Overconfidence and Loss Aversion in Economic Decision Making

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

vorgelegt von
Mara Ewers
aus Berlin

Bonn 2012

Dekan: Prof. Dr. Klaus Sandmann

Erstreferent: Prof. Dr. Matthias Kräkel

Zweitreferent: Prof. Dr. Armin Falk

Tag der mündlichen Prüfung: 23.10.2012

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

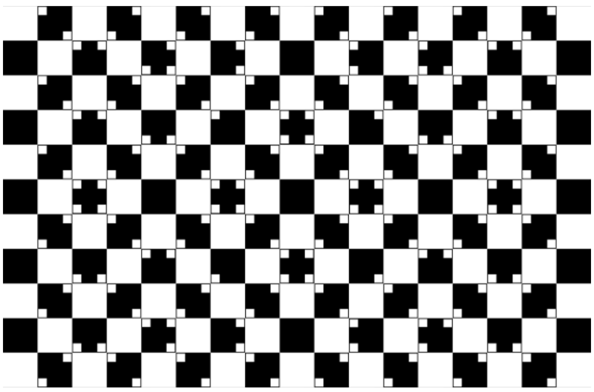
Contents

List of Figures

List of Tables

Introduction

The fact that individuals are not always unperturbed by the moral and emotional environment, information absorbing, equipped with a memory capacity of an infinitely large computer chip and money maximizing has evoked the field of behavioral economics. Individuals that decide about choices and interact in groups are found to have social preferences, do not process all information accurately or suffer from biased perceptions. One example which illustrates that our human senses sometimes perceive information in a biased manner is the following visual trick. If you look at the figure and if you are a human with a not too poor eyesight you will see wavy lines, because your senses trick you and the objective truth, that all lines are straight, is not obvious to you. You might even want to get out a ruler to control your biased perception.



Humans can maximize monetary outcomes and solve mathematical problems, but their decision making depends tremendously on the emotions evoked by the context, the presentation and the complexity. With this dissertation I want to give you an understanding of the interesting phenomena of biased self-assessments like overconfidence by providing a reason for overconfident appearance and survival. Further, an analysis of the consequences of overconfidence and potential de-biasing nudges are discussed. Overconfidence is an omnipresent finding in the literature of psychology and also economics that influences our decision making in economic as well as non-economic situations. For example, when asking university professors whether they are better or worse than the average professor, many more answer that they are better than that they are worse.[1] This biased perception does not only occur for university professors, and also for a wide range of tasks like test taking or car driving. A positive self-image is a desirable and important trait and encourages humans to become involved and withstand new challenges Bénabou and Tirole (2002). However a systematic bias like overconfidence has been blamed as a reason for excessive market entry, costly delays in negotiations, excessive litigation, and excessive stock trading (see e.g. Camerer and Lovallo (1999), Neale and Bazerman (1985), Odean (1998)).

An interesting question is why do we observe such overly optimistic statements and why do humans not learn that not everyone can be better than the other one. My coauthor Florian Zimmermann and I propose image concerns and social approval seeking as an explanation why decision makers forgo monetary incentives and that image concerns prompt them to misreport their private information in an overly positive way. This means that the outcome of the decision making process is highly influenced by whether an audience is listening or not. Another important question concerns the consequences of self-assessment biases on economic situations and how to prevent bad consequences. I analyze the implications of absolute and relative self-assessment biases on the entry choice into a competitive environment and what kind of information can nudge individuals to raise attention to the complexity of the decision problem.

Chapter one considers situations where truthful revelation of private information maximizes monetary utility. My coauthor and I ask whether reports of private information about skills, abilities or achievements are affected by image concerns. Such reports of private information might be important for example for a principal who has to allocate tasks and workers efficiently. Our hypothesis is that image concerns might affect reports in such a way that humans appear overconfident to the audience by revealing private information about ability or skill in an overly optimistic

[1] While this example might even be true if there is one really, really bad professor, it is easy to come up with other evidence (see Englmeier (2004) for a survey overview).

way. Our approach to analyze this problem is by developing a simple model that illustrates how image utility can lead to misreporting of private information in contexts where truthful reports maximize monetary outcomes. In addition, we test the model's predictions in a controlled laboratory experiment. In the experiment, all subjects go through a series of incentivized quiz questions and subsequently have to report a performance measure. They have to assess whether their quiz performance was better than the average performance and receive 5 euros for a correct assessment. We vary if reports are made to an audience or not and find evidence for image effects. In the audience treatment, stated reports are significantly higher than in the private treatment. This suggests that overconfident appearance might be a consequence of social approval seeking. We also find that men state higher self-assessments than women. This gender difference seems to be driven by men responding more strongly to the presence of an audience.

As a consequence of overconfident self-assessments, overestimation of ability and skill, the choices in various decision problems are affected. For example, occupational choices between fix wages or performance dependent wages are biased towards the latter one. Chapter two studies entry choices into a competition game, where the entry choice depends highly on the individual's self-assessment about performance. We study the influence of information on entry choices with a controlled laboratory experiment and investigate whether information provision attracts mainly high performance individuals and reduces competition failure. Competition failure occurs when an individual loses the competition because the opponent holds a higher productivity. In the experiment, subjects face the choice between a competition game and a safe and fix outside option. We analyze subjects' entry behavior with a benchmark treatment without information and three treatments, where we exogenously manipulate the information on the opponents. In one treatment, we disclose the performance distribution of all subjects before each subject decides about entry. In two further treatments, we disclose the matched opponent's true performance or the matched opponent's true performance plus the opponent's performance self-assessment. Information on the opponent is a promising nudge to raise individuals' awareness towards the complexity of the decision problem and to update beliefs about success. We present two main results on the usefulness of information and further results on subject's characteristics like gender and self-assessments that contribute to the literature on self-selection into competition. Our results are, (1) information about the productivity distribution of all potential opponents reduces competition failures by more than 50%, (2) information on the distribution is sufficient, i.e. precise information on the matched opponent's type does not further diminish failure rates. Furthermore, we find a gender difference in competition entry choices in the benchmark treatment which vanishes when giving precise information about the opponent's performance. Several findings concerning the relationship between our different absolute and relative self-assessment measures are also discussed.

In the last chapter my coauthor Roland Eisenhuth and I study loss aversion in auctions. The concept of loss aversion makes the point that losses are perceived worse than equal sized gains; empirical work showed that losses are on average twice as worse as equal sized gains are good. Only framing a sentence like "Chances of survival are 95%." into a loss frame "Chancess of death are 5%." make humans perceive the same situation on average as twice as worse. And since emotions and the environment are a major influence factor for decision making, loss aversion impacts choices in many environments.

In Chapter three, we theoretically and experimentally study independent private value auctions in the presence of bidders who are loss averse in the sense of Köszegi and Rabin (2006, 2007). In one specification, we consider gains and losses in two dimensions separately, about whether they receive the object or not, and how much they pay; in the other specification, we consider gains and losses over the entire risk neutral pay off, i.e. the valuation less the bid. With one dimension, we show that the expected revenue for the auctioneer is higher in the first price auction than in the all pay auction, and with two dimensions, we show that the opposite is true for the revenue ranking between the first price auction and the all pay auction. In order to test the theoretical predictions, we conduct laboratory experiments, in which either money or a real object is auctioned off in both a first price auction and an all pay auction. In both settings, the average revenue is significantly higher in the first price auction, suggesting that bidders may behave according to the one dimensional model, although a real object is auctioned off. Whereas our findings are inconsistent with the two dimensional model, they are consistent with the one dimensional model.

Chapter 1

Image and Misreporting

1.1 Introduction

Individuals hold private beliefs about their performance, skills, abilities and achievements. These beliefs are key determinants in many economic choices, such as investing in education, choosing an optimal health insurance plan, applying for a new position or accepting a new job. Likewise, transmission of this private information is crucial for economic interactions. For instance, efficient allocation of tasks within a firm relies on information about employees' skills and abilities. The same is true for decisions about job promotions or efficient specialization. In this paper, we analyze whether individuals' image concerns can lead them to misreport private information in situations, where truthful revelation would be optimal from a traditional pecuniary persepective. Individuals who care about how they are perceived by their environment, will take this perception into account when making choices or assessing own performance and abilities in front of others. We show with a simple model how the presence of image concerns makes people misreport their own performance, skill or ability. Individuals with low performance will choose to report high performance. Then we provide evidence from a lab experiment. In the experiment we exogenously increase subjects' image concerns with a procedure used by Ariely, Bracha, and Meier (2009), and document the consequences of a desire for a favorable image on statements about own performance.

In our model, decision makers' choice is to publicly report private information about their own type. Correctly stating their private information is optimal in direct monetary terms. However, we assume that decision makers' utility consists of two components, a "standard" part, reflecting monetary concerns and an image part, reflecting reputational concerns (similar as in Bénabou and Tirole (2006) in the context of prosocial behavior). The way we model image concerns is a shortcut that captures all benefits from signaling a high type. The nature of reputational concerns could be strategic. In labor market contexts, signaling of abilities and skills may improve hiring prospects and lead to higher wages or promotion. Benefits could also be in the form of social approval. Alternatively, decision makers could value reputation for hedonic reasons. People simply enjoy being regarded as a high type. We show the existence of a unique Perfect Bayesian Nash equilibrium, where a decision maker misreports private information. Low skilled types choose to signal a high type, if image concerns are relevant. We also show that misreporting increases in the relative importance of image utility. While we focus on social image concerns, our model is also compatible with a self-signaling interpretation where decision makers learn about their own type by inference from own choices (e.g., as in Bénabou and Tirole (2004, 2006)).

We test the main prediction of our model, that image concerns lead to misreporting of private information, in a laboratory experiment. The experiment consists of two stages. In stage 1, subjects go through a series of general knowledge quiz questions. In stage 2, subjects are asked to give a binary and incentivized self-assessment concerning their quiz performance. We study two main treatments: In the audience treatment, we exogenously increase subjects' image concerns in stage 2 by making them report their self-assessment to the other subjects present in the lab. After all subjects have given their binary assessment, one after the other has to stand up and report his or her self-assessment to the group. In the private treatment, subjects do not report their assessment to the group. Our data reveal significant evidence for image effects. In the audience treatment, stated self-assessments are significantly higher than in the private treatment. We also document a gender difference in stated self-assessments, with men reporting higher performance than women. This difference is driven by a stronger response of men to the presence of an audience.

Our findings show that image concerns play an important role in the transmission of private information about skill, ability or performance. Even if truthful reporting is optimal in monetary terms, decision makers misreport. This contributes to a large literature that has documented significant biases in stated self-assessments. If individuals are asked to assess their own type in terms of performance or ability, their self-assessments are frequently overly optimistic. One of the most prominent examples of highly optimistic beliefs is a study by Svenson (1981) on relative self-assessments in the context of car driving skills. He finds, for instance, that 40% of subjects place themselves in the top 20% of car drivers with regard to driving skills. Overconfident self-assessments are also found when subjects are given monetary incentives to correctly evaluate their skill or performance (see Hoelzl and Rustichini (2005)).[1] Our theoretical and experimental results suggest, that documented biases in self-assessments might be produced by a desire to gain a favorable image. By trying to signal a high type, decision makers appear overconfident. This can occur even with perfect knowledge about their own performance, skill or ability. Decision makers can appear overconfident without any inherent biases in self-assessments. The recipient of signals can be an audience, peers, employers etc. In experimental settings, the recipients can be other subjects (as in our experiment), the experimenter, but also the decision-maker himself. In this self-signaling interpretation, the decision-maker learns about his own type from observing his own choices. In the psychology literature, the idea that people construct their own type or self-image from past choices can be found in Bem (1972).

Our findings allow a novel perspective on biases in self-assessments. We show that stated self-assessments can differ between conditions that are identical in all dimensions, except the presence of an audience. This difference cannot be due to a passive, inherent bias or mistake. Instead, the bias in stated self-assessments is actively chosen. Individuals respond to the presence of an audience by making a strategic choice to misreport own performance. In this interpretation, overconfidence is rather the outcome of a preference, e.g., a desire to signal skills or ability, than a mistaken self-perception. This perspective on overconfidence also offers a straightfor-

[1] For a recent overview on empirical studies on overconfidence, see Benoît and Dubra (2011). Several studies examine the consequences of overconfidence for behavior in different contexts. Examples are Dohmen and Falk (2011) in the context of tournament entry, Malmendier and Tate (2008) for CEO behavior, DellaVigna and Malmendier (2006) for overestimation of future gym attendance or Camerer and Lovallo (1999) for excess entry into competitive markets.

ward explanation for why people do not "learn" about their mistaken self-assessment over time. In our setup, individuals can appear and behave overconfident even with perfect knowledge about their own type.

Our experimental findings can also help to provide an explanation for why men appear to be more overconfident than women. A gender difference in self-assessments has been reported in many studies and provides an explanation for gender differences in selection into competition (see for example Gneezy and Rustichini (2004), Gneezy, Niederle, and Rustichini (2003) Niederle and Vesterlund (2007), Dohmen and Falk (2011)). We find that men appear more overconfident than women in the audience treatment, but not in the private treatment. This suggests, that men feel a stronger desire to signal skills or abilities towards others, resulting in overconfident appearance.

Note that other explanations for overconfident behavior have been suggested: Bénabou and Tirole (2002) provide a theoretical argument for a value of self-serving beliefs as these can increase motivation of individuals with imperfect willpower. Other models assume a value of self-confidence and show how overconfident self-assessments can be produced by selectively choosing information or by asymmetrically processing information, putting more weight on positive than on negative information (see for example Brunnermeier and Parker (2005)), Kőszegi (2006) or Möbius et al. (2011)). Recently, several experimental papers have provided support for biases in information-processing and information demand (see Eil and Rao (2011), Möbius et al. (2011) and Charness, Rustichini, and van de Ven (2011)). Benoît and Dubra (2011) provide a different explanation. They argue that most of the evidence for relative overconfidence can in fact be reconciled by correct Bayesian updating from common priors. In other words, evidence in the form "40% of subjects place themselves in the top 20% of good car drivers" should not be interpreted as evidence for overconfident self-assessments as it can be the outcome of correct updating from unbiased information. While all approaches are important and in concert provide a good explanation for documented behavior, our experimental results highlight the crucial role of image concerns for stated self-assessments.

Our findings are also informative from a methodological perspective. They suggest that appropriate monetary incentives alone might not be sufficient to ensure truthful revelation of self-assessments in experiments or surveys. The presence of image concerns creates a trade-off between image concerns and monetary outcomes which leads to biases in stated self-assessments. Minimization of image concerns via, for instance double-blind procedures, might help mitigating this problem. Likewise, our findings are relevant from a mechanism design perspective. They show that mechanisms designed with a purely monetary focus do not necessarily lead to truthful revelation of private information. If people have strong image concerns, these ought to be taken into account when designing optimal mechanisms, e.g., insurance plans or employment contracts. While our focus is on direct transmission of private information, our results apply more generally. In many decision contexts that require prior self-assessment, decision makers' choices allow them to signal skill, ability or performance to others. We discuss this in more detail in section 1.5.

This paper relates to a few recent papers that considered the social signaling component of biases in self-assessments. Burks et al. (2010) compare different explanations of overconfidence in a large survey study with truck drivers. Their results suggest a strong connection between image concerns and overconfidence. Truckers reporting that they care about how others perceive them, significantly overplace their performance in an IQ test and a numeracy task. Charness, Rustichini, and van de Ven (2011) provide experimental evidence that men exploit the possibility to send an exaggerated productivity signal in a strategic interaction of a tournament entry to deter entry of other individuals while women do not. In their paper, they also find evidence for a consumption value from overconfidence.[2] In a related experiment, Reuben et al. (2010) find that subjects exaggerate past performance in order to become a group leader. In contrast, we focus on situations without a monetary incentive to misreport.

More broadly, this paper relates to several papers that study consequences of image concerns on economic decision making in different contexts. So far the literature has mainly analyzed effects of social approval for prosocial decision making. Non-anonymity or the presence of an audience has been shown to increase prosociality (see Gächter and Fehr (1999), Rege and Telle (2004), Andreoni and Petrie (2004) and Ariely, Bracha, and Meier (2009)). Theoretical papers analyzing image concerns

[2]Eil and Rao (2011) and Möbius et al. (2011) also provide evidence for a consumption value from overconfidence.

in a prosocial context include Bénabou and Tirole (2006), Ellingsen and Johannesson (2008) and Andreoni and Bernheim (2009). Closest to our modeling approach is the paper by Bénabou and Tirole (2006). They show how extrinsic incentives can crowd out prosocial behavior, because they destroy the image rewards from prosocial activity. Recently, Falk and Zimmermann (2011) examined the role of image concerns in the context of consistency of behavior. They show that individuals want to behave consistently as this allows the signaling of strength. In our paper, we show that the desire to signal skills or ability can lead to misreporting of private information.

