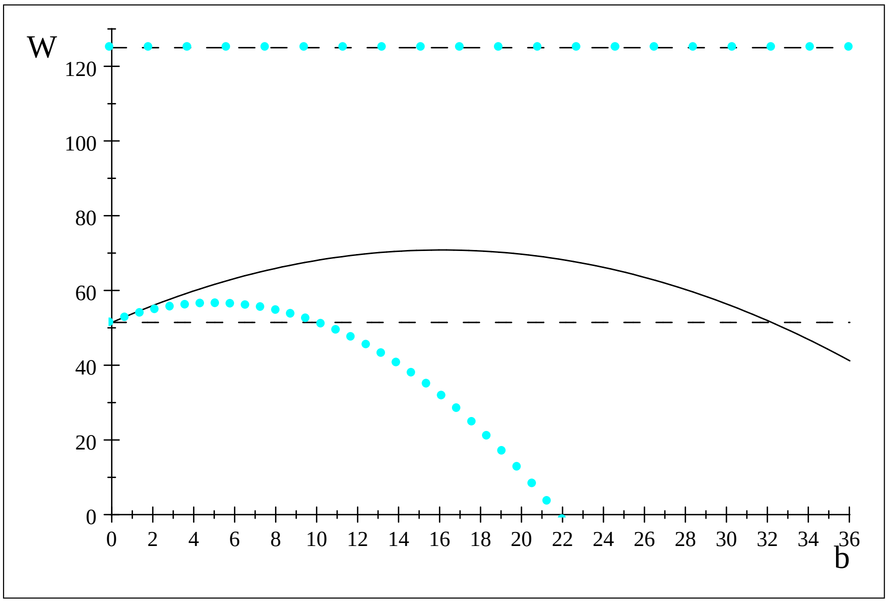

Figure 4.3: Total welfare in equilibrium depending on b

This result is illustrated in Figure 4.3. The solid (dotted) line shows the total welfare in third-best equilibrium when the manager is (not) aware of the agent's biased self-perception as a correspondence of b and the dashed line the total welfare in second-best equilibrium with a fully rational agent of the same true ability. Furthermore, the dot-dashed line depicts the first-best level of total welfare.[13] In both cases whether the manager is aware of the agent's degree of overconfidence, there is a range of self-perception bias in

[13]The figure shows a setting in which $a_M = 10$, $a_A = 5$, and $c = 1.2$.

which the total welfare in the third-best case is above the second-best level, and thus, closer to the first-best level. Furthermore, this result also has several interesting implications: First, note that the agent's overconfidence is not generally harmful and can even be beneficial for the total welfare if it is not too strong. This is due to the fact that the overconfident agent overestimates his own ability, and therefore, exerts higher effort than the fully rational agent of the same true ability, leading to higher total output that benefits all involved parties. With perfect information on overconfidence parameter it also leads to higher helping effort by the manager. Hence, agent overconfidence helps to reduce free-riding. Surprisingly, this positive incentive effect of overconfidence can even over-compensate the negative effect of individually suboptimal effort choice as long as the agent's true ability is sufficiently high relative to his self-perception bias. In particular, the total welfare in equilibrium with an overconfident agent can even be closer to the first-best level than with a fully rational agent of the same true ability.

4.4.3 Optimal degree of agent overconfidence

As depicted in Figure 4.3, both welfare functions of the third-best cases have an inverse U-shaped curve. Therefore, we proceed our analysis with the determination of their local maximum, i.e. the optimal degree of agent overconfidence with respect to the total welfare of all involved parties. In this regard, we can show:

Proposition 5 *With perfect information on ability and overconfidence parameters for the manager, the total welfare of all involved parties is highest if* $b^* = \frac{4c+1}{4c-3}a_A$. *If the manager is not aware of the agent's biased self-perception, the total welfare takes its maximum if the agent's degree of overconfidence is equal to his true ability, i.e.* $\widetilde{b}^* = a_A$.

Proof. See the appendix.

Note that in both cases, the positive welfare effect of agent overconfidence is strictly increasing in the agent's true ability which indicates that the higher the agent's true ability is, the less harmful is his biased self-perception. Interestingly, the optimal degree of overconfidence is equal to the true ability of the agent in the latter case such that the equilibrium effort of the overconfident agent exactly matches its first-best level.

4.5 Persistence of agent overconfidence

In the previous two sections, we have demonstrated that the moderate agent overconfidence can lead to higher effort resulting in higher total welfare for all involved parties. The question that remains to be answered is whether the agent's overconfidence and higher level of contribution and cooperation can persist over time, given the feedback the agent receives from his previous decisions. If the overconfident agent quickly learns to correct his overconfident assessment of the own ability, then overconfidence and higher contribution to team output will not persist. However, if the feedback from previous outcomes cannot be used to correct the biased self-perception, overconfidence can be a persistent phenomenon with a sustained effect on organizational performance.

In this section, we derive the conditions under which overconfidence can persist, even though the agent receives feedback on his previous decisions. We restrict our analysis to the case that the agent only receives feedback on the total output of the firm (or the organizational unit). Obviously, persistence of overconfidence with more exact information, e.g. on all ability and effort parameters, would not be feasible. In the more realistic situation that we analyze, the feedback information is restricted to observing the total output, so that the overconfident agent faces one known parameter (his productive effort), two unknown parameters (the manager's ability and productive effort), and two parameters that he believes to know, but actually does not (his true ability and the manager's helping effort). Under these circumstances, we show that the overconfident agent may not be able to learn that his self-assessment is biased, because he can construct a consistent model that explains the observed total outcome with an overestimated own ability parameter and an underestimated ability parameter for the manager.[14] As

[14] Young (2002) shows that some games cannot be learned by rational players and demonstrates a class of learning environments in which convergence to equilibrium behavior fails

long as the productivity of the manager can be sufficiently underestimated, the agent's overconfidence can persist.

We summarize this result in the following proposition:

Proposition 6 *If* $a_M \geq \frac{4c}{4c-1}\sqrt{(a_A + b)\,b}$, *i.e. the manager's true ability is sufficiently high, the agent's overconfidence persists irrespective of whether the manager is aware of the agent's biased self-perception, because the agent rationalizes the observed outcome information by underestimating the manager's ability. The higher the agent's self-perception bias b is, the greater the underestimation of the manager's ability will be.*

Proof. *1. Persistence of overconfidence when the manager is aware of the agent's biased self-perception*

Our essential assumption is that the agent will not revise his biased assessment of the own ability as long as he observes outcomes that can be rationalized by varying the two unknown parameters, i.e. the manager's ability and productive effort. As long as any observed outcome can be rationalized by the agent, overconfidence is persistent.

First, we derive the sufficient condition for the persistence of overconfidence when the manager is perfectly informed on the agent's overconfidence parameter. Recall that the total output in the case with awareness of the manager is given by

$$Y^{TB} = a_M \cdot e_M^{TB} + a_A \cdot e_A^{TB} + e_A^{TB} \cdot h_M^{TB}.$$

However, the overconfident agent who is convinced that his self-perceived ability is true always overestimates his own contribution to the total output $(a_A \cdot e_A + e_A \cdot h_M)$, and consequently, underestimates the contribution of the manager. In particular, when the overconfident agent observes this total output, he rationalizes the composition of the total output as follows:

$$E_A\left[Y^{TB}\right] = \widehat{a}_M \cdot \widehat{e}_M^{TB} + a_A^{OC} \cdot e_A^{TB} + e_A^{TB} \cdot h_M^{TB}.$$

to occur for any learning process, including the Bayesian updating of objectively correct priors.

where $\widehat{a}_M$ and $\widehat{e}_M^{TB}$ denote the values of the manager's ability and productive effort estimated by the overconfident agent.

This biased model (i.e. the overconfident rationalization of the observed output by underestimating the manager's contribution) is only feasible as long as the actual value of the total output is at least as high as the overconfident agent's estimation of his own contribution, i.e. if the following condition holds:

$$\begin{aligned} & Y^{TB} \geq a_A^{OC} \cdot e_A^{TB} + e_A^{TB} \cdot h_M^{TB} \\ \Leftrightarrow \quad & a_M \cdot \frac{a_M}{2} + a_A \cdot \frac{2a_A^{OC}c}{4c-1} + \frac{2a_A^{OC}c}{4c-1} \cdot \frac{a_A^{OC}}{4c-1} \geq a_A^{OC} \cdot \frac{2a_A^{OC}c}{4c-1} + \frac{2a_A^{OC}c}{4c-1} \cdot \frac{a_A^{OC}}{4c-1} \\ \Leftrightarrow \quad & a_M \geq 2\sqrt{\frac{(a_A+b)bc}{4c-1}}. \end{aligned}$$

Next, we check for the degree of underestimation of the manager's contribution. Let $\tau \equiv a_M - \widehat{a}_M$ denote the underestimation of the manager's ability. As $Y^{TB} = E_A\left[Y^{TB}\right]$ we can determine the level of underestimation by solving the following equation:

$$\begin{aligned} & Y^{TB} = E_A\left[Y^{TB}\right] \\ \Leftrightarrow \quad & \frac{a_M^2}{2} = \frac{\widehat{a}_M^2}{2} + \frac{2(a_A+b)bc}{4c-1} \\ \Leftrightarrow \quad & \tau = \frac{4(a_A+b)bc}{(4c-1)(a_M+\widehat{a}_M)} > 0. \end{aligned}$$

Since τ is always strictly larger than zero, we have established a positive underestimation of the manager's ability that increases in the agent's self-perception bias b, i.e. the larger the agent's self-perception bias the greater the underestimation of the manager's ability by the overconfident agent.

2. Persistence of overconfidence when the manager is not aware of the agent's biased self-perception

Now, we derive the sufficient condition for the persistence of overconfidence when the manager is not aware of the agent's biased self-perception. Recall that the total output in this case is given by

$$\widetilde{Y}^{TB} = a_M \cdot \widetilde{e}_M^{TB} + a_A \cdot \widetilde{e}_A^{TB} + \widetilde{e}_A^{TB} \cdot \widetilde{h}_M^{TB}.$$

Like in the previous case, the overconfident agent rationalizes the observation of this output as follows:

$$E_A\left[Y^{TB}\right] = \widehat{a}_M \cdot \widehat{e}_M^{TB} + a_A^{OC} \cdot \widetilde{e}_A^{TB} + \widetilde{e}_A^{TB} \cdot h_M^{TB}$$

where $\widehat{a}_M$ and $\widehat{e}_M^{TB}$ denote the values of the manager's ability and productive effort estimated by the overconfident agent. Note that the level of helping effort expected by the agent is given by $\widehat{h}_M^{TB} = h_M^{TB}$ as the overconfident agent is convinced of his biased self-perception, and moreover, also assumes that his self-perception is shared by the manager.

However, the overestimation of the own contribution by underestimating the manager's ability is only feasible as long as the following condition holds:

$$\begin{aligned} & \widetilde{Y}^{TB} \geq a_A^{OC} \cdot \widetilde{e}_A^{TB} + \widetilde{e}_A^{TB} \cdot h_M^{TB} \\ \Leftrightarrow \quad & a_M \cdot \tfrac{a_M}{2} + a_A \cdot \tfrac{2a_A^{OC}c}{4c-1} + \tfrac{2a_A^{OC}c}{4c-1} \cdot \tfrac{a_A}{4c-1} \geq a_A^{OC} \cdot \tfrac{2a_A^{OC}c}{4c-1} + \tfrac{2a_A^{OC}c}{4c-1} \cdot \tfrac{a_A^{OC}}{4c-1} \\ \Leftrightarrow \quad & a_M \geq \tfrac{4c}{4c-1}\sqrt{(a_A + b)\, b}. \end{aligned}$$

Analogously to the case with awareness of the manager, we check for the degree of underestimation of the manager's ability. Let $\tau \equiv a_M - \widehat{a}_M$ denote the underestimation of the manager's ability. As $\widetilde{Y}^{TB} = E_A\left[Y^{TB}\right]$ we can determine the level of underestimation by solving the following equation:

$$\begin{aligned} & \widetilde{Y}^{TB} = E_A\left[Y^{TB}\right] \\ \Leftrightarrow \quad & \tfrac{a_M^2}{2} = \tfrac{\widehat{a}_M^2}{2} + \tfrac{8(a_A+b)bc^2}{(4c-1)^2} \\ \Leftrightarrow \quad & \tau = \tfrac{16(a_A+b)bc^2}{(4c-1)^2(a_M+\widehat{a}_M)} > 0. \end{aligned}$$

Again, we find that τ is strictly larger than zero, i.e. the overconfident agent always underestimates the manager's ability. In particular, the underestimation increases in the agent's self-perception bias b.