The remainder of the paper is organized as follows. The next section introduces our model. Section 1.3 presents the experimental design, section 1.4 the results from our experiment and section 1.5 concludes.

1.2 Model

We provide a simple framework that allows illustrating how image concerns can influence reports of private information.

Consider decision makers D that differ in a parameter p which is an element of $P = \{0, 1, ..., \bar{p}\}$. Depending on the context, p captures D's ability, skill, performance or achievement. p is D's private information but is commonly known to be distributed according to a probability function f defined over P. At first, we assume that decision makers have perfect knowledge about p. In Appendix 1.A we provide a version of the model where decision makers have imperfect knowledge about their type and show that this produces qualitatively the same results. Decision makers' choice x is to report some measure related to p in public. We assume a binary report: is p larger than some value $\overline{r}$? This report could be absolute (is p higher than a certain number?), or relative to others (is p higher than the average or the median performance of other decision makers?). Thus, we have that $x \in \{Yes, No\}$. Decision makers win a monetary prize y if their stated report is correct, otherwise they earn 0. Thus, choice x and prize y reflect contexts where truthful reporting of private information is optimal in direct monetary terms. In experimental settings, choice x and prize y simply capture an incentivized self-assessment. More generally, choice x could be a decision that depends on p, e.g., the choice to enter a tournament, and the prize y

reflects direct monetary consequences from that choice. Note that the prize y might also capture direct nonmonetary utility consequences from misreporting, e.g, costs of lying.[3]

We assume that utility has two sources, direct (monetary) payoffs and image utility. Money enters linearly in the utility function and the two components are additively separable. Thus utility is given by

$$U(x) = y\mathbf{1}(x) + \alpha\beta E(p \mid x).$$

The first part captures direct utility over money. $\mathbf{1}(x)$ is an indicator function taking the value 1 if the stated report is correct and 0 otherwise. The second part incorporates image utility. $E(p \mid x)$ is the public's expectation about D's performance, skill or ability p, conditional on D's choice x. Thus, the public infers decision makers' p from their reports, and social approval depends on that. α and β specify the strength of image concerns. α is an individual parameter, i.e., decision makers differ in α. Some D care more about their image or respond more strongly to social approval than others. α is assumed to be constant across contexts and environments. While α is D's private knowledge, it is commonly known to be drawn from a distribution described by a density function g over $[0, \overline{\alpha}]$ with $g(\alpha) > 0,\ \forall\, \alpha \in [0, \overline{\alpha}]$. We assume that performance or ability p and the desire for social approval α are drawn independently. β instead, is identical for all decision makers and we assume $\beta > 0$. β might depend on the context of the decision problem, e.g., the size of the public, the social distance between D and the public or the strategic value of a favorable image. Thus, β is the parameter that is exogenously manipulated in our experiment. An alternative interpretation of decision makers' image concerns is a desire for a positive self-image (similar as in Bénabou and Tirole (2004, 2006)).[4] In this case, decision makers receive a private signal about their performance or ability prior to their decision. Thus, when deciding, they hold information about their p. However, for their later self-evaluation, this knowledge is not available for example due to reasons of

[3]Gneezy (2005) and Fischbacher and Heusi (2008) examine lying behavior in different contexts. They find evidence that subjects lie, but also that there is some cost of lying that prevents subjects from lying 100%.

[4]In the psychology literature, the idea that people construct their self-image from past actions can be found in Bem (1972).

imperfect recall. Since actions are easier to recall than signals, decision makers base their self-evaluation on past stated reports.

We are now ready to state the following two Propositions (proofs are provided in Appendix 1.A):

Proposition 1. *If $\overline{\alpha}$ is sufficiently large, i.e., $\overline{\alpha}\beta\left[\sum_{p>\overline{r}} \frac{f(p)p}{\sum_{p>\overline{r}} f(p)} - \sum_{p\leq\overline{r}} \frac{f(p)p}{\sum_{p\leq\overline{r}} f(p)}\right] > y$, there exists a unique Perfect Bayesian Equilibrium where decision makers with $p < \overline{r}$ and $\alpha > \alpha^*$ choose $x = Yes$. Decision makers with $p > \overline{r}$ choose $x = Yes$ and those with $p < \overline{r}$ and $\alpha < \alpha^*$ choose $x = No$.*

Proposition 1 shows, under which conditions exists a unique Perfect Bayesian Equilibrium where decision-makers misreport their private information. If image concerns are large enough, i.e., image gains from choosing $x = Yes$, $\overline{\alpha}\beta\left[\sum_{p>\overline{r}} \frac{f(p)p}{\sum_{p>\overline{r}} f(p)} - \sum_{p\leq\overline{r}} \frac{f(p)p}{\sum_{p\leq\overline{r}} f(p)}\right]$, outweigh the monetary costs y, decision makers with low performance ($p < \overline{r}$) and relatively high image concens ($\alpha > \alpha^*$) will report $x = Yes$ in order to signal a high type.

Proposition 2. *An increase in β reduces the threshold type α^*. Consequently, more decision makers with $p < \overline{r}$ misreport by choosing $x = Yes$.*

Proposition 2 shows how reports change in β, for example, when the size of the public, the social distance between D and the public, or the strategic value of reputation changes. Our model predicts that an exogenous increase in image concerns increases the number of decision makers that misreport information. Consequently, reports become less informative. This is the comparative static we exploit with our experiment.

1.3 Experimental Design

Our model suggests that the desire for social approval will tempt decision makers to misreport their private information in public. To test this hypothesis, we introduced a simple choice environment where subjects held private information about their skill or performance. Then, we manipulated image concerns exogenously by varying whether private information is reported to a public or not.

Table 1.1 summarizes our experimental between-subjects design. We study two main treatments, an *audience treatment* and a *private treatment*. In both treatments, the experiment started with a short introductory game. Subjects, one after the other, were asked to stand up and provide the group with some personal information such as name, age, and field of study.[5] The main part of the experiment consisted of two stages. In stage 1, subjects were asked to answer 20 multiple-choice quiz questions. The questions covered various general knowledge topics like history, economics, math, or art. Subjects were given four possible answers and had to select one. We incentivized the quiz, such that subjects earned 40 cents for every correct answer. The number of correctly answered questions serves as our measure of performance. Subjects received no feedback regarding the number of correctly answered quiz questions. Therefore, they held private but not necessarily perfect information about their performance. In stage 2, subjects faced a simple incentivized self-assessment task.[6] We asked them to compare their own performance to the average quiz-performance of a group of other participants. The group of other participants consisted of 95 different subjects who also performed the quiz. We asked: "Do you think your quiz-performance was better or worse than the average performance of another group?"[7] Subjects received 5 euros for a correct self-assessment. Thus, monetary incentives to tell the truth were strong. The only difference between our two treatments was the following: In the *audience treatment*, all subjects entered their self-assessment into the computer, and then reported their self-assessment to the other subjects present in the lab. Subjects knew in advance that they had to report the assessment to the other subjects. Thus, after all subjects privately assessed their relative quiz-productivity and entered it in the computer, one after the other had to stand up, say their name

[5]The purpose of the introductory game was to reduce the social distance between partcipants. Gächter and Fehr (1999) show in the context of a public goods game that social approval incentives are only effective in combination with a procedure to increase familiarity among group members.

[6]Subjects were only informed about the self-assessment task after they finished stage 1.

[7]Studies that want to document relative overconfidence usually use comparisons to percentiles such as the median. For our question, identifying overconfidence is not the main goal, because we are particularly interested in the treatment effect on reported self-assessments. Therefore, we decided to use the simpler and more comprehensive average as measure of comparison.

and report their self-assessment to the group.[8] This procedure of introducing an audience to increase image concerns has been used for example in Ariely, Bracha, and Meier (2009) in the context of pro-social behavior. The *private treatment* was identical to the audience treatment, however subjects did not state their self-assessment towards the other subjects.

Table 1.1: Design of the experiment

Stage 1	Stage 2	Treatments	Questionnaire
Multiple-choice quiz - Number of correct answers is our measure of performance - 40 cents / correct answer	Self-assessment - Are you better or worse than the average? - 5 euros / correct self-assessment	1. Private: no further action 2. Audience: reporting self-assessment to an audience	- Risk - Survey questions - Demographics

1.3.1 Experimental Procedures

47 subjects participated in the private treatment, 48 in the audience treatment. We were interested in potential gender differences and therefore invited an equal amount of women and men to each session. All sessions of the experiment were conducted in the *B*onn*E*con*L*ab, subjects were recruited via ORSEE (Greiner (2004)) and the experiment was run using the experimental software z-Tree (Fischbacher (2007)). A session took on average 50 minutes and subjects earned 9.50 euros on average.[9] We distributed the instructions for stage one and two immediately before the stage started and they were read aloud.

[8]While subjects reported their private information (self-assessments) in front of the audience, their previously entered self-assessment was also shown on their computer screen to make sure subjects could not lie about their entered self-assessment.

[9]1 euro was worth about 1.4 dollar at the time.

1.3.2 Hypothesis

In the experiment, we systematically increase image concerns of subjects by introducing an audience. When comparing the private and the audience treatment, by Proposition 2 of our model, reported self-assessments should be higher in the audience treatment compared to the private treatment.

Hypothesis. *Subjects choose "better than average" more frequently in the audience treatment than in the private treatment.*

1.4 Results

In section 1.4.1, we compare reports of the audience and the private treatment. In addition, we show the influence of gender on our treatment effect and analyse individuals' perceptions of others' stated self-assessments. In section 1.4.2, we present results from an additional control treatment we conducted.

1.4.1 Main Results

Result 1. *There is a treatment difference in stated self-assessments: Subjects in the audience treatment report "better than average" significantly more often compared to subjects in the private treatment.*

We find that 68% of subjects in the audience treatment report to be "better than average", compared to 48% of subjects in the private treatment. This sizable effect is also statistically significant in probit regressions. Table 1.2 reports the marginal effects of three probit regressions (columns 1-3), regressing a treatment dummy and several controls on reported self-assessment, where 1 indicates a report "better than average".[10] Column 1 of Table 1.2 shows that the raw treatment effect is significant at the 5% level. Our finding is robust when controlling for different measures of quiz performance. In column 2, we take the number of correctly solved quiz questions

[10]Columns 4 and 5 of Table 1.2 are discussed later.

Table 1.2: Determinants of stated self-assessment

Dependent variable: Relative self-assessment= $\begin{cases} 1 & \text{if report is ``better than average''} \\ 0 & \text{if report is ``worse than average''} \end{cases}$

	(1) All	(2) All	(3) All	(4) Private	(5) Audience
Dummy treatment	0.20**	0.25**	0.27**		
	(0.09)	(0.12)	(0.12)		
Quiz performance		0.07***		0.07*	0.08**
		(0.03)		(0.04)	(0.04)
Dummy quiz performance			0.16		
			(0.11)		
Dummy gender		-0.31***	-0.37***	-0.21	-0.28**
		(0.11)	(0.10)	(0.18)	(0.12)
Controls		included	included	included	included
N	95	95	95	47	48
-LL	62	50	53	24	18

Notes: Probit estimates. Marginal effects (evaluated at the mean of independent variables) reported; robust standard errors are in parentheses. Significance at the 1, 5, and 10 percent level is denoted by ***, **, and *, respectively. *Dummy treatment* =1 if audience treatment and 0 if private treatment. *Dummy gender* =1 if female. *Dummy quiz performance* =1 if better than average. *Controls* include a survey based risk measure, a measure of image concerns, age, and relationship status.

as a control for quiz performance. In column 3, we use a different measure: we create a performance dummy, taking the value one if performance was actually better than average and zero otherwise. In both regressions, the treatment effect remains significant at the 5% level.[11] In addition, nonparametric testing with a Fisher-exact test also confirms result 1 ($p - value = 0.06$, two-sided).

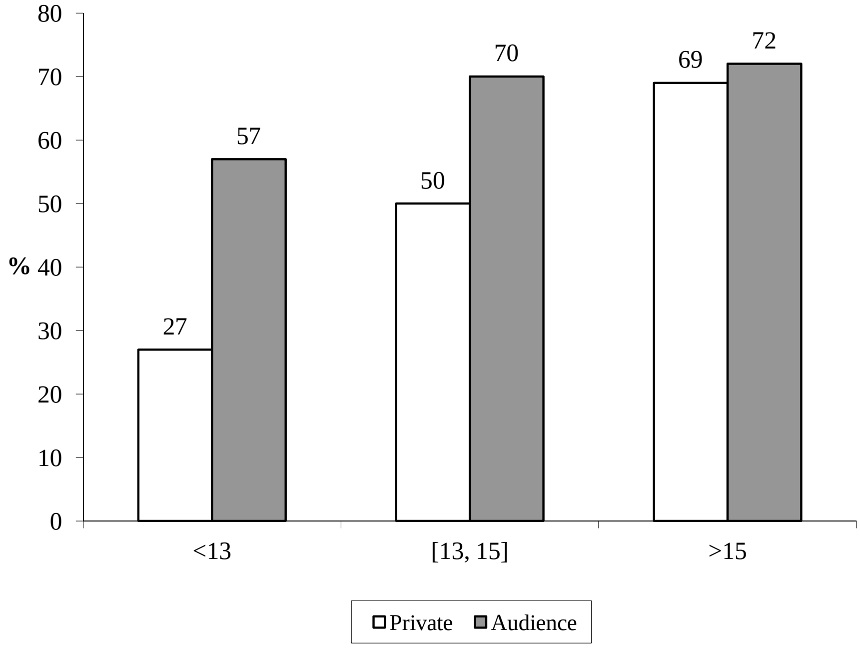
A different way to look at our data is to analyze the treatment effect for different

[11]Note that the average quiz performance over all treatments is 14.4 correctly solved quiz questions. The distributions of quiz performance do not significantly differ across treatments ($p - values > 0.34$ of Kolmogorov-Smirnov tests). The comparison group of 95 participants had an average quiz performance of 14.3 which is also not significantly different from performances of subjects in our treatments.

intervals of actual quiz performance. According to our model, the treatment effect should be driven by subjects who place themselves below the average, when privately evaluating own performance, but want to signal high quiz performance towards others. With perfect, as well as with imperfect knowledge of their performance (see Appendix 1.A), most subjects with low quiz performance privately place themselves below average, while those with high performance, mostly place themselves above average. Consequently, our model predicts that stated self-assessments for subjects with rather high quiz performance should be similar between treatments, while reports for subjects with low quiz performance should differ between treatments. This is indeed what we find. Figure 1.1 depicts the percentages of subjects in the audience and the private treatment who report to be better than average for different intervals of actual quiz performance, centered around the average of the comparison group (14.3 questions). Among subjects that clearly solved more questions than average (more than 15 correctly solved questions), 72 % report to be better than average in the audience treatment, compared to 69 % in the private treatment. For subjects with low quiz performance (less than 13 correctly solved questions), however, we have a very pronounced treatment difference. While 57 % report to be better than average in the audience treatment, only 27 % do so in the private treatment. This suggests, in line with our model, that our treatment effect is mainly driven by subjects who privately place themselves below average, but want to signal high performance towards others.

Additional, more indirect evidence that high reported self-assessments are associated with social approval comes from two survey questions we asked at the end of the experiment. First, we were interested in whether subjects enjoyed the quiz ("How much did you enjoy the quiz?"). Second, we asked whom subjects would hire if they were the boss of a firm on the basis of reported self-assessments. The three possible answers were: Somebody who reports 'better than average', 'worse than average', and 'I do not care'. We find that 64% of our subjects enjoyed the quiz or enjoyed it very much. Only 10% indicated they did not like the quiz. Also, none of the subjects was willing to hire a worker that reports "worse than average" in the audience treatment and only 13% would do so in the private treatment.

Figure 1.1: Percentage of "better than average" reports for high, low and close to average quiz performance subjects in the audience treatment and private treatment.



Result 2. *There is a gender difference in reported self-assessments: Men report "better than average" significantly more often. This difference seems to be driven by a stronger response of men to the presence of an audience.*

We find a gender difference in reported self-assessments. By inspection of Table 1.2 we find in regressions (2) and (3) that the probability to choose "better" is higher for men than women. The marginal effect of the gender dummy is significantly different from zero. A gender difference in self-assessments has been reported in many studies and provides a possible explanation for the gender difference in selection into competitive environments (see for example Gneezy and Rustichini (2004), Gneezy, Niederle, and Rustichini (2003) Niederle and Vesterlund (2007), Dohmen and Falk (2011)). Columns 4 and 5 of Table 1.2 show separate Probit regressions for the private and the audience treatment. The data indicate that the gender effect is mostly driven by more men overreporting in the audience treatment. While men report to be "better than average" significantly more often than women in the audience treatment, the effect is not significant in the private treatment. This finding might provide a possible explanation to gender differences in overconfidence. It suggests, that men feel a stronger desire to signal skills or abilities towards others which results in overconfident appearance.[12]

Result 3. *The public is aware of misreporting due to image concerns when evaluating subjects' reports.*

Does the audience anticipate that the report "better than average" might be driven by image concerns? To answer that question, we asked our subjects in the questionnaire about their perception and beliefs regarding the reported self-assessments of the other participants. We asked: "When subjects stated their self-assessment, do you think the other participants overestimated, underestimated or correctly estimated their performance?" Table 1.3 summarizes the answers. We find that a majority of subjects in the audience treatment (56%) thinks that others misreport and state too optimistic assessments. Only 26% hold a similar view in the private treatment. A Fisher-exact test confirms a significant difference ($p = 0.01$), where we categorize subjects' perceptions in "overreport" or not. Thus, we find evidence that the audience anticipates misreporting and adjusts beliefs accordingly. This finding supports the mechanism of our model. The decision maker chooses to signal a high self-assessment, the public anticipates this and adjusts beliefs about the decision maker's type downwards.[13]

1.4.2 Additional Control - Feedback Treatment

Our model predicts a treatment difference in reported self-assessments between the audience and the private treatment, which is caused by subjects' desire to signal a high type towards the other subjects in the audience treatment. To verify that signaling is really the key driver of our treatment effect, we conducted an additional

[12]Note, however, that this interpretation should be taken with caution. In Appendix 1.B we report the marginal effects of a probit regression with interactions of a gender dummy and treatment dummy (I_Treatment*Women). The marginal effect of this interaction describes the difference of the gender effect in the audience treatment compared to the private treatment. The difference is negative. In line with our interpretation, men report especially in the audience treatment that they are better than average, however not significantly more often than in the private treatment.

[13]Ludwig and Nafziger (2011) explore subjects' beliefs about other subjects' confidence bias and find that the majority believes that others are unbiased, and only few think that others are overconfident.

Table 1.3: Subjects' beliefs about the other participants' self-assessments

	Private treatment	Audience treatment
Overreport	26 %	56 %
Correct	40 %	42 %
Underreport	34 %	2 %

treatment that allowed us to control for the signaling of ability. The *feedback treatment* was identical to the audience treatment. The only difference was that after subjects reported their self-assessment to the audience, the experimenter informed the audience whether the assessment was correct or not. In this situation, the public learns the true relative performance and therefore subjects can no longer use their reports to signal ability. Thus, if misreporting in the audience treatment is driven by the signaling of skills or ability, the effect should vanish once we take the possibility to signal ability away. Consequently, we should observe that reported self-assessments do not differ between the private and the feedback treatment. 48 subjects participated in the feedback treatment and procedures were identical to the two other treatments.

Result 4. *Reported self-assessments do not differ significantly between the feedback treatment and the private treatment.*

We find that 56% of subjects choose to report "better than average". Compared to the private treatment with a frequency of 48%, there is no significant difference (p-value=0.54 using a Fisher exact test). Table 1.4 reports the marginal effects of probit regressions with and without controls for the private and feedback treatment. The treatment effect is insignificant in all regressions.

Result 4 provides additional evidence that image concerns, via the signaling of a high type, can lead to misreporting and overconfident appearance. While subjects used the opportunity to signal ability in the audience treatment, our treatment effect vanishes, once we take away the possibility to signal ability via our feedback treatment.

Table 1.4: Determinants of stated self-assessment in the private and feedback treatment

Dependent variable: Relative self-assessment= $\begin{cases} 1 & \text{if better} \\ 0 & \text{if worse} \end{cases}$

	(1)	(2)	(3)
Dummy treatment	-0.07	-0.03	-0.05
	(0.10)	(0.12)	(0.11)
Dummy gender		-0.09	-0.22*
		(0.13)	(0.12)
Quiz performance		0.11***	
		(0.03)	
Dummy quiz performance			0.21*
			(0.11)
Controls		included	included
N	95	95	95
-LL	65	48	54

Notes: Probit estimates. Marginal effects (evaluated at the mean of independent variables) reported; robust standard errors are in parentheses. Significance at the 1, 5, and 10 percent level is denoted by ***, **, and *, respectively. *Dummy treatment* =1 if feedback treatment and 0 if private treatment. *Dummy gender* =1 if female. *Dummy quiz performance* =1 if better than average. *Controls* include the survey based risk measure, image concerns, age, and relationship status.

1.5 Concluding Remarks

In this paper we studied the consequences of image concerns on reports of private information. We illustrated with a simple model how a desire for social approval can give rise to overconfident behavior. In addition, we conducted a controlled lab experiment that supports predictions of our model. In the experiment, subjects stated a higher self-assessment when an audience is present than in private. We also find that men choose more often than women to signal ability and confidence especially when an audience is present.

Our findings show that biases in self-assessments might be produced by image concerns. As a consequence, decision makers can appear overconfident even with perfect knowledge about their own performance, skill or ability, in other words, without inherent biases in self-assessments. This is also an explanation why overconfidence is persistent. Receiving feedback and learning one's type over time might not prevent decision makers from appearing overconfident.

In our experiment, we manipulated image concerns by letting subjects report their self-assessment to an audience. The audience was mainly composed of students that did not know each other and thus social distance between decision makers and the public was rather high. We expect that in more intense social contexts, e.g. talking to one's supervisor, boss, parents or friends, the magnitude of our finding might be even larger. Furthermore, we did not provide direct strategic reasons for image or reputational concerns. An interesting set-up to implement an instrumental value of appearing skilled or able would be as follows: subjects would randomly be assigned to the roles of principals and agents. In each session there would be twice as many agents as principals. Agents would go through our quiz questions and then (anonymously) state a self-assessment towards the principal assigned to them. The principal has to select one of the two agents for an additional quiz and has incentives to select the agent he thinks is most able. Agents would be given incentives for being selected. We suspect that agents would overstate self-assessments to increase the likelihood of being selected by the principal. Therefore, stated self-assessments in such a treatment should be higher compared to our control treatment.

While the main focus of the paper is on social image concerns, our model is also compatible with a self-signaling interpretation. Instead of signaling skill or confidence to others, decision makers care about how they perceive their own self. In this interpretation, self-image is built from past actions. While beliefs about performance are available when making choices, later self-evaluation is built on past actions because actions are easier to recall than beliefs (e.g., Bénabou and Tirole (2004, 2006)). This is important, as much of the evidence on overconfidence has been in rather (though usually not perfect) private environments. In such environments, self-signaling (as well as signaling towards the experimenter) might be a key driver of overconfidence. Although this is not explicitly modeled in our framework, the self-signaling interpretation might give rise to inherent biases in self-assessment. Interestingly, these biases would not stem from selective choice of information or asymmetric information processing (like for example in Brunnermeier and Parker (2005), Kőszegi (2006) or Möbius et al. (2011)) but from self-evaluation based on biased past choices.

Finally, while the main focus of this paper is on direct transmission of private information, our results apply more generally. In many choice environments that require prior self-assessment, decision makers' choices allow the signaling of skill, ability or performance towards others. Consider the choice to enter a tournament. The decision to enter or not clearly depends on individuals' private self-assessment. The money-maximizing choice for individuals with low skills and abilities is usually not to enter the tournament. In the presence of image concerns, however, individuals with low skills might yet decide to enter, as this allows them to signal skill and abilities to others. In the context of participation in welfare programs, image concerns and social approval seeking might lead to low participation rates due to fear of reputation losses. Moffitt (1983) presents data from different welfare programs in the U.S. in the 1970's. He reports that as much as 30 - 60 % of the citizens who are eligible for welfare do not apply and argues that this is a consequence of the fear of stigmatisation of welfare recipients.

1.6 Appendix

Appendix 1.A

Proof of Proposition 1:

In equilibrium, all D with $p > \overline{r}$ choose $x = Yes$. It is straightforward to show that this is optimal, given that it maximizes both monetary outcomes and image utility. For decision makers with $p < \overline{r}$, behavior depends on the strength of image concerns. There exists a threshold type α^*, such that all D with $p < \overline{r}$ and $\alpha > \alpha^*$ will choose $x = Yes$ and those with $p < \overline{r}$ and $\alpha < \alpha^*$ choose $x = No$. The threshold type α^* with $p < \overline{r}$ must be indifferent between potential image gains from choosing $x = Yes$ and monetary losses from reporting incorrectly. We have the following indifference condition:

$$\alpha^*\beta\left[\sum_{p>\overline{r}} f(p)p + \int_{\alpha^*}^{\overline{\alpha}} g(z)dz \sum_{p\leq\overline{r}} f(p)p\right]\frac{1}{\sum_{p>\overline{r}} f(p) + \int_{\alpha^*}^{\overline{\alpha}} g(z)dz \sum_{p\leq\overline{r}} f(p)} \tag{1.1}$$
$$= y + \alpha^*\beta\sum_{p\leq\overline{r}}\frac{f(p)p}{\sum_{p\leq\overline{r}} f(p)}.$$

The left hand side captures image utility in case D chooses $x = Yes$, which is simply a weighted average of the average performance, skill or ability of decision makers with $p > \overline{r}$ and those with $p < \overline{r}$, with weights depending on how many Ds misreport. The right hand side captures image utility when choosing $x = No$, which is simply the average performance or ability of Ds with $p < \overline{r}$ plus the prize y for reporting correctly. Rearranging equation 1.1 leads the following:

$$\alpha^*\beta\left[\frac{1}{\sum_{p>\overline{r}} f(p) + \int_{\alpha^*}^{\overline{\alpha}} g(z)dz \sum_{p\leq\overline{r}} f(p)}\left(\sum_{p>\overline{r}} f(p)p + \int_{\alpha^*}^{\overline{\alpha}} g(z)dz \sum_{p\leq\overline{r}} f(p)p\right)\right. \tag{1.2}$$
$$\left. - \sum_{p\leq\overline{r}}\frac{f(p)p}{\sum_{p\leq\overline{r}} f(p)}\right] = y.$$

One can see from equation 1.2 that decision makers with $\alpha < \alpha^*$ and $p < \overline{r}$

optimally choose $x = No$. As the expression in square brackets remains unchanged but the strength of image concerns is smaller ($\alpha\beta < \alpha^*\beta$), image gains in total are smaller than monetary losses, i.e., they will state a truthful report $x = No$. D's with $\alpha > \alpha^*$ instead optimally choose $x = Yes$ as their image gains loom larger than their monetary losses. Note also, that if $\overline{\alpha}$ is sufficiently large, the threshold type α^* and thus the equilibrium, always exists. To see this, take the left hand side of equation 1.2 and vary α^*. If α^* approaches zero, the left hand side approaches zero as well. As α^* approaches $\overline{\alpha}$, the left hand exceeds y by assumption. Furthermore, the left hand side is continuous and strictly increasing in α^*. Consequently, there necessarily exists an α^* for which equation 1.2 holds.

To show uniqueness of the equilibrium, first note that in every equilibrium, the types with very low image concerns ($\alpha \to 0$) will always choose the money-maximizing outcome, i.e., those with $p < \overline{r}$ optimally choose $x = No$ and those with $p > \overline{r}$ optimally choose $x = Yes$. Next we show that there cannot be an equilibrium where decision makers with $p > \overline{r}$ do not choose $x = Yes$. Suppose there would be such an equilibrium. Then the image utility from choosing $x = No$ necessarily would need to be greater than the image utility from $x = Yes$. In that case however, all D with $p < \overline{r}$ would also choose $x = No$. This leads to a contradiction because then the public will infer a lower p from $x = No$ than from $x = Yes$ and consequently image utility from $x = Yes$ would be higher. Thus in every equilibrium, some D with $p < \overline{r}$ and low values of α will choose $x = No$ and all D with $p > \overline{r}$ choose $x = Yes$. Also, by assumption $\overline{\alpha}$ is large enough such that some D with $p < \overline{r}$ choose $x = Yes$. From that it is easy to see that every equilibrium has a threshold type α^*, such that decision makers with $p < \overline{r}$ and $\alpha > \alpha^*$ will choose $x = Yes$ and those with $\alpha < \alpha^*$ will choose $x = No$. From equation 1.2 we see that α^* and consequently the equilibrium described above is unique.

Proof of Proposition 2:

The proof is straightforward. Considering equation 1.2, one can see that a change in β affects the threshold type α^*. An increase in β reduces the threshold type, in other words, more decision makers with $p < \overline{r}$ will choose $x = Yes$.

Model with Imperfect Knowledge:

So far, we assumed that decision makers perfectly know their p. However, one could argue that in most real-life situations, individuals only have imperfect knowledge about their skills or abilities. Also, in our experiment subjects are likely to be uncertain about their performance. In this section, we analyze what happens if decision makers have imperfect knowledge about their type but know more than the public. The crucial difference to the case with perfect knowledge is that type-uncertainty weakens the informativeness of decision makers choices. Intuitively, it is more difficult for the public to infer ability from choices, if decision makers themselves are uncertain about their ability.

The set-up is identical to above. The only difference is that decision makers do not perfectly know their p. Instead, they hold a point belief $\hat{p} \in \{0, 1, ..., \bar{p}\}$ and $\hat{p}$ is (potentially) different from p.[14] D's choice x is again to report whether p is larger than some value $\bar{r}$, i.e., $x \in \{Yes, No\}$. Given their imperfect knowledge about p, it is possible that decision makers wrongly assess whether their p is larger or smaller than $\bar{r}$. We specify the imperfect knowledge about p as follows. Let $\phi(p)$ denote the likelihood that decision makers point belief $\hat{p}$ is larger (smaller) than $\bar{r}$ although the true p is smaller (larger). Thus $\phi(p)$ is the probability that $\hat{p} > \bar{r}$ although $p < \bar{r}$ or $\hat{p} < \bar{r}$ although $p > \bar{r}$. We make the following assumptions about $\phi(p)$. First of all, we naturally assume that $\phi(p) < \frac{1}{2}$ for all p. Second, we assume that $\phi(p)$ is strictly increasing in p for $p < \bar{r}$, and strictly decreasing in p for $p > \bar{r}$. In other words, the likelihood that Ds think that their p is larger (smaller) than $\bar{r}$, although it is smaller (larger) increases the smaller the difference between p and $\bar{r}$.

We now show that decision makers still have incentives to misreport their private information $\hat{p}$. The key difference between a set-up with imperfect knowledge and one with perfect knowledge is, that the public's inference about performance from choices x changes. Since the public is aware that decision-makers only have imperfect knowledge about their performance, the informativeness of reports x about performance p is reduced. However, the informativeness does not vanish. One can show that if all decision makers report truthfully, i.e. they maximize monetary utility in

[14]To focus on the effect of type uncertainty on the informativeness of choices, we abstract from risk by assuming point beliefs about ability.

the absence of image concerns, the public infers higher ability from reports $x = Yes$ compared to reports $x = No$, that is $E(p \mid x = Yes) > E(p \mid x = No)$. We have that $E(p \mid x = Yes) = \frac{\sum_{p>\overline{r}}(1-\phi(p))f(p)p+\sum_{p<\overline{r}}\phi(p)f(p)p}{\sum_{p>\overline{r}}(1-\phi(p))f(p)+\sum_{p<\overline{r}}\phi(p)f(p)}$ is greater than $E(p \mid x = No) = \frac{\sum_{p<\overline{r}}(1-\phi(p))f(p)p+\sum_{p>\overline{r}}\phi(p)f(p)p}{\sum_{p<\overline{r}}(1-\phi(p))f(p)+\sum_{p>\overline{r}}\phi(p)f(p)}$.