3. General conditions for both cases

Taking the results of the two parts together, we can show that for any positive self-perception bias $b > 0$, the condition for persistence in the second case (without awareness of the agent's biased self-perception) is generally more restrictive than in the first case as

$$\frac{4c}{4c-1}\sqrt{(a_A+b)\,b} > 2\sqrt{\frac{(a_A+b)bc}{4c-1}}.$$

Hence, if $a_M \geq \frac{4c}{4c-1}\sqrt{(a_A+b)\,b}$ is true, then the condition $a_M \geq 2\sqrt{\frac{(a_A+b)bc}{4c-1}}$ also holds, allowing us to use the former as a general condition in the proposition. ■

There are two corollaries following directly from the proposition above:

Corollary 1 *No (positive) level of agent overconfidence is persistent if* $a_M = 0$.

Corollary 2 *Agent overconfidence is more likely to persist if the manager is aware of the overconfident agent's self-perception bias.*

Corollary 1 simply points out that there is always some level of overconfidence that is persistent, unless the manager's ability is zero. Intuitively, it is clear that overconfidence can only persist, as long as the agent has the possibility to justify his overestimated own ability by underestimating the ability of the manager. The range of possible underestimation drops if the manager's ability decreases, leaving less and less room for the overconfident agent. If the manager's ability is zero, persistent overconfidence is no longer feasible. However, note that the extreme case of zero ability has no empirical relevance, because it describes a situation in which the manager cannot contribute anything to the firm's output by own productive effort. Hence, the corollary implies that for any situation with empirical relevance, there is at least some level of persistent overconfidence, leading to higher contribution of overconfident agents. Corollary 2 indicates the fact that persistent overconfidence is more likely to be observed, if the manager also knows that the agent is overconfident. When the manager anticipates the agent's higher effort as a consequence of his biased self-perception, she can adapt her helping effort, respectively. This leads to higher total output which, in turn, allows more room for the overconfident agent to underestimate the manager's contribution. Hence, the persistence of agent overconfidence also crucially depends on the information setting.

4.6 Conclusion

We analyzed the effects of agents' overconfidence on organizational performance. We have shown that overconfident agents usually overestimate their abilities, and therefore, exert higher effort to the team production and increase the total output, even though their effort choices are not individually optimal. Although overconfidence may negatively affect the agents' individual utility, moderately overconfident agents generally tend to increase the firm's output and the total welfare of all involved parties above the level that fully rational agents would achieve. This seemingly counter-intuitive finding results, on the one hand, from higher level of effort exerted by overconfident agents, and on the other hand, also from higher level of helping effort chosen by managers due to the complementarity of agents' productive effort and managers' helping effort. However, the effect of overconfidence on individuals' utility and total welfare crucially depends on the information setting, i.e. whether managers are aware of the agents' biased self-perception. When managers know that the agents are overconfident, they can adapt the level of their helping effort, and thus, increase the total output. In this regard, agent overconfidence can be considered as a commitment device reducing free-riding and increasing the managers' incentives to provide more supervising or monitoring functions, leading to a higher level of cooperation in teams. If managers are not aware of the agent's biased self-perception, the positive output effect only results from the overconfident agents' higher productive effort. Hence, our results imply that firms should not generally avoid hiring overconfident agents. In contrast, the employment of overconfident agents may even be advantageous for the firm's performance and profit. However, to fully exploit the advantages from the employment of overconfident agents, managers should have well-founded knowledge on agents' biased self-perception.

Furthermore, we have shown that agent overconfidence can survive in an organization, as long as the overconfident agents can rationalize their overly optimistic estimation of the own ability by underestimating the ability of the managers. Interestingly, the persistence of overconfidence does not only depend on the agents' degree of overconfidence, but also on the absolute level of the managers' true abilities. In particular, the higher the ability levels in a workplace, the more likely it is to observe persistent agent overconfidence. This is due to the fact that high-ability managers can be underestimated to a greater extent than low-ability managers. The more a manager's ability can be underestimated, the easier it is for an overconfident agent to rationalize the observed output without having to adapt the overestimation of his own ability. Moreover, agent overconfidence has a greater chance to survive when managers are aware of the agents' biased self-perception. The reason is that managers may prefer to work with overconfident agents because of their higher contribution to the team production. For this reason, managers may have an interest to sustain agent overconfidence by exerting higher level of helping effort. Consequently, the total outcome is higher which, in turn, provides the overconfident agents more room to rationalize their biased self-assessment resulting in higher probability for persistent agent overconfidence.

4.7 Appendix

4.7.1 Derivation of the first-best effort choices

In the first-best case the total welfare of all involved parties should be maximized:

$$\max_{e_M, e_A, h_M} W = (a_M \cdot e_M + a_A \cdot e_A + e_A \cdot h_M) - \frac{1}{2}e_M^2 - \frac{c}{2}h_M^2 - \frac{1}{2}e_A^2.$$

The first order conditions are given by

$$\frac{\partial W}{\partial e_M} \doteq 0 \Leftrightarrow a_M - e_M = 0 \Leftrightarrow e_M = a_M \tag{4.5}$$

$$\frac{\partial W}{\partial h_M} \doteq 0 \Leftrightarrow e_A - ch_M = 0 \Leftrightarrow h_M = \frac{e_A}{c} \tag{4.6}$$

$$\frac{\partial W}{\partial e_A} \doteq 0 \Leftrightarrow a_A + h_M - e_A = 0 \Leftrightarrow e_A = a_A + h_M \tag{4.7}$$

The first-best effort follows directly by inserting (4.6) into (4.7) and some algebraic transformation. ■