Thus, we can state the following proposition:

Proposition 3. *If $\overline{\alpha}$ is sufficiently large, there exists a unique Perfect Bayesian Equilibrium where decision makers with $\hat{p} < \overline{r}$ and $\alpha > \alpha^*$ choose $x = Yes$. Decision makers with $\hat{p} > \overline{r}$ choose $x = Yes$ and those with $\hat{p} < \overline{r}$ and $\alpha < \alpha^*$ choose $x = No$.*

Proposition 3 corresponds to Proposition 1 in the set-up with perfect knowledge.[15] The condition for $\overline{\alpha}$ being sufficiently, however, is more demanding compared to the perfect knowledge case. $\overline{\alpha}$ needs to be large enough such that the image gains from choosing $x = Yes$, $\overline{\alpha}\beta\left[\frac{\sum_{p>\overline{r}}(1-\phi(p))f(p)p+\sum_{p<\overline{r}}\phi(p)f(p)p}{\sum_{p>\overline{r}}(1-\phi(p))f(p)+\sum_{p<\overline{r}}\phi(p)f(p)} - \frac{\sum_{p<\overline{r}}(1-\phi(p))f(p)p+\sum_{p>\overline{r}}\phi(p)f(p)p}{\sum_{p<\overline{r}}(1-\phi(p))f(p)+\sum_{p>\overline{r}}\phi(p)f(p)}\right]$ outweigh the monetary costs y. The reason that this condition is more demanding than that in the case of perfect information is that type uncertainty reduces the reputational gains from choosing $x = Yes$. Therefore image concerns need to be higher in the case of imperfect knowledge of own type. Proposition 3 shows that also with imperfect knowledge, decision makers have incentives to misreport private information. The intuition is simple. Although decision makers are not perfectly informed about their own skills, performance or ability, they know more than the public. Consequently reports x have some informative value for the public and thus the signaling motive for decision makers still exists.

For variations in common image utility β, the same comparative statics hold as in section 2.2.2.

Proposition 4. *An increase in β reduces the threshold type α^*. Consequently, more decision makers with $\hat{p} < \overline{r}$ misreport by choosing $x = Yes$.*

[15]The logic of the proof is the same as for Proposition 1.

Appendix 1.B

Table 1.5: Determinants of relative self-assessment in the private and audience treatment with interactions

Dependent variable: Relative self-assessment= $\begin{cases} 1 & \text{if better} \\ 0 & \text{if worse} \end{cases}$	
Dummy treatment	1.36
	(1.39)
Gender dummy	-0.19
	(0.16)
I_Treatment*Gender dummy	-0.23
	(0.28)
Quiz performance	0.06*
	(0.03)
I_Treatment*Quiz performance	0.05
	(0.06)
Controls	included
N	95
-LL	42

Notes: Probit estimates. Marginal effects reported; robust standard errors are in parentheses. Significance at the 1, 5, and 10 percent level is denoted by ***, **, and *, respectively. *Dummy treatment* =1 if audience treatment and 0 if private treatment. *Dummy gender* =1 if female. *Controls* include the survey based risk measure, image concerns, age, relationship status, and interactions of the *Dummy treatment* with each variable.

Appendix 1.C

Instructions, translated into English. General instructions and instructions for the first part of the experiment were identical across treatments. Instructions for the second part of the experiment differed across treatments.

GENERAL INSTRUCTIONS

You are taking part in a decision-making experiment in which you have the opportunity to earn money. The amount of money you earn is paid to you upon completion of the experiment. Please read the instructions carefully. The instructions are identical for all participants. If you have any questions, please raise your hand. The experimenter will answer your question at your place. During the experiment, you have to remain silent. Violation of this rule leads to immediate exclusion from the experiment and all payments.

All monetary units in the experiment are measured in tokens, and 100 tokens = 1 euro.

This experiment consists of two parts. In both parts, you can earn money. Your payoff from the experiment results from the sum of your payoffs in both parts. In the following we will go through the instructions for the first part of the experiment. After the first part is completed, we will provide you with the instructions of the second part.

INSTRUCTIONS FOR THE FIRST PART OF THE EXPERIMENT

In the first part of the experiment you will be asked 20 quiz questions. You will always be offered 4 possible answers of which exactly one will be correct. Please always select one of the four possible answers. For each correct answer you get 40 tokens. After you have answered the first 10 questions, please click on the OK button. Then a new screen with 10 more questions will appear. Please confirm your responses again with the OK button.

Do you have any questions?

INSTRUCTIONS FOR THE SECOND PART OF THE EXPERIMENT - *(Private Treatment)*

All participants have answered 20 quiz questions in the first part of the experiment. In this part of the experiment, you need to assess whether your quiz result is better or worse than the average result of another group of participants. If your assessment is correct, you get 500 tokens; if your assessment is wrong, you get 0 tokens. This will be further explained below in more detail.

The quiz questions you were asked in the first part of the experiment, were also answered by a group of 95 participants (all of which (like you) participated in an experiment in the BonnEconLab) some time ago. You now need to assess whether your performance in the quiz was better or worse than the average performance of the group of 95 participants. You get 500 tokens for a correct assessment, otherwise you get 0 tokens.

Please read these instructions again carefully.

An input box appears soon on your screen into which you can enter your decision.

Do you have any questions?

INSTRUCTIONS FOR THE SECOND PART OF THE EXPERIMENT - *(Audience Treatment)*

All participants have answered 20 quiz questions in the first part of the experiment. In this part of the experiment, you need to assess whether your quiz result is better or worse than the average result of another group of participants. If your assessment is correct, you get 500 tokens; if your assessment is wrong, you get 0 tokens. This will be further explained below in more detail.

Note the following: After all participants entered their assessment into the computer, all participants must report their assessment to the other participants. Every participant will be called up individually one after the other. Once it is your turn, you have to stand up, say your name and report your assessment.

So if you stated that you think your quiz result was better than the average of the other group, then you have to stand up after you were called and say: "My name is ... and I think I was better than the average of the other group."

If you stated that you think your quiz result was worse than the average of the other group, then you have to stand up after you were called and say: "My name is ... and I think I was worse than the average of the other group."

Below we will explain your decision in more detail.

The quiz questions you were asked in the first part of the experiment, were also answered by a group of 95 participants (all of which (like you) participated in an experiment in the BonnEconLab) some time ago. You now need to assess whether your performance in the quiz was better or worse than the average performance of the group of 95 participants. You get 500 token for a correct assessment, otherwise you get 0 token.

Please read these instructions again carefully.

An input box appears soon on your screen into which you can enter your decision.

Do you have any questions?

INSTRUCTIONS FOR THE SECOND PART OF THE EXPERIMENT - *(Feedback Treatment)*

All participants have answered 20 quiz questions in the first part of the experiment. In this part of the experiment, you need to assess whether your quiz result is better or worse than the average result of another group of participants. If your assessment is correct, you get 500 tokens; if your assessment is wrong, you get 0 tokens. This will be further explained below in more detail.

Note the following: After all participants entered their assessment into the computer, all participants must report their assessment to the other participants. Every participant will be called up individually one after the other. Once it is your turn, you have to stand up, say your name and report your assessment. Also note: After you have reported your assessment, the experimenter will tell you and the other participants, whether your quiz result was actually better or worse than the average score of the other group.

So if you stated that you think your quiz result was better than the average of the other group, then you have to stand up after you were called and say: " My name is ... and I think I was better than the average of the other group." If your quiz result was indeed better than the average performance of the other group, the experimenter will announce: "The quiz result of Mr. / Ms. XY was better than the average score of the other group." If your quiz result was indeed than the average performance of the other group, the experimenter will announce: "The quiz result of Mr. / Ms. XY was worse than the average score of the other group."

If you stated that you think your quiz result was worse than the average of the other group, then you have to stand up after you were called and say: "My name is ... and I think I was worse than the average of the other group." If your quiz result was indeed worse than the average performance of the other group, the experimenter will announce: "The quiz result of Mr. / Ms. XY was worse than the average score of the other group.” If your quiz result was indeed better than the average performance of the other group, the experimenter will announce: ”The quiz result of Mr. / Ms. XY was better than the average result of the other group.”

Below we will explain your decision in more detail.

The quiz questions you were asked in the first part of the experiment, were also answered by a group of 95 participants (all of which (like you) participated in an experiment in the BonnEconLab) some time ago. You now need to assess whether your performance in the quiz was better or worse than the average performance of the group of 95 participants. You get 500 token for a correct assessment, otherwise you get 0 token.

Please read these instructions again carefully.

An input box appears soon on your screen into which you can enter your decision.

Do you have any questions?

Chapter 2

Information and Competition Entry

2.1 Introduction

Competition is a major force of economic behavior and interactions. Examples of competition are business formation, job promotion, occupational choice, or sports tournaments. The decision to enter a competition clearly depends on individuals' private self-assessment about performance. However, absolute and relative self-assessments about performance are often inaccurate (e.g. Weinstein (1980), Taylor and Brown (1988)). As a consequence of overconfident self-assessments and neglecting the performance of the opponents, Camerer and Lovallo (1999) find excessive market entry in a laboratory experiment. Too many subjects entered the market and therefore the market share is lower than an outside option. Similarly, *competition failure* often yields less monetary utility than an outside option, where competition failure occurs if an individual loses the competition, because the opponent holds a higher performance. For example, new businesses frequently fail after inception as a result of overconfidence and entrepreneurs earn less money than in a paid job according to their performance (e.g. Koellinger, Minniti, and Schade (2007)).

The goal of our study is to explore whether information influences entry decisions in a competition game. We ask whether information on the opponents reduces competition failure by preventing entry of overconfident individuals and by attract-ing mainly high performance types. Competition failure is an especially interesting criteria, because it informs us about the frequency of individuals that waste money due to losing the competition instead of choosing a higher outside option. Evidently, information is beneficial in various contexts for the decision making process. In the competition game, information can nudge individuals to update their performance beliefs and think more carefully about the complexity of the decision problem and the chances of success. In consequence, we expect fewer competition failures. However, the updating process might not be correctly applied and it is possible that information even fosters the overconfidence bias. We briefly discuss the effects of absolute and relative self-assessment biases on entry choices in the competition game and when information might be beneficial. The latter results as an outcome of the magnitude of self-assessment biases, individuals' updating process and it's interactions, which we explore empirically with a lab experiment.

For the ideal empirical analysis of the effect of information on entry choices, we need to have control over the information available to each individual. In addition, we need clean measures of individuals' performance and self-assessments about performance. Therefore, we make use of the advantages of a controlled laboratory experiment instead of a field experiment. In the competition game, success depends on the performance of an ex ante performed quiz task. We elicit absolute and relative self-assessments about subjects' quiz performance. Subjects are randomly matched in groups of two and decide about entering the competition or opting for an outside payment. A subject wins the competition, if the matched opponent did not enter, or if he or she has a higher performance than the opponent. We set up a benchmark treatment, *No Info,* in which subjects receive no additional information and three information treatments with a between-subject design. In treatment *Distribution*, the performance distribution of all subjects in the session is revealed, before subjects made their decision to enter. In treatment *True*, we reveal more precise data on the matched opponent, that is the matched opponent's true performance. To study whether there is a discouraging or encouraging impact of the opponent's over- or underconfidence, we study in treatment *True & Belief*, the entry choice by disclosing information on the matched opponent's performance and his or her absolute performance self-assessment. This variation of aggregated and precise information allows us to exactly measure the frequency change of entry choices and competition failure rates for different types of information disclosure.

Competition entry in the information treatments is significantly lower compared to the benchmark treatment. We find evidence for competition failure especially in the benchmark treatment, *No Info*, which is mainly driven by overplacement, i.e. neglecting the performance of the opponents (similar to Camerer and Lovallo (1999)), instead of overestimation, gender or willingness to take risks. Analyzing the data of the information treatments reveals sizable and significant improvements by information. The two major findings are firstly, competition failures decrease by 57%, when providing information on the performance distribution and secondly, more precise information does not further improve entry choices. This implies that simple and aggregated information on the performance distribution is sufficient to decrease competition failure rates by a striking value. Our data show also that the decision to enter depends strongly on performance in the information treatments, but not in the benchmark treatment. This complements the usefulness of information disclosure to attract high performance individuals for the competition. Disclosing the opponent's performance has a high and significant influence on entry. In addition, the knowledge of an overconfident opponent discourages entry, but not in a significant way. Note that we do not analyze a strategic choice of self-assessments on the opponent's entry choice. Reuben et al. (2010) show that exaggerating one's self-assessment strategically in a team environment, helps to become the leader of a team. And Charness, Rustichini, and van de Ven (2011) find a similar result for a competition environment where the disclosure of the opponent's high self-assessment discourages competition entry. However, both papers do not disclose the opponent's true performance at the same time when they disclose the self-assessment. Furthermore, we confirm the finding of Niederle and Vesterlund (2007), who point out that due to a gender difference in overconfidence, women shy away from competitive environments more frequently compared to men, in the benchmark treatment and treatment *Distribution*. The gender difference in competition entry vanishes in the treatments where the performance of the matched opponent is disclosed.

A large literature in psychology and experimental economics emphasizes the finding of self-assessment biases (e.g. Svenson (1981), Weinstein (1980)).[1] The consequences have been studied in various economic environments like business contexts. For example, Cooper, Woo, and Dunkelberg (1988) conclude that entrepreneurs overestimate their chances of success with their new business, which in consequence leads to competition failure. Also, Dunning, Meyerowitz, and Holzberg (1989) and Baldwin (1995) report business failures shortly after market entry. For a similar survey study and a recent overview see Koellinger, Minniti, and Schade (2007). In addition, overconfidence has been highlighted as a major force in costly delays in labor negotiations, excessive litigation, excessive stock trading and subsequent market volatility, (see, e.g. Neale and Bazerman (1985), Odean (1998), Daniel, Hirshleifer, and Subrahmanyam (2001), Malmendier and Tate (2008)). Our findings contribute to this literature by showing how consequences of self-assessment biases can be mitigated in a competition environment. In addition, our study contributes to the literature on sorting behavior in competitive environments. A competition is usually set up to attract high performance types. Dohmen and Falk (2011) conclude that in addition, relative self-assessments, gender, and willingness to take risks are vital personal attitudes that effect competition entry choices when studying decision making of students in a laboratory experiment and also of a representative sample of the German population. Experimental studies on entry decisions by Camerer and Lovallo (1999), Niederle and Vesterlund (2007), Bartling et al. (2009) also show that subjects with high relative self-assessments self-select into the competition more frequently.

In the last years, a growing literature on libertarian paternalism by psychologists and economists aims at encouraging and supporting individuals in economic and non-economic decision finding (e.g. Thaler and Sunstein (2003, 2008)). Examples are the analysis of optimal default options (e.g. Choi et al. (2003)) or school interventions. Our study contributes to this literature by showing that providing simple and inexpensive information helps our subjects to make better decisions in the competition game. Our controlled laboratory findings might have similar effects in real life situations of competitive environments. An established application where information

[1] Several explanations, for why self-assessment biases are present in many contexts and still persist have been proposed, for instance self-image concerns (Bénabou and Tirole (2002), Kőszegi (2006)), asymmetrical processing of positive and negative information (e.g. Brunnermeier and Parker (2005), Möbius et al. (2011)), or social image concerns (e.g. Burks et al. (2010), Ewers and Zimmermann (2011)).

of past performances is used to decrease allocation and application costs, is the university place allocation system *ZVS* in Germany. Here, prospective students learn the distribution of school grades of former accepted students for the respective field and university of the last year *before* applying. Other possible implementations in practice might be interventions at employment agencies or firms. For instance, public and private employment agencies could emphasize the importance of the opponents performance in a competition by providing information about market characteristics, where startup businesses want to engage in. Firms often collect data about workers' performance, effort, cognitive and non-cognitive skills. Before announcing (internal) job-promotions, firms could disclose anonymously the performance outcomes of former workers to reduce irrelevant applications. Banks or venture capital companies, that finance credits for startup businesses could disclose probabilities of success for the market segment.

The remainder of the paper is organized as follows. In Section 2.2, we present the experiment with the design and hypotheses. Section 2.3 provides the results. Finally, Section 2.4 concludes.

2.2 Experiment

In the ideal experimental set up for a clean analysis of competition entry decisions, we need to know subject's beliefs about their own and relative performance, their true performance, and further individual characteristics like willingness to take risks and gender. We conduct our experiment in the laboratory instead of conducting a field experiment to secure control about the information each subject has and we can precisely manipulate and vary the information provision. Confounding factors that might influence subjects beliefs are ruled out and additional information about further key variables is available, e.g. risk attitudes and entry costs.