4.7.2 Derivation of the third-best effort choices without information on the agent's overconfidence parameter

Now, the manager believes that the agent is fully rational while the overconfident agent is still convinced of his overly optimistic self-perception. Moreover, the agent is convinced that his self-perception is shared by the manager. Hence, the individuals' optimization problems are now given by

$$\begin{aligned}
\max_{e_M, h_M} U_M &= \frac{1}{2}(a_M \cdot e_M + a_A \cdot e_A + e_A \cdot h_M) - \frac{1}{2}e_M^2 - \frac{c}{2}h_M^2 \\
s.t.\ e_A &= \arg\max \frac{1}{2}(a_M \cdot e_M + a_A \cdot e_A + e_A \cdot h_M) - \frac{1}{2}e_A^2
\end{aligned}$$

and

$$
\begin{aligned}
\max_{e_A} U_A &= \frac{1}{2}\left(a_M \cdot e_M + a_A^{OC} \cdot e_A + e_A \cdot h_M\right) - \frac{1}{2}e_A^2 \\
s.t.\ h_m &= \arg\max \frac{1}{2}\left(a_M \cdot e_M + a_A^{OC} \cdot e_A + e_A \cdot h_M\right) - \frac{1}{2}e_M^2 - \frac{c}{2}h_M^2.
\end{aligned}
$$

Let U_i^j denotes the maximization problem of individual i $(= M, A)$ that is expected by individual j $(= M, A)$. The first order conditions are now given by

$$\frac{\partial U_M^M}{\partial e_M} \doteq 0 \Leftrightarrow \frac{1}{2}a_M - e_M = 0 \Leftrightarrow e_M = \frac{1}{2}a_M \tag{4.8}$$

$$\frac{\partial U_M^M}{\partial h_M} \doteq 0 \Leftrightarrow \frac{1}{2}e_A - ch_M = 0 \Leftrightarrow h_M = \frac{e_A}{2c} \tag{4.9}$$

$$\frac{\partial U_M^A}{\partial h_M} \doteq 0 \Leftrightarrow \frac{1}{2}e_A - ch_M = 0 \Leftrightarrow h_M = \frac{e_A}{2c} \tag{4.10}$$

$$\frac{\partial U_A^M}{\partial e_A} \doteq 0 \Leftrightarrow \frac{1}{2}(a_A + h_M) - e_A = 0 \Leftrightarrow e_A = \frac{1}{2}(a_A + h_M) \tag{4.11}$$

$$\frac{\partial U_A^A}{\partial e_A} \doteq 0 \Leftrightarrow \frac{1}{2}\left(a_A^{OC} + h_M\right) - e_A = 0 \Leftrightarrow e_A = \frac{1}{2}\left(a_A^{OC} + h_M\right) \tag{4.12}$$

The effort levels described by (4.4) can be derived by inserting (4.9) into (4.11) and (4.10) into (4.12) and some algebraic transformation, respectively. ■

4.7.3 Proof of Proposition 2

The manager's utility in equilibrium with a fully rational agent is given by

$$
\begin{aligned}
U_M^{SB} &= \tfrac{1}{2}\left(a_M \cdot e_M^{SB} + a_A \cdot e_A^{SB} + e_A^{SB} \cdot h_M^{SB}\right) - \tfrac{1}{2}e_M^{SB2} - \tfrac{c}{2}h_M^{SB2} \\
&= \tfrac{1}{2}\left(a_M \cdot \tfrac{a_M}{2} + a_A \cdot \tfrac{2a_Ac}{4c-1} + \tfrac{2a_Ac}{4c-1} \cdot \tfrac{a_A}{4c-1}\right) - \tfrac{1}{2}\tfrac{a_M^2}{4} - \tfrac{c}{2}\left(\tfrac{a_A}{4c-1}\right)^2 \\
&= \tfrac{a_M^2}{8} + \tfrac{(8c-1)a_A^2c}{2(4c-1)^2}.
\end{aligned}
$$

With an overconfident agent, and further, given that the manager is also aware of the agent's self-perception bias b, her utility in equilibrium is described by

$$\begin{aligned}
U_M^{TB} &= \tfrac{1}{2}\left(a_M \cdot e_M^{TB} + a_A \cdot e_A^{TB} + e_A^{TB} \cdot h_M^{TB}\right) - \tfrac{1}{2}e_M^{TB2} - \tfrac{c}{2}h_M^{TB2} \\
&= \tfrac{1}{2}\left(a_M \cdot \tfrac{a_M}{2} + a_A \cdot \tfrac{2a_A^{OC}c}{4c-1} + \tfrac{2a_A^{OC}c}{4c-1} \cdot \tfrac{a_A^{OC}}{4c-1}\right) - \tfrac{1}{2}\tfrac{a_M^2}{4} - \tfrac{c}{2}\left(\tfrac{a_A^{OC}}{4c-1}\right)^2 \\
&= \tfrac{a_M^2}{8} + \tfrac{(8a_Ac-a_A+b)(a_A+b)c}{2(4c-1)^2}.
\end{aligned}$$

Finally, with an overconfident agent, and given that the manager is not aware of the agent's self-perception bias, her utility in equilibrium is equal to

$$\begin{aligned}
\widetilde{U}_M^{TB} &= \tfrac{1}{2}\left(a_M \cdot \widetilde{e}_M^{TB} + a_A \cdot \widetilde{e}_A^{TB} + \widetilde{e}_A^{TB} \cdot \widetilde{h}_M^{TB}\right) - \tfrac{1}{2}\widetilde{e}_M^{TB2} - \tfrac{c}{2}\widetilde{h}_M^{TB2} \\
&= \tfrac{1}{2}\left(a_M \cdot \tfrac{a_M}{2} + a_A \cdot \tfrac{2a_A^{OC}c}{4c-1} + \tfrac{2a_A^{OC}c}{4c-1} \cdot \tfrac{a_A}{4c-1}\right) - \tfrac{1}{2}\tfrac{a_M^2}{4} - \tfrac{c}{2}\left(\tfrac{a_A}{4c-1}\right)^2 \\
&= \tfrac{a_M^2}{8} + \tfrac{(8a_Ac-a_A+8bc)a_Ac}{2(4c-1)^2}.
\end{aligned}$$

It is straightforward to see that $U_M^{TB} > U_M^{SB}$ and $\widetilde{U}_M^{TB} > U_M^{SB}$ for any strictly positive self-perception bias b. Hence, the manager is always better off with an overconfident agent irrespective of whether she is aware of the agent's overconfidence.

Finally, it is easy to show that the manager's utility in equilibrium with an overconfident agent is always strictly higher if she is aware of the agent's self-perception bias as

$$U_M^{TB} - \widetilde{U}_M^{TB} = \tfrac{b^2c}{2(4c-1)^2} > 0. \blacksquare$$

4.7.4 Proof of Proposition 3

The fully rational agent's utility in equilibrium is given by

$$\begin{aligned}
U_A^{SB} &= \tfrac{1}{2}\left(a_M \cdot e_M^{SB} + a_A \cdot e_A^{SB} + e_A^{SB} \cdot h_M^{SB}\right) - \tfrac{1}{2}e_A^{SB2} \\
&= \tfrac{1}{2}\left(a_M \cdot \tfrac{a_M}{2} + a_A \cdot \tfrac{2a_Ac}{4c-1} + \tfrac{2a_Ac}{4c-1} \cdot \tfrac{a_A}{4c-1}\right) - \tfrac{1}{2}\left(\tfrac{2a_Ac}{4c-1}\right)^2 \\
&= \tfrac{a_M^2}{4} + \tfrac{2a_A^2c^2}{(4c-1)^2}.
\end{aligned}$$

With perfect information on ability and overconfidence parameters for the manager, the overconfident agent's utility in equilibrium is described by

$$\begin{aligned} U_A^{TB} &= \tfrac{1}{2}\left(a_M \cdot e_M^{TB} + a_A \cdot e_A^{TB} + e_A^{TB} \cdot h_M^{TB}\right) - \tfrac{1}{2} e_A^{TB2} \\ &= \tfrac{1}{2}\left(a_M \cdot \tfrac{a_M}{2} + a_A \cdot \tfrac{2a_A^{OC}c}{4c-1} + \tfrac{2a_A^{OC}c}{4c-1} \cdot \tfrac{a_A^{OC}}{4c-1}\right) - \tfrac{1}{2}\left(\tfrac{2a_A^{OC}c}{4c-1}\right)^2 \\ &= \tfrac{a_M^2}{4} + \tfrac{(2a_Ac-2bc+b)}{(4c-1)^2}\left(a_A + b\right)c. \end{aligned}$$

Finally, if the manager is not aware of the agent's overconfidence, the overconfident agent's utility in equilibrium is equal to

$$\begin{aligned} \widetilde{U}_A^{TB} &= \tfrac{1}{2}\left(a_M \cdot \widetilde{e}_M^{TB} + a_A \cdot \widetilde{e}_A^{TB} + \widetilde{e}_A^{TB} \cdot \widetilde{h}_M^{TB}\right) - \tfrac{1}{2} \widetilde{e}_A^{TB2} \\ &= \tfrac{1}{2}\left(a_M \cdot \tfrac{a_M}{2} + a_A \cdot \tfrac{2a_A^{OC}c}{4c-1} + \tfrac{2a_A^{OC}c}{4c-1} \cdot \tfrac{a_A}{4c-1}\right) - \tfrac{1}{2}\left(\tfrac{2a_A^{OC}c}{4c-1}\right)^2 \\ &= \tfrac{a_M^2}{4} + \tfrac{(2a_Ac-2bc)}{(4c-1)^2}\left(a_A + b\right)c. \end{aligned}$$

By comparing U_A^{SB} and U_A^{TB}, the first part of the proposition can be proved:

$$\begin{aligned} & U_A^{TB} \geq U_A^{SB} \\ \Leftrightarrow\ & b \leq \tfrac{a_A}{2c-1}. \end{aligned}$$

By comparing U_A^{SB} and $\widetilde{U}_A^{TB}$, the second part of the proposition can be shown:

$$\begin{aligned} & U_A^{SB} > \widetilde{U}_A^{TB} \\ \Leftrightarrow\ & a_A^2 > (a_A - b)(a_A + b) \\ \Leftrightarrow\ & b^2 > 0. \end{aligned}$$

Finally, the last part of the proposition can be shown by comparing U_A^{TB} and $\widetilde{U}_A^{TB}$:

$$\begin{aligned} & U_A^{TB} \geq \widetilde{U}_A^{TB} \\ \Leftrightarrow\ & \tfrac{b}{(4c-1)^2}(a_A + b)\, c \geq 0. \ \blacksquare \end{aligned}$$

4.7.5 Proof of Proposition 4

The total welfare in equilibrium with a fully rational agent is given by

$$\begin{aligned}
W^{SB} &= \left(a_M \cdot e_M^{SB} + a_A \cdot e_A^{SB} + e_A^{SB} \cdot h_M^{SB}\right) - \tfrac{1}{2} e_M^{SB2} - \tfrac{c}{2} h_M^{SB2} - \tfrac{1}{2} e_A^{SB2} \\
&= \left(a_M \cdot \tfrac{a_M}{2} + a_A \cdot \tfrac{2a_A c}{4c-1} + \tfrac{2a_A c}{4c-1} \cdot \tfrac{a_A}{4c-1}\right) - \tfrac{1}{2}\tfrac{a_M^2}{4} - \tfrac{c}{2}\left(\tfrac{a_A}{4c-1}\right)^2 - \tfrac{1}{2}\left(\tfrac{2a_A c}{4c-1}\right)^2 \\
&= \tfrac{3a_M^2}{8} + \tfrac{(12c-1)a_A^2 c}{2(4c-1)^2}.
\end{aligned}$$

With perfect information on ability and overconfidence parameters for the manager, the total welfare in equilibrium with an overconfident agent is equal to

$$\begin{aligned}
W^{TB} &= \left(a_M \cdot e_M^{TB} + a_A \cdot e_A^{TB} + e_A^{TB} \cdot h_M^{TB}\right) - \tfrac{1}{2} e_M^{TB2} - \tfrac{c}{2} h_M^{TB2} - \tfrac{1}{2} e_A^{TB2} \\
&= \left(a_M \cdot \tfrac{a_M}{2} + a_A \cdot \tfrac{2a_A^{OC} c}{4c-1} + \tfrac{2a_A^{OC} c}{4c-1} \cdot \tfrac{a_A^{OC}}{4c-1}\right) - \tfrac{1}{2}\tfrac{a_M^2}{4} - \tfrac{c}{2}\left(\tfrac{a_A^{OC}}{4c-1}\right)^2 - \tfrac{1}{2}\left(\tfrac{2a_A^{OC} c}{4c-1}\right)^2 \\
&= \tfrac{3a_M^2}{8} + \tfrac{(4(3a_A-b)c-(a_A-3b))(a_A+b)c}{2(4c-1)^2}.
\end{aligned}$$