2.2.1 Experimental Design

We present now the four main parts of our experiment. In the first part, we pin down subjects performance for the competition. In part two, subjects assess their absolute and relative beliefs about performance. In part three, subjects play the competition entry game and receive information or not about the opponent previous to their entry choice. In part four we analyze subjects' willingness to take risks. Finally, the experiment ends with a general questionnaire after the risk task.

In the first part, subjects' performance is determined by a multiple-choice quiz of 20 questions and no time limit. The quiz contained questions concerning history, arts, economics, and orthography. Subjects received 20 points for a correct quiz answer. All monetary quantities of the experiment are denominated in points; 100 points are equal to one euro. Subjects did not receive feedback on the number of correct quiz answers or the amount of earned money. Thereby, subjects could not learn, and will never learn during the experiment, their true quiz performance, where quiz performance is defined as the sum of correct quiz answers. We decided for the quiz task instead of a real effort task, because we are mainly interestes in the effect of information on entry decisions and not on performance changes. While the influence of information and feedback on performance changes is worthy to study, we want to exclude these additional impacts to gain control about the pure effect of information on entry decisions in a competition. The performance in the quiz is only determined by knowing the answers and exerting effert is a minor factor.

We employed five incentivized self-assessment questions to elicit subjects' absolute and relative self-assessments on quiz performance. Not all five, but only one measure was paid to impede any hedging motives. We studied different questions to receive overestimation and overplacement measures and to perform consistency and robustness checks. A subject exhibits overestimation (underestimation) if her absolute self-assessment is better (worse) than the true quiz performance and a subject exhibits overplacement (underplacement) if her relative self-assessment is better (worse) than the true relative quiz performance.[2] The first two questions were used to analyze overestimation: (i) How many quiz questions have you solved correctly? (ii) For the second overestimation measure, subjects had to distribute 100 points into 21 categories. Every category was associated with the number of correct quiz answers and all points had to be distributed.[3] Our three questions on overplacement are: (iii)

[2]The studies by Healy and Moore (2008a) and Healy and Moore (2008b) point out that the distinction of overconfidence in overestimation and overplacement is crucial for the decision making process and much of the previous literature confuses these two concepts.

[3]For example, a subject that is sure to have answered more than one question correct should place zero points into category "0" and "1".

Is your amount of correct quiz questions one of the best 12 or worst 12 quiz performances in the room?, (iv) How many of the other participants in the room solved more quiz questions correct than you?, (v) How many of the other participants in the room have less correct answers than you? A subject's payoff for (iii) was 100 points for a correct assessment. Subject's payoff for (i), (ii), (iv) and (v) was 200 points for a correct assessment and 50 points for guessing one category next to the correct answer. We define overconfidence and underconfidence by calculating the difference between the actual quiz performance and the five statements (see Table 2.5 in the Appendix 2.A for more details).

After subjects' quiz assessment, we conducted the competition game in the third part of the experiment. We employed a between subject design for the study of four treatments including a benchmark treatment and three information treatments. Subjects were randomly matched in pairs of two and decided to enter the competition or not. Before they made their decision, we provided information according to the treatment they participated in. In the benchmark treatment, *No Info,* they received no additional information. In the *Distribution* treatment, we provided information about the quiz performance distribution of all subjects in the room by showing a table with the amount of subjects that had 0, 1, 2,...,20 questions correct on the screen. In treatment *True*, each subject received information about the quiz performance of his or her matched opponent and in treatment *True & Belief,* each subject received information on the matched opponent's quiz performance and, in addition, the performance belief from the self-assessment question (i) (see also Table 2.1 for an overview of all treatments). Note, that this was not known to subjects when they answered the self-assessment questions.[4] We can then study exactly what kind of information reduces or increases competition entry and competition failure.

The rules of the competition game are the following: If a subject does not enter the competition, he or she receives an outside option of 200 points. If only one subject enters the competition, he or she wins the competition automatically and receives the winner prize of 400 points. If both subjects enter the competition, the subject with the highest quiz performance wins the competition and the loser receives the loser

[4]For strategic choices of reports of self assessments see for example Ewers and Zimmermann (2011) or Charness, Rustichini, and van de Ven (2011).

prize of 100 points. The loser prize is of course smaller than the outside option, otherwise all subjects enter the competition. If subjects with the same performance enter, a subject wins either the winner or loser prize with probability 0.5. Thus, subjects' performance of the previous task is relevant for the probability of winning the competition. The only action of the subjects is to decide about the entry decision. Subjects do not perform another quiz task. To secure the understanding of the game, we asked several control questions before the game started.

The experiment ended with a task to elicit risk attitudes. Subjects made 30 decisions between a lottery and a secure payoff. The lottery was always the same: It provides a 50% chance to win 400 points and a 50% chance to win 100 points. The secure payoff increased from 0 to 400 points. A subject's switching point was used as an indicator for her willingness to take risks.

All sessions of the experiment were conducted at the BonnEconLab at the University of Bonn, subjects were recruited via ORSEE (Greiner (2004)) and we used the software z-Tree by Fischbacher (2007). We conducted eight sessions in March 2010 with 190 subjects from various fields of study and tried to have an equal amount of women and men in every session to analyze a gender effect. Subjects answered all questions and tasks at the computer. At the beginning of the experiment, all subjects knew that the experiment consists of four parts and that they receive the instructions of each of the four parts individually and immediately before the task started. The average duration of a session was 50 minutes and the average payoff was 9 euros.

2.2.2 Overconfidence Hypothesis and Information Efficiency

If subjects hold a systematic bias about their absolute or relative quiz performance, they might over- or undervalue their subjective utility of the competition game. In consequence, overconfident subjects enter too often and underconfident subjects respectively enter too rarely which leads to our hypothesis on entry behavior that we explore with the laboratory experiment:

Hypothesis. *If a subject is overconfident (underconfident), then she enters the competition more (less) often than an unbiased subject holding everything else con-*

Table 2.1: Information treatments

Treatment	Information
No Info	No additional information
Distribution	Information about the performance distribution of subjects in the room
True	Information about the opponent's performance
True & Belief	Information about the opponent's performance and performance belief

stant.

The main purpose of this paper is to detect what information is a simple and inexpensive way to decrease competition failure. Information is obviously a strong tool to improve decision making in various contexts. Receiving information on the opponent reduces uncertainty in the game and attracts the attention to the fact that the decision problem depends highly on the opponents performance. In addition, individuals can update their self-assessment, which might reduce the self-assessment bias. However, in our two player game both subjects receive information at the same time. Therefore, additional information might not necessarily be beneficial. The way how individuals update their self-assessment is also crucial. Whether an overconfident individual updates her belief downwards or rather fosters the overconfident belief has an influence on the entry choice. Furthermore, the investigation of self-assessment biases does always entail a discussion of higher order beliefs. The belief about what others believe is of interest in this two player game, too. E.g., if a player assumes that other players are underconfident, she will enter the competition too often.

If subjects hold perfect beliefs about their performance and the performance distribution, we expect no difference in entry decisions by providing information about performance distribution in treatment *No Info*. However, the literature on self-assessment biases emphasizes that many individuals hold too optimistic beliefs about their absolute and relative performance. On the one hand, knowing the distribution about performance reduces uncertainty and we expect less wrong entry decisions. On the other hand, it is very important how individuals actually use their new information to update beliefs. For example, a player with a correct absolute self-assessment and an overconfident relative self-assessment should optimally update the relative self-assessment by updating it downwards. An overconfident player might instead prefer to update her actually correct absolute self-assessment upwards.

In a similar vein, for unbiased subjects, the effect of information on the opponent's performance in treatment *True* should result in efficient entry decisions. The player with the higher performance always enters and the other one chooses the outside option. However, biased subjects might not choose efficiently.

In addition, we study how the entry choice is influenced by over- and underconfidence of the opponent by disclosing the true performance of the opponent, and also, we disclose the performance self-assessment of the opponent in treatment *True & Belief*. The information of an overconfident opponent indicates that the opponent enters very certainly and might discourage a player to enter. Or an underconfident opponent encourages the entry decision although the opponent has a higher performance.

Obviously, we need to make several assumptions on 1. absolute and relative self-assessment biases, 2. belief updating and 3. higher order beliefs, to claim that information is beneficial. An analysis of these channels and its interaction is an interesting task to deepen the understanding of individuals' processing with uncertainty. However, it is not the focus of our paper. We will analyze the effect of the information on entry choices empirically within the experiment.

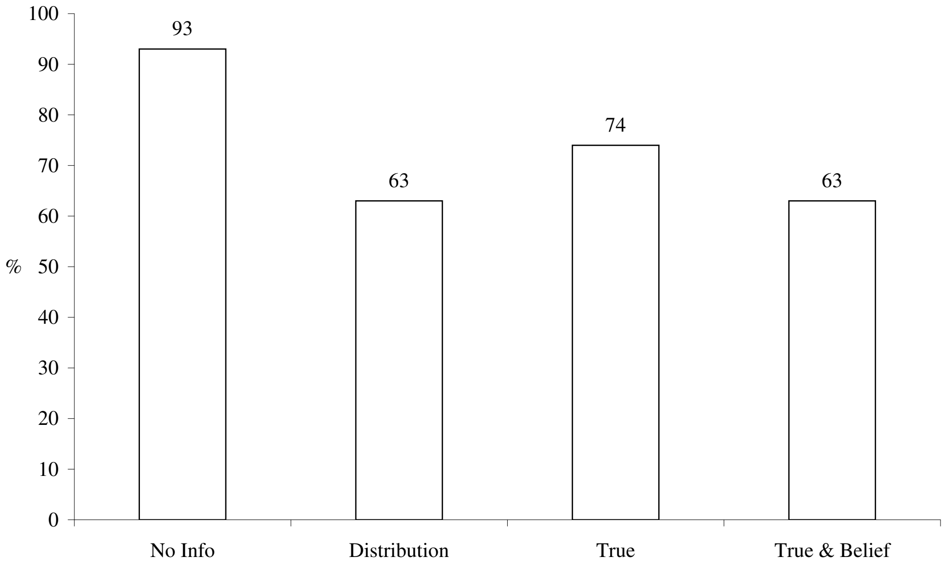
2.3 Results

We first present results on the entry frequency across treatments and the effect of information on competition failure. All tests we use are two-sided. In 2.3.2 we present the determinants of entry on performance and personal attitudes and in 2.3.3, we discuss how overconfidence is distributed and correlated to personal attitudes.

2.3.1 Entry Frequency and Information Efficiency

Result 1. *The frequency of entry choices decreases significantly with information compared to the benchmark treatment.*

Figure 2.1: Percentage of competition entry by treatments



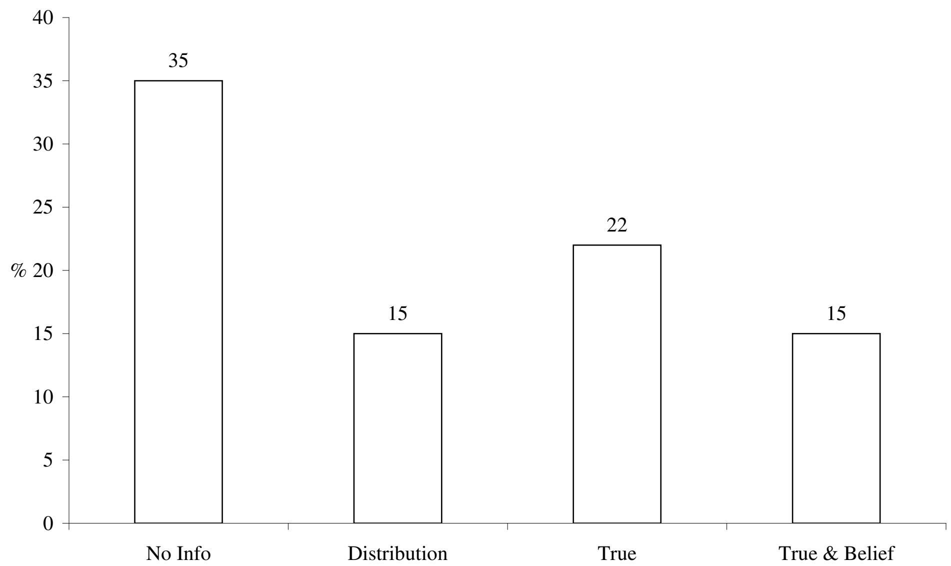
In the benchmark treatment, *No Info,* a striking number of 93% of our subjects chose to enter the competition, in treatment *Distribution,* only 63% entered, in treatment *True* 74% and 63% in treatment *True & Belief*, respectively (see Figure 2.1).[5] The large amount of entry decisions in the benchmark treatment shows that it is rather a minor problem that underconfident, but high performance subjects do not enter the competition. The difference of the benchmark treatment compared to the information treatments is sizable. We test whether the large differences in entry frequencies are significant using Fisher-exact tests and find that the entry frequency of each information treatment is significantly different from the benchmark treatment with a $p-value = 0.001$ for either *Distribution* or *True & Belief* as comparison and

[5] Observing such high entry rates may be the result of this particular design with small competition prizes. However, incentives are the same in all treatments and therfore the between treatment analysis is not affected.

$p = 0.03$ for treatment *True* as comparison. The difference is however not significant when testing the frequencies of the information treatments against each other. This implies that the provision of information intensely decreases entry choices, and it is sufficient to provide aggregated data on the distribution of performance instead of precise information on the opponent. This is an interesting finding for application purposes, because aggregated data might be easier to collect than precise information on the opponent. In the following, we analyze whether the decrease of the entry frequency due to the provision of information, does also decrease competition failure and improve efficient entry choices.

Result 2. *The frequency of competition failure is significantly lower in treatment Distribution and True & Belief compared to the benchmark treatment without information. The main determinant of competition failure is overplacement instead of overestimation, willingness to take risks or a gender difference.*

Figure 2.2: Percentage of competition failure by treatments



The goal of this study is to investigate whether information is useful to reduce wrong entry decisions. To answer this question we focus on the criterion *competition* *failure*. This criterion is particularly useful, because it measures how often individuals earn less money than they could actually earn. The data on competition failure reveal a similar picture to the finding of entry frequencies and underline the important result that information does not only decrease entry choices, but more essentially, information prohibits wrong entry choices. We observe 35% of competition failures in the benchmark treatment, while only 15% of subjects fail in treatment *Distribution*, 22% in treatment *True* and 15 % in *True & Belief* (see Figure 2.2). The difference of failure rates in the benchmark compared to *Distribution* or *True & Belief* is significant in a Fisher-exact test with $p = 0.03$ and $p = 0.17$ for treatment *True* vs. *No Info*. Note that 35% is a very high number, because the highest possible failure rate would be that 50% fail.

The main driver of competition failure is overplacement, which is in the fashion of a result by Camerer and Lovallo (1999), where neglecting the opponents performance leads to market failures. Table 2.2 presents the determinants of competition failure with probit regressions. The dummy variable competition failure takes the value 1 if a subject entered the competition although the subject's performance is lower than that of the entering matched opponent. We analyze the influence of overestimation and overplacement on competition failure by using subjects reports to question (i) and (iv)[6] of the self-assessment stage and compare them with the true performance and performance ranking.[7] We also control for risk attitudes and gender. We find a significant influence of overplacement in treatment *No Info*. Subjects that believe they are better than others experience a competition failure significantly more often. This effect vanishes when providing information and it is only weakly significant in treatment *True & Belief*. We do not find that willingness to take risks or gender has explanatory power in determining competition failure. The general self-selection analysis will be discussed in the next section.

[6] "How many quiz questions have you solved correctly?" and "How many of the other participants in the room solved more quiz questions correct than you?"

[7] Using any of the other overconfidence definitions does not change the qualitative results.