If the manager is not aware of the agent's biased self-perception, the total welfare in equilibrium with an overconfident agent is equal to

$$\begin{aligned}
\widetilde{W}^{TB} &= \left(a_M \cdot \widetilde{e}_M^{TB} + a_A \cdot \widetilde{e}_A^{TB} + \widetilde{e}_A^{TB} \cdot \widetilde{h}_M^{TB}\right) - \tfrac{1}{2} \widetilde{e}_M^{TB2} - \tfrac{c}{2} \widetilde{h}_M^{TB2} - \tfrac{1}{2} \widetilde{e}_A^{TB2} \\
&= \left(a_M \cdot \tfrac{a_M}{2} + a_A \cdot \tfrac{2a_A^{OC} c}{4c-1} + \tfrac{2a_A^{OC} c}{4c-1} \cdot \tfrac{a_A}{4c-1}\right) - \tfrac{1}{2}\tfrac{a_M^2}{4} - \tfrac{c}{2}\left(\tfrac{a_A}{4c-1}\right)^2 - \tfrac{1}{2}\left(\tfrac{2a_A^{OC} c}{4c-1}\right)^2 \\
&= \tfrac{3a_M^2}{8} + \tfrac{\left(4(3a_A-b)(a_A+b)c-a_A^2\right)c}{2(4c-1)^2}.
\end{aligned}$$

By comparing W^{SB} and W^{TB}, the first part of the proposition can be proved:

$$W^{TB} - W^{SB} = \tfrac{(8a_A c+2a_A+3b-4bc)bc}{2(4c-1)^2}$$

$$\begin{aligned}
&\tfrac{(8a_A c+2a_A+3b-4bc)bc}{2(4c-1)^2} \geq 0 \\
\Leftrightarrow\quad & b \leq \tfrac{2(4c+1)}{4c-3} a_A.
\end{aligned}$$

By comparing W^{SB} and $\widetilde{W}^{TB}$, the second part of the proposition can be shown:

$$\widetilde{W}^{TB} - W^{SB} = \frac{2(2a_A-b)bc^2}{(4c-1)^2}$$

$$\begin{aligned} & \frac{2(2a_A-b)bc^2}{(4c-1)^2} \geq 0 \\ \Leftrightarrow \quad & b \leq 2a_A. \end{aligned}$$

Finally, the last part of the proposition can be shown by comparing W^{TB} and $\widetilde{W}^{TB}$:

$$W^{TB} - \widetilde{W}^{TB} = \frac{(2a_A+3b)bc}{2(4c-1)^2} > 0. \blacksquare$$

4.7.6 Proof of Proposition 5

With perfect information on ability and overconfidence parameters for the manager, the total welfare in the third-best equilibrium is given by

$$W^{TB} = \frac{3a_M^2}{8} + \frac{((3a_A-b)4c-(a_A-3b))(a_A+b)c}{2(4c-1)^2}.$$

Solving the first order condition ($\frac{\partial W^{TB}}{\partial b} \doteq 0$) leads to

$$b^* = \frac{4c+1}{4c-3}a_A.$$

The second order condition is always satisfied as

$$\frac{\partial^2 W^{TB}}{\partial b^2} = \frac{(3-4c)c}{(4c-1)^2} < 0.$$

If the manager is not aware of the agent's biased self-perception, the total welfare is equal to

$$\widetilde{W}^{TB} = \frac{3a_M^2}{8} + \frac{\left(4(3a_A-b)(a_A+b)c-a_A^2\right)c}{2(4c-1)^2}.$$

Again, solving the first order condition ($\frac{\partial \widetilde{W}^{TB}}{\partial b} \doteq 0$) leads to

$$\widetilde{b}^* = a_A.$$

The second order condition is also always satisfied as

$$\frac{\partial^2 \widetilde{W}^{TB}}{\partial b^2} = -\frac{4c^2}{(4c-1)^2} < 0. \blacksquare$$

Bibliography

Abeler, J., S. Altmann, S. Kube, and M. Wibral (2010). Gift exchange and workers' fairness concerns: When equality is unfair. *Journal of the European Economic Association 8*, 1299–1324.

Adams, J. (1965). Inequality in social exchange. In L. Berkowitz (Ed.), *Advances in Experimental Psychology*, pp. 267–299. New York: Academic Press.

Agell, J. and P. Lundborg (1995). Theories of pay and unemployment: Survey evidence from swedish manufacturing firms. *Scandinavian Journal of Economics 97*(2), 295–307.

Akerlof, G. A. and J. L. Yellen (1990). The fair wage-effort hypothesis and unemployment. *The Quarterly Journal of Economics 105(2)*, 255–283.

Alchian, A. A. and H. Demsetz (1972). Production, information costs, and economic organization. *American Economic Review 62*, 777–795.

Alesina, A. and G.-M. Angeletos (2005). Fairness and redistribution. *American Economic Review 95*(4), 960–980.

Anderson, L., J. Mellor, and J. Milyo (2008). Inequality and public good provision: An experimental analysis. *Journal of Socio-Economics 37*(3), 1010–1028.

Andreoni, J. (1990). Impure altruism and donations to public goods: A theory of warm-glow giving. *The Economic Journal 100*, 464–477.

Barber, B. M. and T. Odean (2000). Trading is hazardous to your wealth: the common stock investment performance of individual investors. *Journal of Finance 55*, 773–806.

Barber, B. M. and T. Odean (2001). Boys will be boys: gender, overconfidence, and common stock investment. *Quarterly Journal of Economics 116*, 261–292.

Bartling, B. and F. von Siemens (2010). Equal sharing rules in partnerships. *Journal of Institutional and Theoretical Economics 166(2)*, 299–320.

Ben-David, I., J. Graham, and C. Harvey (2007). Managerial overconfidence and corporate policies. *NBER Working Paper 13711*.

Bengtsson, C., M. Persson, and P. Willenhag (2005). Gender and overconfidence. *Economics Letters 86*, 199–203.

Berger, J., C. Herbertz, and D. Sliwka (2011). Incentives and cooperation in firms: Field evidence. *IZA Discussion Paper No. 5618*.

Bergstrom, T. C., L. Blume, and H. Varian (1986). On the private provision of public goods. *Journal of Public Economics 29*, 25–49.

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

Blinder, A. S. and D. H. Choi (1990). A shred of evidence on theories of wage stickiness. *The Quarterly Journal of Economics 105*(4), 1003–1015.

Blume, A. and A. Ortmann (2007). The effects of costless pre-play communication: Experimental evidence from games with Pareto-ranked equilibria. *Journal of Economic Theory 132*(1), 274–290.

Bolton, G. E. and A. Ockenfels (2000). ERC - a theory of equity, reciprocity and competition. *American Economic Review 90*, 166–193.

Buckley, E. and R. Croson (2006). Income and wealth heterogeneity in the voluntary provision of linear public goods. *Journal of Public Economics 90*(4-5), 935–955.

Burks, S., J. Carpenter, and L. Goette (2009). Performance pay and worker cooperation: Evidence from an artefactual field experiment. *Journal of Economic Behavior & Organization 70*, 458–469.

Camerer, C. (2003). *Behavioral Game Theory: Experiments on Strategic Interaction*. Princeton: Princeton University Press.

Camerer, C. and D. Lovallo (1999). Overconfidence and excess entry: An experimental approach. *American Economic Review 89*(1), 306–318.

Campbell, C. and K. Kamlani (1997). The reasons for wage rigidity: Evidence from a survey of firms. *The Quarterly Journal of Economics 112*, 759–789.

Cappelen, A. W., A. D. Hole, E. O. Sorensen, and B. Tungodden (2007). The pluralism of fairness ideals: An experimental approach. *American Economic Review 97*(3), 818–827.

Carpenter, J. P. and E. Seki (2006). Competitive work environments and social preferences: Field experimental evidence from a japanese fishing community. *The B.E. Journal of Economic Analysis & Policy 5*(2).

Chan, K., S. Mestelman, R. Moir, and R. Muller (1996). The voluntary provision of public goods under varying income distributions. *Canadian Journal of Economics 29*(1), 54–69.

Chan, K., S. Mestelman, R. Moir, and R. Muller (1999). Heterogeneity and the voluntary provision of public goods. *Experimental Economics 2*(1), 5–30.

Cherry, T., S. Kroll, and J. Shogren (2005). The impact of endowment heterogeneity and origin on public good contributions: evidence from the lab. *Journal of Economic Behavior & Organization 57*(3), 357–365.

Cooper, R., D. DeJong, R. Forsythe, and T. Ross (1992). Communication in coordination games. *The Quarterly Journal of Economics 107*(2), 739–771.

Correll, S. J. (2001). Gender and the career choice process: the role of biased self-assessments. *American Journal of Sociology 106*, 1691–1730.

De Bondt, W. F. M. and R. H. Thaler (1996). Financial decision-making in markets and firms: A behavioral perspective. In *Handbook in Operations Research and Management Science, Vol. 9*. North Holland.

Deaves, R., E. Lüders, and G. Luo (2008). An experimental test of the impact of overconfidence and gender on trading activity. *Review of Finance 13*, 555–575.

Demichelis, S. and J. Weibull (2008). Language, meaning, and games: A model of communication, coordination, and evolution. *American Economic Review 98*(4), 1292–1311.

Demougin, D., C. Fluet, and C. Helm (2006). Output and wages with inequality averse agents. *Canadian Journal of Economics 39*(2), 399–413.

Drago, R. and G. T. Garvey (1998). Incentives for helping on the job: Theory and evidence. *Journal of Labor Economics 16*(1), 1–25.

Drago, R. and G. K. Turnbull (1988). Individual versus group piece rates under team technologies. *Journal of the Japanese and International Economies 2*, 1–10.

Dur, R. and A. Glazer (2008). Optimal contracts when a worker envies his boss. *Journal of Law, Economics, and Organization 24*(1), 120–137.

Dur, R. and J. Sol (2010). Social interaction, co-worker altruism, and incentives. *Games and Economic Behavior 69*, 293–301.

Durante, R. and L. Putterman (2009). Preferences for redistribution and perception of fairness: An experimental study. *Working Paper*.

Encinosa, W., M. Gaynor, and J. B. Rebitzer (2007). The sociology of groups and the economics of incentives: Theory and evidence on compensation systems. *Journal of Economic Behavior & Organization 62*, 187–214.

Engelmann, D. and M. Strobel (2004). Inequality aversion, efficiency and maximin preferences in simple distribution experiments. *American Economic Review 94*, 857–869.

Englmaier, F. (2011). Commitment in r&d tournaments via strategic delegation to overoptimistic managers. *Managerial and Decision Economics 32*, 63–69.

Englmaier, F. and A. Wambach (2010). Optimal incentive contracts under inequity aversion. *Games and Economic Behavior 69(2)*, 312–328.

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

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

Fehr, E. and K. M. Schmidt (2002). Theories of fairness and reciprocity - evidence and economic applications. In M. Dewatripont, L. Hansen, and S. Turnovsky (Eds.), *Advances in Economics and Econometrics - 8th World Congress.*

Galasso, A. and T. S. Simcoe (2011). CEO overconfidence and innovation. *Management Science 57*, 1469–1484.

Gervais, S. and I. Goldstein (2007). The positive effects of biased self-perceptions in firms. *Review of Finance 11*, 453–496.

Goerg, S. J., S. Kube, and R. Zultan (2010). Treating equals unequally - incentives in teams, workers' motivation and production technology. *Journal of Labor Economics 28*, 747–772.

Grieco, D. and R. M. Hogarth (2009). Overconfidence in absolute and relative performance: The regression hypothesis and bayesian updating. *Journal of Economic Psychology 30*, 756–771.

Groves, T. (1973). Incentives in teams. *Econometrica 41*(4), 617–631.

Grund, C. and D. Sliwka (2005). Envy and compassion in tournaments. *Journal of Economics and Management Strategy 14*, 187–207.

Gürtler, O. (2008). On delegation under relational contracts. *International Journal of the Economics of Business 15*, 85–98.

Hackbarth, D. (2008). Managerial traits and capital structure decisions. *Journal of Financial and Quantitative Analysis 43*, 843–881.

Hardin, G. (1968). The tragedy of the commons. *Science 162*, 1243–1248.

Heywood, J. S., U. Jirjahn, and G. Tsertsvadze (2005). Getting along with colleagues - does profit sharing help or hurt? *Kyklos 58*, 557–573.

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

Holmström, B. and P. Milgrom (1991). Multitask principal-agent analyses: Incentive contracts, asset ownership and job design. *Journal of Law, Economics and Organization 7*, 24–52.

Homans, G. C. (1958). Social behavior as exchange. *The American Journal of Sociology 63*(6), 597–606.