Table 2.2: Determinants of competition failure

Dependent variable: Competition failure= $\begin{cases} 1 \quad \text{if} \quad \text{both enter and the performance is worse} \\ 0 \quad \text{otherwise} \end{cases}$

	No Info	*Distribution*	*True*	*True & Belief*
Overplacement (iv)	0.15***	0.10	0.08	0.18*
	(0.04)	(0.08)	(0.06)	(0.10)
Overestimation (i)	-0.09	0.10	0.05	0.12
	(0.09)	(0.08)	(0.10)	(0.19)
Willingness to take risks	0.02	-0.05	0.08	-0.31
	(0.10)	(0.14)	(0.15)	(0.27)
Dummy gender	0.24	0.32	-0.92*	1.13
	(0.46)	(0.53)	(0.51)	(0.84)
Constant	-1.11	-0.25	-2.70	4.57
	(2.3)	(3.26)	(3.55)	(6.09)
N	48	48	46	48
-LL	23.17	16.25	19.51	11.62

Notes: Coefficients of Probit estimates. Robust standard errors are in parentheses. Significance at the 1, 5, and 10 percent level is denoted by ***, **, and *, respectively. The variable *Willingness to take risks* indicates subjects' switching point from the lottery choice task of part four of the experiment. A high value indicates higher willingness to take risks. The variable *Dummy gender* takes the value 1 if male.

2.3.2 Self-Selection

Result 3. *The gender difference in competition is only significant in the benchmark and Distribution treatment.*

In Table 2.3 we present the determinants of entry decisions separately for each treatment and find similar results to previous findings of the literature on sorting behavior in a competitive environment. For the analysis, we classify a subject as overconfident with the following procedure. First, we generate one measure out of our five overconfidence measures with a principal-component-analysis. All of our five overconfidence measures are highly and positively correlated (see Table 2.4 in the Appendix 2.A). We extract a linear combination that explains most of the composition of all five measures. The first component score has an eigenvalue of 2.8 and therefore explains 56% of the composed analysis. Also, all five variables have a similarly large influence on the first component, such that the first component is an appropriate measure. Then, we define a subject as overconfident if the value of the first score is larger than zero and the subject is underconfident if it is smaller than zero.

In the benchmark treatment, all male subjects enter and therefore only female subjects can be included in the analysis of the *No Info* treatment. We find that performance has no significant influence on entry in the benchmark treatment. In all other treatments high performance types enter significantly more often. This is particularly interesting, because it complements our previous results that information is beneficial and is an additional criterion next to competition failure. Due to the provision of information, the competition game attracts significantly often high performance subjects. The average performance of subjects that enter in treatment *No Info* is 12.95 and is lower compared to treatment *Distribution*, where the average performance of entrants is 14.23. In addition, our data show that the well documented gender difference in competition entry behavior (see e.g. Gneezy, Niederle, and Rustichini (2003), Niederle and Vesterlund (2007), Dohmen and Falk (2011)) is also present in treatment *No Info* and *Distribution*, but vanishes in treatments *True* and *True & Belief*. This suggests that women are not per se less competitive and information, which reduces uncertainty about the opponent, decreases womens' reluctance to enter the competition. [8]

In treatment *True & Belief* we disclose the opponent's self-assessment bias to study whether it has an encouraging or daunting effect on competition entry. Our data reveal that subjects enter the competition less often when facing a very confident subject, however, the effect is not significant. In our experiment, subjects could not choose their self-assessment strategically. Charness, Rustichini, and van de Ven (2011) analyze the effects of a strategic self-assessment choice on entry behavior of the opponent in a tournament environment. They find that subjects are intimidated by a high self-assessment of the opponent and enter less often. Anyhow, they do not reveal the true performance of the opponent. Table 2.3 reports that the true performance of the opponent has a highly significant and sizable effect on entry choices in both treatments *True* and *True & Belief*. This information is seriously taken into account, such that the probability to enter is less likely if the opponent has a high performance.

[8] Similarly, Wozniak, Harbaugh, and Mayr (2010) analyze gender differences in compensation decisions with a within subject design of several treatments, where they disclose subject's own performance and the performance of all opponents in one treatment and find that the gender difference decreases.

Table 2.3: Determinants of stated self-assessment

Dependent variable: Competition entry= $\begin{cases} 1 & \text{if subject enters} \\ 0 & \text{otherwise} \end{cases}$

	No Info	*Distribution*	*True*	*True & Belief*
Overconfidence	0.33	0.46**	1.46*	0.21
	(0.30)	(0.21)	(0.79)	(0.21)
Performance	0.04	0.39**	1.52**	0.24*
	(0.26)	(0.18)	(0.67)	(0.15)
Opponent's performance			-1.18**	-0.45***
			(0.47)	(0.18)
Opponent's performance belief				-0.10
				(0.12)
Willingness to take risks	0.79**	0.30**	0.29	-0.01
	(0.38)	(0.13)	(0.27)	(0.13)
Dummy gender	dropped^{+}	1.75**	0.44	0.93
		(0.75)	(0.71)	(0.72)
Constant	-17.52*	-11.96***	-9.43	4.68
	(9.57)	(4.09)	(9.40)	(3.88)
N	24	48	46	48
-LL	6.4	16	7.7	18

Notes: Probit estimates. Robust standard errors are in parentheses. Significance at the 1, 5, and 10 percent level is denoted by ***, **, and *, respectively. The variable *Overconfidence* is the first component score of the principal component analysis of our five self-assessment measures. The variable *Willingness to take risks* indicates subjects' switching point from the lottery choice task of part four of the experiment. A high value indicates higher willingness to take risks. The variable *Dummy gender* takes the value 1 if male. $^{+}$Gender dummy predicts entry perfectly.

2.3.3 Existence and Robustness of Overconfidence

71% of our subjects think their quiz performance is better than the median performance. We observe the existence of overconfidence and underconfidence (see Figure 2.3 in the Appendix 2.A for histograms of the distribution of all five measures of overconfidence). However, mean and median of the overestimation measures are not significantly different from zero. Only the mean, but not the median of our overplacement measures (iv) and (v) is significantly larger than zero with $p = 0.02$ and $p = 0.06$ of a t-test. The mean and distribution of performance does not systematically vary across treatments ($p-values > 0.4$ of Wilcoxon-ranksum tests and t-tests). The average quiz performance is 13.5, the worst performance was 6 correct answers and the best one was 19 correct answers. This indicates a mediocre task difficulty level. Our finding coincides with the results by Healy and Moore (2008a) who find, that easy tasks produce underestimation, difficult tasks produce overestimation, and mediocre tasks produce on average no estimation bias. We find that all of our five elicited measures of overconfidence are significantly and positively correlated. Table 2.4 in the Appendix 2.A reports a correlation table of all confidence measures and performance. There exists a significantly negative correlation of performance and overconfidence for the relative and absolute measures, which is robust when controlling for ceiling effects. Due to the definition of overplacement and overestimation, the best performing subjects can never be overconfident and the worst ones can never be underconfident. The negative correlation of overestimation and overplacement with performance still persists and is significant ($p - values < 0.05$ for Spearman-rank tests) when restricting the sample to productivities lower than 20, 19, 18, 17 or 16.[9] Overconfidence in our setting is linked to an especially optimistic belief about one's performance type, while psychologists describe optimism as a positive view towards uncertain future events or concentrating on the good sides of life. We employ a 10 item questionnaire on optimism at the end of all main tasks of the experiment to analyze a connection to the economic definition of overconfidence. Indeed, we find that the psychological measure of optimism is positively correlated with overplacement

[9]The low performance subjects can always underestimate their performance. However the worst subjects can not underplace themselves. When omitting the best and worst subjects, the negative correlation of overplacement and performance is still significant.

($p = 0.08$ of a Spearman-rank test).

2.4 Conclusion

Well adjusted absolute and relative self-assessments about skill, ability or achievements are important to decide about competition entry. However, individuals frequently misestimate their own skill, ability or achievements. In this paper, we study a competition game where subjects might have an absolute or relative self-assessment bias about their performance. We ask whether subjects benefit from information about the opponents by updating their performance beliefs which in consequence should reduce competition failure. With a laboratory experiment, we can exogenously vary information that we provide to subjects before they decide about entering the competition or taking an outside option. Subjects enter the competition significantly less often in our information treatments compared to the benchmark treatment. The two main contributions of this study are firstly, simple aggregated information on the performance distribution significantly reduces competition failure by up to 57%, and secondly, aggregated information is sufficient and more detailed information does not further reduce inefficiency. The information provision does not directly disperse the self-assessment bias, but it raises the awareness for the opponent's potentially high performance and mitigates the consequences. Our findings suggest a simple and inexpensive way to reduce decision making that yields less monetary utility by increasing individuals' appreciation of the decision problem.

An interesting extension in the field would be an implementation at an employment agency or job center. The agency could provide aggregated information for startup businesses like Ich-AGs in Germany (You-Inc.).[10] For example, an entrepreneur who wants to start a restaurant could receive information on the current amount of restaurants and restaurant failures. We expect less business openings and less frequent business failures. In a similar vein, firms could provide a distribution of the qualification of their workers before the hiring process to reduce the amount of applications and organizational costs. Inexpensive information which is simple to

[10] Caliendo et al. (2007) show that 20-40% of You-Inc. start ups ("Ich-AG") do not exists anymore after 16 month and particular groups like facility managers fail especially often, because the demand is already exhausted.

understand might also be valuable in other economic contexts. Providing information about competitors by disclosing past performances could reduce the baneful effects of overconfidence in economic environments like labor negotiations, litigation, or stock trading.

2.5 Appendix

Appendix 2.A

Figure 2.3: Histograms of overconfidence measures (i) to (v)

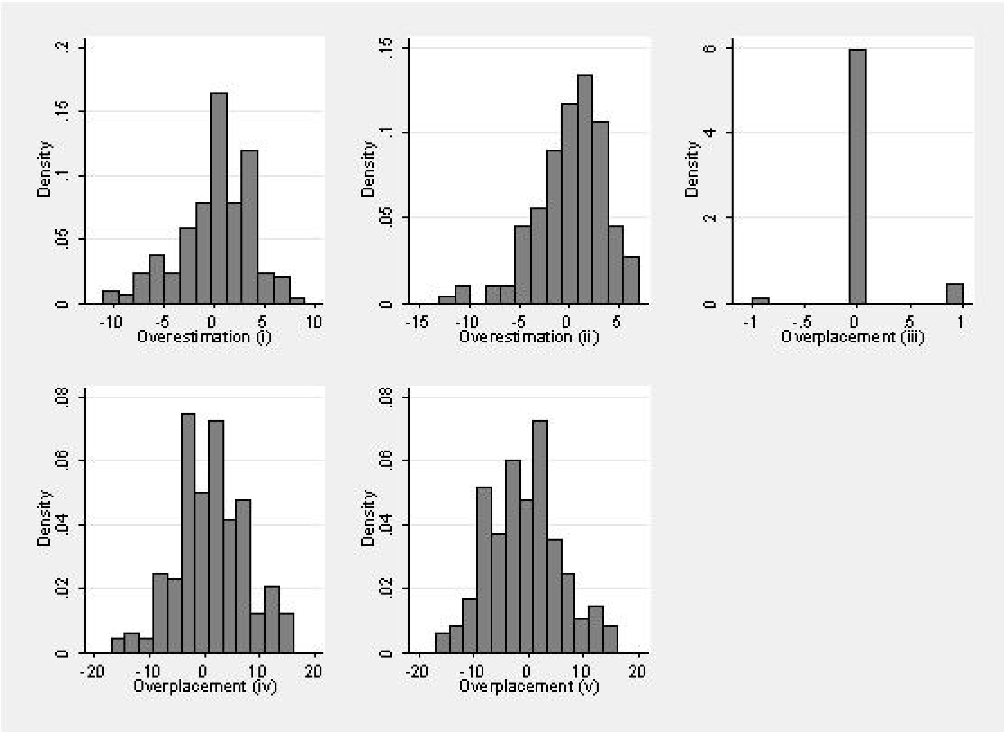


Table 2.4: Correlation table of our five measures of overconfidence and quiz performance

	Overestimation		Overplacement		
	(i)	(ii)	(iii)	(iv)	(v)
(i)	1				
(ii)	0.77***	1			
(iii)	0.26***	0.22***	1		
(iv)	0.41***	0.44***	0.46***	1	
(v)	0.36***	0.38***	0.44***	0.90***	1
Quiz performance	-0.14*	-0.19***	-0.25***	-0.68***	-0.68***

Notes: N=190. Significance of the Spearman-rank test at the 1, 5, and 10 percent level is denoted by ***, **, and *, respectively.

Table 2.5: Measures of absolute and relative overconfidence

Measure	Level of overconfidence
(i) Degenerate	= self-assessment - quiz performance (QP)
(ii) Token distribution task	$=\sum_{i=0}^{20} t_i \cdot i/20 - QP$, where t_i is the amount of points for category i
(iii) Median	$= \begin{cases} 1 & \text{if worse than median although indicated better} \\ 0 & \text{if median comparison is correct} \\ -1 & \text{if better than median although indicated worse} \end{cases}$
(iv) Upward distribution	= amount of subjects that are better - self-assessment
(iv) Downward distribution	= self-assessment - amount of subjects that are worse

Appendix 2.B

Instructions, translated into English. General instructions and instructions for first, second, and fourth part of the experiment were identical across treatments. Instructions for the third part of the experiment differed across treatments.

GENERAL INSTRUCTIONS

You are taking part in a decision-making experiment in which you have the opportunity to earn money. The amount of money you earn is paid to you upon completion of the experiment. Please read the instructions carefully. The instructions are identical for all participants. If you have any questions, please raise your hand. The experimenter will answer your question at your place. During the experiment, you have to remain silent. Violation of this rule leads to immediate exclusion from the experiment and all payments.

All monetary units in the experiment are measured in points, and 100 points = 1 Euro.

This experiment consists of four parts. In all parts, you can earn money. Your payoff from the experiment results from the sum of your payoffs in all parts. In the following we will go through the instructions for the first part of the experiment. After the first part is completed, we will provide you with the instructions of the second part.

INSTRUCTIONS FOR THE FIRST PART OF THE EXPERIMENT

In the first part of the experiment you will be asked 20 quiz questions. You will always be offered 4 possible answers of which exactly one will be correct. Please always select one of the four possible answers. You get 20 points for each correct answer. After you have answered the first 10 questions, please click on the OK button. Then a new screen with 10 more questions will appear. Please confirm your responses again with the OK button.

Do you have any questions?

INSTRUCTIONS FOR THE SECOND PART OF THE EXPERIMENT (Only on screen)

We will ask you five questions concerning your estimation of your quiz performance. The questions are on the screen and will not be read together. The questions are identical for all participants. For each question, you can again earn points. The points will be added to the points earned in the quiz. Please give your answers on the screen and confirm them with the OK button. If you have any questions, please raise your hand, we come to your place.

Question 1: Please estimate now as good as possible the number of your correct quiz answers.

Your payment will look like this: You earn 200 points if your estimate is correct and 50 points, if your estimate is a number next to the correct number.

Question 2: Please estimate now again as good as possible the number of your correct quiz answers.

This time you do not have to commit to a number. Instead, you get 100 points which you can allocate into any of 21 fields. Each field of 0-20 is the number of correct answers. Your payment is as follows: The number of points that you have allocated to the field that corresponds to your correct number of quiz answers will be doubled and paid. The points, on a field next to the correct field, will be paid to 50%. For example, participant xy has 10 correct answers and distributes the coins as follows:

Correct quiz questions	...	9	10	11	12	13
Token	0	10	50	40	0	0

then the participant gets: $0.5 \cdot 10 + 2 \cdot 50 + 0.5 \cdot 40 = 125$ points.

Please enter a "0" in any field, where you do not want to allocate points to.

Question 3: Is your amount of correct quiz questions one of the best 12 or worst

12 quiz performances in the room?

For a correct estimate you will get 100 points.

Question 4: What do you think, how many of the other 23 participants answered more quiz questions correct than you?

You earn 200 points if you have guessed correctly and 50 points, if your estimated number is one number next to the correct number.

Question 5: What do you think, how many of the other 23 participants have less correct quiz answers than you?

You earn 200 points if you have guessed correctly and 50 points, if your estimated number is one number next to the correct number.

If you have any questions, please raise your hand.

INSTRUCTIONS FOR THE THIRD PART OF THE EXPERIMENT

In part 3 you will be randomly assigned to another participant of the experiment. You will learn at no time of the experiment the identity of the assigned participant. You decide in part 3, if you want to participate in a competition against your randomly matched participant or not.

If you do not want to participate, you will get 200 points, whatever decision your opponent takes.

If you want to participate, you will produce an individual output. The output corresponds to your number of correct answers in the quiz. You compete against your randomly assigned opponent with your output.

$$\text{Output} = \text{number of correct quiz answers}$$

Example: Player xy has solved 1 of 20 questions in the quiz correctly. If player xy participates, he or she produces an output of 1.