Huck, S. and P. Rey-Biel (2006). Endogenous leadership in teams. *Journal of Institutional and Theoretical Economics (JITE) 162*(2), 253–261.

Itaya, J., D. De Meza, and G. Myles (1997). In praise of inequality: public good provision and income distribution. *Economics Letters 57*(3), 289–296.

Itoh, H. (1991). Incentives to help in multi-agent situations. *Econometrica 59*(3), 611–636.

Itoh, H. (1992). Cooperation in hierachical organizations: An incentive perspective. *Journal of Law, Economiocs, and Organization 8*, 321–345.

Itoh, H. (1994). Job design, delegation and cooperation: A principal-agent analysis. *European Economic Review 38*, 691–700.

Itoh, H. (2001). Job design and incentives in hierarchies with team production. *Hitotsubashi Journal of Commerce and Management 36*, 1–17.

Itoh, H. (2004). Moral hazard and other-regarding preferences. *The Japanese Economic Review 55*(1), 18–45.

Kandel, E. and E. P. Lazear (1992). Peer pressure and partnerships. *Journal of Political Economy 100*, 801–17.

Kinari, Y., N. Mizutani, F. Ohtake, and H. Okudaira (2011). Overconfidence increases productivity. *Working Paper*.

Klayman, J., J. B. Soll, C. González-Vallejo, and S. Barlas (1999). Overconfidence: It depends on how, what, and whom you ask. *Organizational Behavior and Human Decision Processes 79*, 216–247.

Konow, J. (2000). Fair shares: Accountability and cognitive dissonance in allocation decisions. *American Economic Review 90*(4), 1072–1091.

Konow, J., T. Saijo, and K. Akai (2009). Morals and mores? experimental evidence on equity and equality. *Working Paper*.

Kragl, J. and J. Schmid (2009). The impact of envy on relational employment contracts. *Journal of Economic Behavior & Organization 72*(2), 766–779.

Kölle, F., D. Sliwka, and N. Zhou (2011). Inequality, inequity aversion, and the provision of public goods. *IZA Discussion Paper No. 5514*.

Lazear, E. P. (1989). Pay equality and industrial politics. *Journal of Political Economy 97*, 561–580.

Lichtenstein, S., B. Fischhoff, and L. D. Phillips (1982). Calibration of probabilities: The state of the art to 1980. In D. Kahneman, P. Slovic, and A. Tversky (Eds.), *Judgement under uncertainty: Heuristics and biases*, pp. 306–334. New York: Cambridge University Press.

Ludwig, S., P. C. Wichardt, and H. Wickhorst (2011a). On the positive effects of overconfident self-perception in teams. *Munich Discussion Paper No. 2011-11*.

Ludwig, S., P. C. Wichardt, and H. Wickhorst (2011b). Overconfidence can improve an agent's relative and absolute performance in contests. *Economics Letters 110*(3), 193–196.

Malmendier, U. and G. Tate (2005). CEO overconfidence and corporate investment. *The Journal of Finance 60*, 2661–2700.

Malmendier, U. and G. Tate (2008). Who makes acquisitions? CEO overconfidence and the market's reaction. *Journal of Financial Economics 89*(1), 20–43.

Mohnen, A., K. Pokorny, and D. Sliwka (2008). Transparency, Inequity Aversion, and the Dynamics of Peer Pressure in Teams: Theory and Evidence. *Journal of Labor Economics 26*(4), 693–720.

Neilson, W. and J. Stowe (2010). Piece-rate contracts for other-regarding workers. *Economic Inquiry 48(3)*, 575–586.

Nieken, P., A. Sadrieh, and N. Zhou (2011). Overconfidence and managers' responsibility hoarding. *Working Paper*.

Olson, M. (1965). *The Logic of Collective Action*. Cambridge University Press.

Ostrom, E., R. Gardner, and J. Walker (1994). *Rules, Games, and Common-Pool Resources*. University of Michigan Press.

Palomino, F. and E. Peyrache (2010). Psychological bias and gender wage gap. *Journal of Economic Behavior & Organization 76*, 563–573.

Palomino, F. and A. Sadrieh (2011). Overconfidence and delegated portfolio management. *Journal of Financial Intermediation 20*(2), 159–177.

Prendergast, C. J. (1995). A theory of responsibility in organizations. *Journal of Labor Economics 13*, 387–400.

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

Rey-Biel, P. (2008). Inequity aversion and team incentives. *Scandinavian Journal of Economics 110*(2), 297–320.

Roth, A. and J. Kagel (1995). Bargaining experiments. In *Handbook of Experimental Economics*. Princeton University Press.

Russo, J. E. and P. J. H. Schoemaker (1992). Managing overconfidence. *Sloan Management Review 33*, 7–17.

Santos-Pinto, L. (2008). Positive self-image and incentives in organisations. *The Economic Journal 118*, 1315–1332.

Santos-Pinto, L. (2010). Positive self-image in tournaments. *International Economic Review 51*(2), 475–496.

Sautmann, A. (2011). Contracts for agents with biased beliefs: Some theory and an experiment. *Working Paper*.

Sobel, J. (2005). Interdependent preferences and reciprocity. *Journal of Economic Literature 43*, 392–436.

Soll, J. B. (1996). Determinants of Overconfidence and Miscalibration: The Roles of Random Error and Ecological Structure. *Organizational Behavior and Human Decision Processes 65*, 117–137.

Squintani, F. (2006). Equilibrium and mistaken self-perception. *Economic Theory 27*, 615–641.

Svenson, O. (1981). Are we all less risky and more skillful than our fellow drivers? *Acta Psychologica 47*, 143–148.

Thurow, L. (1971). The income distribution as a pure public good. *The Quarterly Journal of Economics 85*(2), 327–336.

Van Dijk, F., J. Sonnemans, and F. Van Winden (2002). Social ties in a public good experiment. *Journal of Public Economics 85*(2), 275–299.

Warr, P. G. (1983). The private provision of a public good is independent of the distribution of income. *Economics Letters 13*, 207–211.

Weinberg, B. A. (2009). A model of overconfidence. *Pacific Economic Review 14*(4), 502–515.

Winter, E. (2004). Incentives and discrimination. *American Economic Review 94*(3), 764–773.

Young, H. P. (2002). Bounded rationality and learning. on the limits to rational learning. *European Economic Review 46*, 791–799.

Zhou, N. (2011). Overconfidence, helping effort, and team performance. *Working Paper*.