If you participate, your payoff depends on your own output and the output of your opponent in case he or she participates, too. The rules of the competition are as follows:

1. If you participate but your opponent does not, you win automatically.
2. If you and your opponent participate, you win if you have a higher output than your opponent. But if you have a lower output than your opponent, you lose.
3. If you and your opponent participate and you both have the same output, the computer randomly decides who wins and who loses.

The winner receives 400 points and the loser gets 100 points. You will find the payoff matrix below on all of your decision screens.

	Participation	No participation
Winner prize	400	200
Loser prize	100	200

DECISION

The only thing you need to do is to decide for or against participation in the competition. You make your decision on the screen and confirm it with the button "Confirm decision".

CONTROL QUESTIONS

Before part 3 starts, we ask you a few control questions to ensure that all participants understand part 3 of the experiment. You earn no money for correct answers of the control questions, nor do we take money away from you for a wrong answer. However, part 3 starts only when all participants have answered all control questions correctly.

Example: Player 1 solved 10 questions correct and player 2 solved 8 questions correct.

Question 1: What is the output of player 1 if he enters the competition?

Question 2: How many points does player 1 receive if he does not enter the competition?

Question 3: What is the output of player 2 if he enters the competition?

Question 4: How many points does player 3 receive if he does not enter the competition?

Question 5: Who wins the competition if both enter?

Question 6: How many points does player 1 get if both enter the competition?

Question 7: How many points does player 2 get if both enter the competition?

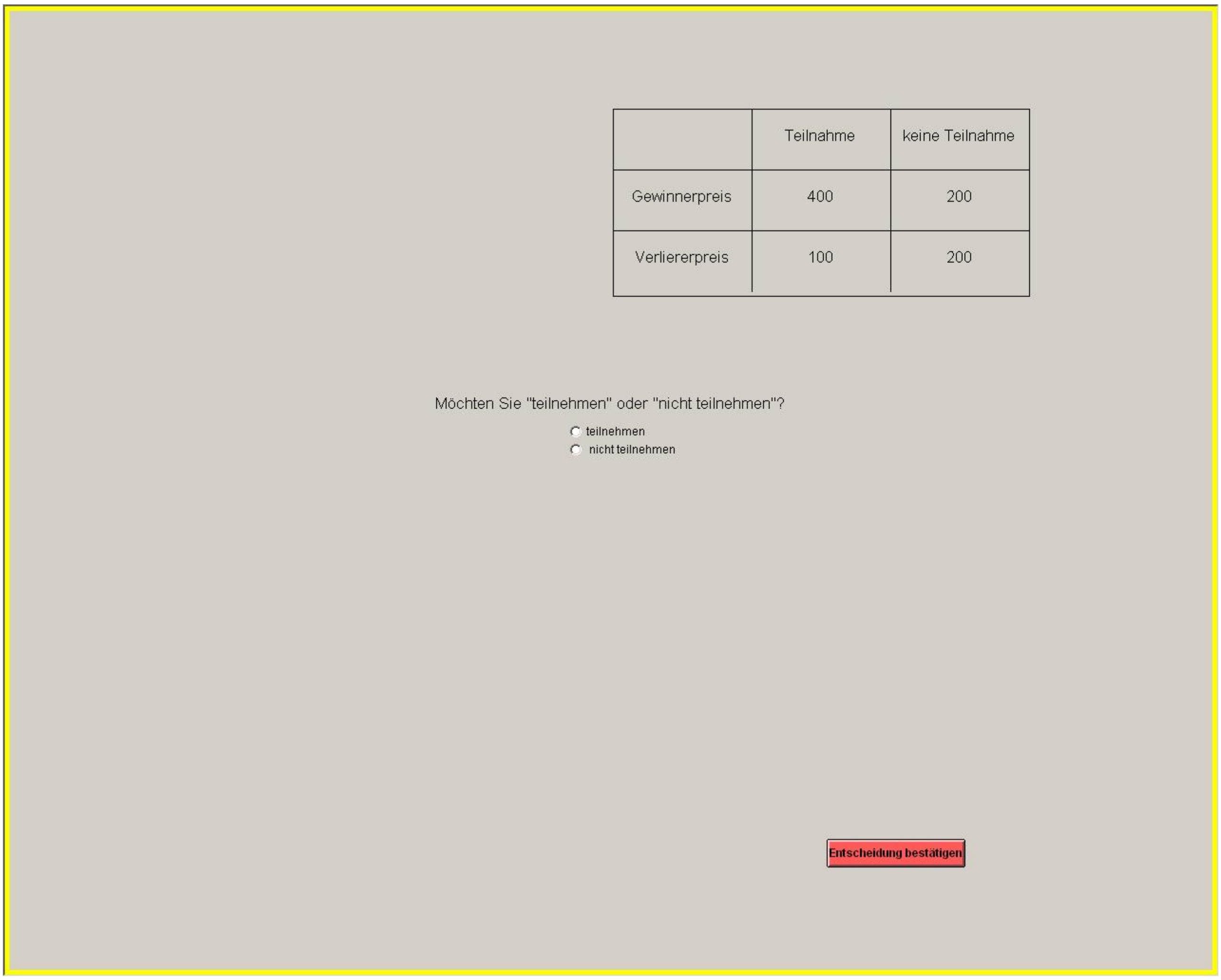
Question 8: How many points does player 1 get if only he enters the competition?

Question 9: How many points does player 2 get if only he enters the competition?

Any questions?

We present now the additional instructions four each treatment. Note that each subject participated only in one treatment.

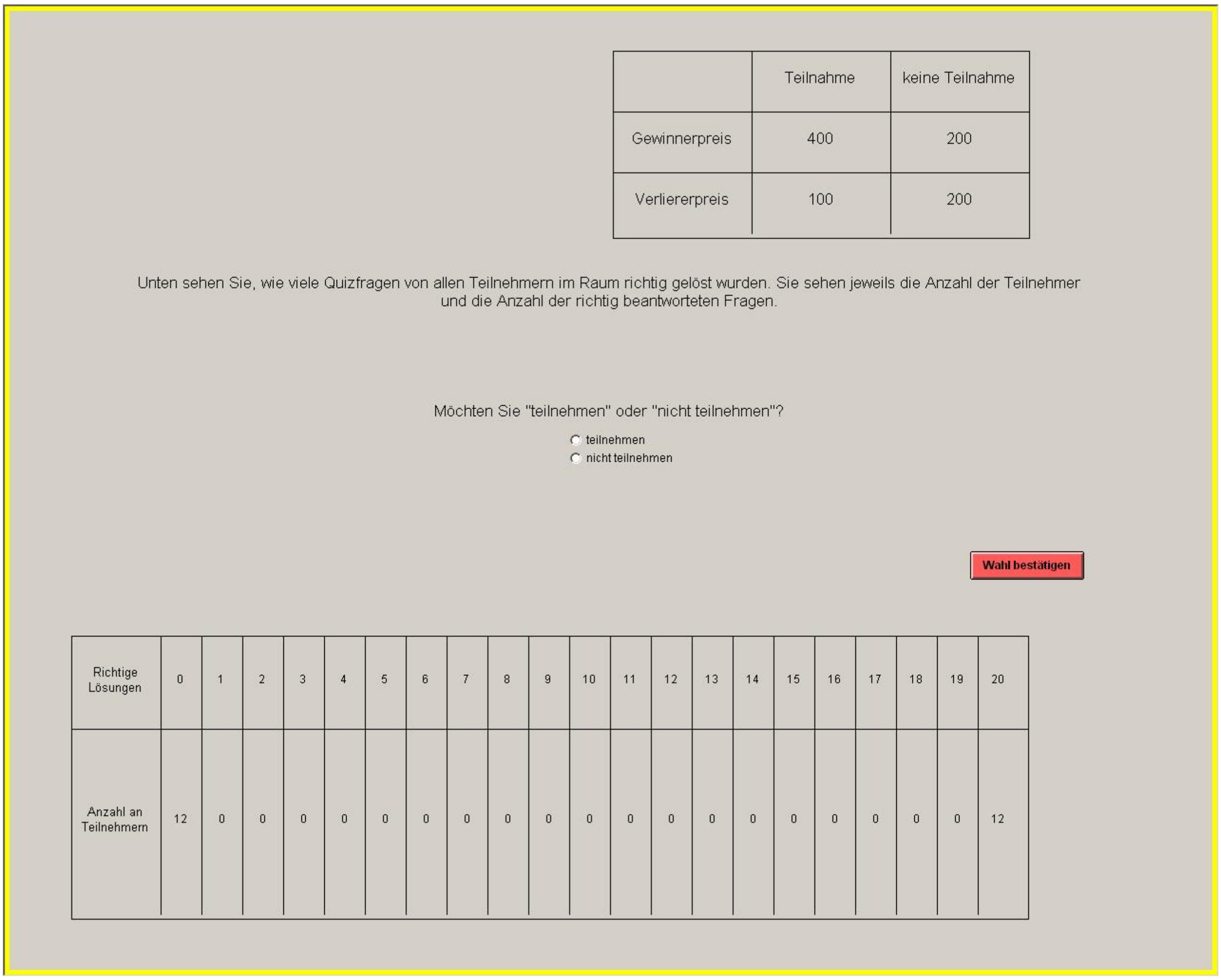
Treatment No Info You will be randomly assigned to one participant. You both make your decision to participate in the competition. You receive no information about the participant that is randomly matched to you and likewise your randomly matched participant does not receive any information about you before you both make your decision about participation in the competition. Your decision screen looks like this:



Please confirm your decisions by pressing the button "Confirm decision ". Any questions?

Treatment Distribution You will be randomly assigned to one participant. You both make your decision to participate in the competition. Before that, you get summarized information on all participants in the room, including you. All participants receive this information on the screen. On the screen, you will see a table with the number of correctly solved quiz questions from 0-20 and the number of participants that solved the corresponding number of quiz questions correctly.

Example: 12 participants answered all questions wrong, and the remaining 12 participants answered all questions correct. Your screen would look like this:



	Teilnahme	keine Teilnahme
Gewinnerpreis	400	200
Verliererpreis	100	200

Unten sehen Sie, wie viele Quizfragen von allen Teilnehmern im Raum richtig gelöst wurden. Sie sehen jeweils die Anzahl der Teilnehmer und die Anzahl der richtig beantworteten Fragen.

Möchten Sie "teilnehmen" oder "nicht teilnehmen"?

- teilnehmen
- nicht teilnehmen

Wahl bestätigen

Richtige Lösungen	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Anzahl an Teilnehmern	12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	12

Please confirm your decisions by pressing the button "Confirm decision". Any questions?

Treatment True You will be randomly assigned to one participant. You both make your decision to participate in the competition. Before that, you receive the information how many quiz questions your randomly matched participant answered correctly. This information is shown on your decision screen. Likewise, your randomly matched participant will be informed about your number of correctly solved quiz questions. Then both of you take the decision to participate. Your decision screen looks like this with the respective number of correct quiz answers of your matched participant :

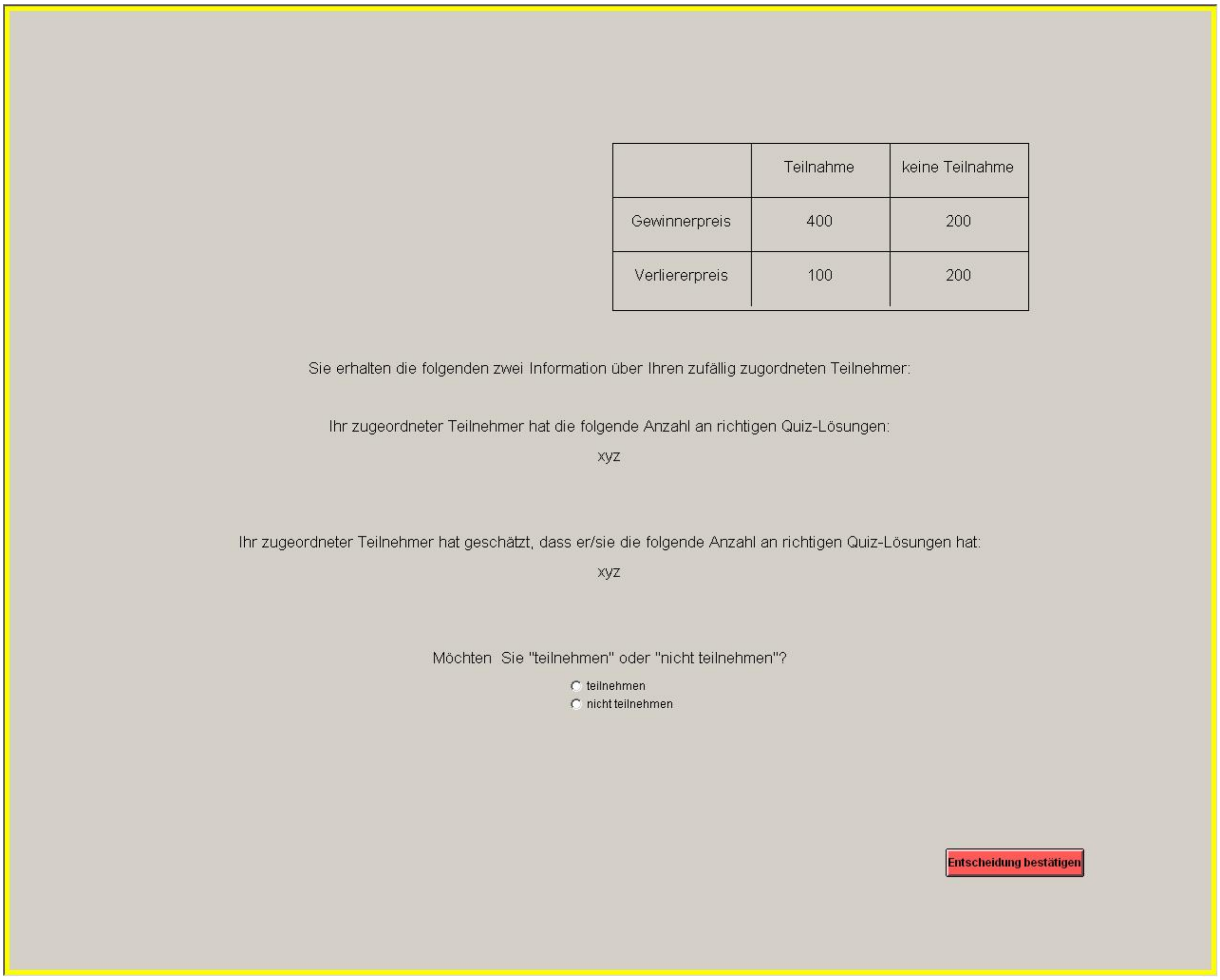
Please confirm your decisions by pressing the button "Confirm decision ". Any questions?

Treatment True & Belief You will be randomly assigned to one participant. You both make your decision to participate in the competition. Before that, you receive two pieces of information about your randomly matched participant.

1. how many questions he/she has correctly answered in the quiz

2. how many questions he/she estimated that he/she answered correctly.

This information is shown on your decision screen. Likewise, your randomly matched participant will be informed about your number and estimation of your correctly solved quiz questions. Then both of you take the decision to participate. Your decision screen looks like this with the respective number and estimation of correct quiz answers of your matched participant :



Please confirm your decisions by pressing the button ”Confirm decision”. Any questions?

INSTRUCTIONS FOR THE FOURTH PART OF THE EXPERIMENT

In the fourth part of the experiment, you make 30 decisions. You have to choose 30 times between Option A and Option B. Under Option A, you will receive a secure payment that starts with 0 points in the first decision round and increases to 400 points in the last decision round. Under Option B you will receive a lottery. In the lottery, you get with 50% probability either 100 points or 400 points. The lottery is in all 30 decisions the same.

We pay one of the 30 decisions, which is chosen randomly. You confirm your 30 decisions with the OK button. Do you have any questions?

Chapter 3

Auctions with Loss Averse Bidders

3.1 Introduction

Since Kahneman and Tversky (1979), loss aversion and reference dependent preferences have been applied to a variety of empirical and theoretical economic problems. When applying models of loss aversion, the modeller is required to decide over what individuals have feelings of gains and losses. This is the problem of narrow versus wide bracketing. To illustrate the problem, consider the series of experiments conducted by Kahneman, Knetsch, and Thaler (1990), who study the endowment effect in competitive markets. When subjects are given actual goods, the endowment effect has an impact on trading volumes; if, however, subjects are endowed with money rather than a good, they observe no endowment effect. The explanation given is that when trading money for coffee mugs, there is a friction caused by a loss in one and a gain in the other dimension. When money is traded for money, this friction disappears. Köszegi and Rabin (2006) propose a model which rationalizes the experimental findings mentioned, using the concept of consumption dimensions, over which individuals have gain loss utility in an additively separable manner. Applying the model of Köszegi and Rabin (2006, 2007), we derive the equilibrium bidding behavior in the first price auction (FPA) and in the all pay auction (APA) for general environments with independent private values (IPV). In addition, we study the behavioral implications of loss aversion on bidding strategies, and compare the revenue across auction formats. In one specification, we consider gains and losses in two dimensions separately, about whether they receive the object or not, and how much they pay. In the other specification, we consider gains and losses over the entire risk neutral payoff, i.e. the valuation less the bid. The first specification represents narrow bracketing, while the second one represents wide bracketing. With one dimension, we show that the expected revenue for the auctioneer is higher in the FPA than in the APA, and with two dimensions, we show that the opposite is true for the revenue ranking between the FPA and the all pay auction.

In order to test the theoretical predictions, we conduct laboratory experiments, in which either money or a real object is auctioned in both a FPA and an APA. We find that in both settings, the average revenue is significantly higher in the first price auction, suggesting that bidders may behave according to the one dimensional model, although a real object is auctioned. Whereas our findings are inconsistent with the two dimensional model, they are consistent with the one dimensional model.

The paper contributes to the literature on loss aversion and reference dependent preferences in several ways. Comparing our results to the ones in Kahneman, Knetsch, and Thaler (1990), we conclude that whether individuals do a narrow or a wide form bracketing of gains and losses depends on the environment under consideration. While competitive markets and auctions are similar in many ways, the degree of uncertainty is a lot higher in auctions. Additionally, we provide an estimate for the ratio of marginal disutility of losses to marginal utility of gains of 1.42, using the generalized method of moments for the data obtained in the induced value experiments. Furthermore, we show that when applying the Köszegi and Rabin (2007) model, the theoretical predictions depend crucially on the modeller's decision how to define the consumption dimensions over which individuals experience gains and losses.

Finally, our experimental data shows that there is no measurable difference between auctioning an actual good or simply money in auctions with induced valuation. This result is important from a methodological view for experimenters that choose to conduct auction experiments in the laboratory or in the field.

3.1.1 Related Literature

Auction Theory and Risk Preferences

Riley and Samuelson (1981), Maskin and Riley (1984), Matthews (1987), and Fibich, Gavious, and Sela (2006) study the implications of risk averse bidders in auction settings. Lange and Ratan (2010) consider the case of loss averse bidders for the FPA and the Vickrey auction and show that the FPA yields higher expected revenue than the Vickrey auction, independent of whether bidders consider gambles in one or two dimensions. Shunda (2010) shows that under a different notion of reference dependence, the auctioneer can increase his expected revenue by introducing a buy now price. In the present paper, we focus on a specific class of hybrid auctions, incorporating both the FPA and the APA, and study the bidders' behavior and revenue (non) equivalence across different auction formats. Furthermore, while the revenue ranking of the FPA and the Vickrey auction in both models is the same (Lange and Ratan (2010)), our analysis provides another testable implication of reference dependence with revenue data alone. Using a general mechanism design approach in the spirit of Myerson (1981), Eisenhuth (2012) shows that in the one dimensional model, the FPA maximizes the expected revenue among all efficient auctions, and that in the two dimensional model, any optimal auction is fully all pay. In light of these results, we focus on the optimal auction (in the class of efficient auctions) in each case when considering the FPA and the APA.

Experimental Economics

The empirical literature on the APA is small, as it is not a commonly used auction format. To the best of our knowledge, Noussair and Silver (2006) provide the only empirical analysis comparing the APA and the FPA in a laboratory setting with independent private values. They replicate the environment in Cox, Smith, and Walker (1982) and Cox, Roberson, and Smith (1988), who study the FPA, and compare the revenue data from these studies to their revenue data on the APA. Their finding is that the APA yields significantly higher revenue than the FPA. One confounding effect is that they provide subjects with an initial endowment of nearly seven times as much as Cox, Smith, and Walker (1982) and Cox, Roberson, and Smith (1988).

Thereby, Noussair and Silver (2006) lose some control over their data comparison Furthermore, they observe bids of 0 for the lowest types in either auction format. Real object auctions are not studied. Lucking-Reiley (1999) studies real object field auctions using the FPA, the Vickrey auction, the English auction, and the Dutch auction with Magic cards and refutes revenue equivalence; an analysis of the APA is missing. Moreover, as the data are collected through online auctions, bidders do not know how many opponents they are facing in the auction. We contribute to the experimental literature by studying revenue equivalence between the APA and the FPA, explicitly differentiating between auctioning money and an actual object.

3.2 Model

3.2.1 Preferences

We consider two different specifications of reference dependent preferences. The first one is a specification according to which the bidders consider gambles for the object and money separately; the second specification treats the difference between the valuation for the object and the amount paid as one dimension, and gambles are evaluated over this difference only. As proposed in Köszegi and Rabin (2006), the first specification of bidders' preferences is given by

$$u(c^g, c^m | r^g, r^m, \theta) := \underbrace{\theta c^g + c^m}_{\text{intrinsic utility}} + \underbrace{\eta^g \mu^g(\theta(c^g - r^g)) + \eta^m \mu^m(c^m - r^m)}_{\text{gain loss utility}},$$

where $c^g, r^g \in \{0, 1\}$ captures the good dimension, $c^m, r^m \in \mathbb{R}$ captures the money dimension. For $l \in \{g, m\}$, c^l is true consumption, r^l is the reference level of consumption, $\eta^l > 0$, measures the weight attached to gain loss utility in dimension l, and $\theta \geq 0$ is the bidder's intrinsic valuation for the good. The second specification is given by

$$u(c^g, c^m | r^g, r^m, \theta) := \underbrace{\theta c^g + c^m}_{\text{intrinsic utility}} + \underbrace{\eta \mu(\theta c^g + c^m - (\theta r^g + r^m))}_{\text{gain loss utility}}.$$

Moreover,

$$\mu^l(x) := \begin{cases} x, & \text{if } x \geq 0 \\ \lambda^l x, & \text{if } x < 0, \end{cases}$$

where $\lambda^l > 1$, $l \in \{g, m\}$, and the second specification with only one dimension is implied when the index on the parameters is suppressed. These preferences capture loss aversion through the Kahneman and Tversky (1979) value function, μ^l, $l \in \{g, m\}$. A deviation from the reference point is disliked more if it is a loss than it is liked if it is a gain.

3.2.2 Auction Rules

A single, indivisible object is sold among $N \geq 2$ loss averse bidders who share the same η^l and λ^l, $l \in \{g, m\}$, and whose valuations, $\{\theta_i\}_{i=1}^N$, are the realizations of N independent draws from the continuous distribution function, $F : \Theta \to [0, 1]$, where $\Theta := [\theta_{\min}, \theta_{\max}] \subset \mathbb{R}_+$, with strictly positive density everywhere. The valuation of bidder i, θ_i, is bidder i's private information. The bidders and the auctioneer share the same prior beliefs. We consider the following class of auctions with all pay component, $\alpha \in [0, 1]$. Bidders simultaneously submit their bid, and the bidder with the highest bid wins the object and pays his entire bid. All other bidders walk away without the object but have to pay α of their bid. In case of a winning tie, the winner is selected among the highest bidders with equal probability. For $\alpha = 0$, we have the FPA and for $\alpha = 1$, the APA. This formulation, incorporating both the APA and the FPA, appears first in Siegel (2010). Other common auction formats, as, for instance, the Vickrey auction are excluded from our analysis, partially because Lange and Ratan (2010) study the Vickrey auction in the same setting and partially because Eisenhuth (2012) shows that this is without loss of generality, using a general mechanism design approach in the spirit of Myerson (1981). In our theoretical analysis, we focus on symmetric equilibrium bidding functions.

3.2.3 Solution Concept

In the above described auction setting, each bidder learns his valuation before submitting his bid and therefore, maximizes his interim expected utility. Using Kőszegi and Rabin (2006)'s notation, if the distribution of reference points is G, and the distribution of actual consumption outcomes is H, the decision maker's interim expected utility is given by

$$U(H|G,\theta) := \int_{\{(c^g,c^m)\}} \int_{\{(r^g,r^m)\}} u(c^g, c^m|r^g, r^m, \theta) dG(r^g, r^m|\theta) dH(c^g, c^m|\theta).$$

Definition 1. *(Köszegi and Rabin (2007)) Conditional on the realization of the type, θ, for any choice set, D, $H \in D$ is an interim CPE if $U(H|H,\theta) \geq U(H'|H',\theta)$, for all $H' \in D$.*

Fixing all other bidders' behavior, each bidder's bid, b_i, induces a distribution, $H_i(\mathcal{A}|b_i, b_{-i})$, over the set of alternatives, $\mathcal{A} := \{0,1\}^N \times \mathbb{R}^N$. Therefore, the definition can be modified in the following way to match the auction setting under consideration.

Definition 2. *Conditional on the realization of the type, θ_i, $b : \Theta \to \mathbb{R}_+$ is a symmetric interim CPE bidding function if for all $i, \theta_i, \theta_{-i}, b' \geq 0$,*

$$\begin{aligned} &U\left(H_i(\mathcal{A}|b(\theta_i), b_{-i} = b(\theta_{-i}))|H_i(\mathcal{A}|b(\theta_i), b_{-i} = b(\theta_{-i})), \theta_i\right) \\ &\quad \geq U\left(H_i(\mathcal{A}|b', b_{-i} = b(\theta_{-i}))|H_i(\mathcal{A}|b', b_{-i} = b(\theta_{-i})), \theta_i\right). \end{aligned}$$

The interpretation of CPE is that each bidder understands that once consumption occurs, i.e. once the auction is over, he evaluates this consumption outcome against the actual lottery as his reference lottery. As laid out in Köszegi and Rabin (2007), CPE is the most appropriate solution concept for decisions under risk, whose uncertainty is resolved long after the decision is made. An alternative solution concept, *choice unacclimating personal equilibrium* (UPE), requires the decision to be optimal, given the expectations at the time the decision is made. Below we show that UPE is equivalent to CPE in the auction setting under consideration. This equivalence, however, cannot be generalized to other setting, as documented by Köszegi and Rabin (2007). We will state all formal results using CPE, but show that replacing CPE by UPE, the validity of all results is retained. For the following analysis, it is convenient to define $\Lambda^l := \eta^l(\lambda^l - 1) > 0$, $l \in \{g, m\}$, which can be viewed as an overall measure of the degree of loss aversion in the respective dimension. The following assumption, as proven in the Appendix 3.A, guarantees that all bidders participate in the auction for any realization of their own type, and that their equilibrium bidding functions derived below are strictly increasing.

Assumption 1. ***(No Dominance of Gain Loss Utility)*** $\Lambda^g \leq 1$.

This assumption places, for a given η (λ), an upper bound on λ (η). In Herweg, Müller, and Weinschenk (2010), this assumption is referred to as *no dominance of gain loss utility*. In the following, we consider each specification of reference dependent preferences, one and two dimensional, at a time.

3.3 Analysis

3.3.1 Two Dimensions

Consider the ex post utility of bidder i when his bid is x, and x_{-i} is the vector of all other bidders' bids. Let $q_i(x) = P(i \text{ wins } | x, x_{-i}) = P(x > \max_{j \neq i}\{x_j\})$ be the probability that bidder i wins the auction, conditional on his own and all other bidders' bids. When he ends up with the object and pays x, his utility is

$$\underbrace{\theta_i - x}_{\text{intrinsic utility}} + \underbrace{\eta^g \left(1 - q_i(x)\right) \theta_i - \eta^m \lambda^m \left(1 - q_i(x)\right)(1 - \alpha)x}_{\text{gain loss utility}}.$$

The first term represents intrinsic utility, and the second term captures gain loss utility. Compared to the situation in which the bidder does not win the auction, which happens with probability $(1 - q_i(x))$, he experiences a gain in the good dimension and a loss in the money dimension. In case bidder i ends up without the object and his bid is x, his utility is

$$\underbrace{-\alpha x}_{\text{intrinsic utility}} + \underbrace{\eta^g \lambda^g q_i(x)(-\theta_i) + \eta^m q_i(x)(1 - \alpha)x}_{\text{gain loss utility}},$$

since, compared to the situation in which he wins the auction, which happens with probability, $q_i(x)$, this is considered a loss in the good dimension and a gain in the money dimension. Therefore, bidder i's interim expected utility is

$$\begin{aligned} &q_i(x)\left(\theta_i - x + \eta^g\left(1-q_i(x)\right)\theta_i - \eta^m\lambda^m\left(1-q_i(x)\right)(1-\alpha)x\right) \\ &\quad + \left(1-q_i(x)\right)\left(-\alpha x - \eta^g\lambda^g q_i(x)\theta_i + \eta^m q_i(x)(1-\alpha)x\right). \end{aligned}$$

By varying their bid, x, each bidder changes the probability of winning, and therefore his reference lottery. We look for strictly increasing, symmetric equilibrium bidding functions. Hence, dropping the i subscript, the bidder's program is

$$\begin{aligned} V(\theta) &:= \max_{x\in\mathbb{R}_+}\Big\{q(x)\left(\theta - x + \eta^g\left(1-q(x)\right)\theta - \eta^m\lambda^m\left(1-q(x)\right)(1-\alpha)x\right) \\ &\qquad + \left(1-q(x)\right)\left(-\alpha x - \eta^g\lambda^g q(x)\theta + \eta^m q(x)(1-\alpha)x\right)\Big\} \\ &= F^{N-1}(\theta)\left(1-\Lambda^g\left(1-F^{N-1}(\theta)\right)\right)\theta \\ &\quad -F^{N-1}(\theta)\left(1+\Lambda^m\left(1-F^{N-1}(\theta)\right)\right)(1-\alpha)b_\alpha(\theta) - \alpha b_\alpha(\theta), \end{aligned} \tag{3.1}$$

where the ultimate equality follows from independence of the types, b_α being strictly increasing (and hence, invertible), and the definition of Λ^l, $l \in \{g, m\}$. Since, $V(\theta_{\min}) = 0$, applying the envelope theorem yields the following expression for the symmetric CPE bidding function:

$$b_\alpha(\theta) = \frac{\beta(\theta)\theta - \int_{\theta_{\min}}^{\theta}\beta(s)ds}{(1-\alpha)F^{N-1}(\theta)\Delta(\theta)+\alpha},$$

where $\Delta(\theta) := (1+\Lambda^m(1-F^{N-1}(\theta))) \geq 1$ and $\beta(\theta) := F^{N-1}(\theta)\left(1-\Lambda^g\left(1-F^{N-1}(\theta)\right)\right)$. Furthermore, by the envelope theorem, the bidder's payoff is independent of the auction format, α, and depends, as in the risk neutral case, only on the probability of winning the auction.

Proposition 1. *Suppose assumption 1 holds. Then, $b_\alpha(\theta)$ is strictly increasing, for almost all θ and constitutes the unique symmetric pure strategy CPE bidding function.*

Proposition 1 and all formal results which follow are proven in the Appendix 3.A. In order to study the equilibrium bidding behavior of loss averse bidders, it is instructive to first consider the case in which bidders are only loss averse in the money dimension ($\Lambda^m > 0$) and risk neutral in the good dimension ($\Lambda^g = 0$). Letting b_α^{RN} denote the equilibrium bidding function with risk neutral bidders in the same environment, the CPE bidding function then reads

$$b_\alpha(\theta) \quad = \quad \frac{1}{\psi_\alpha(\theta)} b_\alpha^{RN}(\theta),$$

where

$$\psi_\alpha(\theta) := \frac{(1-\alpha)F^{N-1}(\theta)\Delta(\theta) + \alpha}{(1-\alpha)F^{N-1}(\theta) + \alpha} \geq 1.$$

If bidders only have gain loss considerations in the money dimension, then the equilibrium bid is the distorted risk neutral bid. Regarding the comparative statics results with respect to the parameter, Λ^m, the following result holds.

Proposition 2. *$b_\alpha(\theta)$ is strictly decreasing in Λ^m, for almost all θ, for all $\alpha \in [0, 1)$. For $\alpha = 1$, Λ^m has no effect on $b_\alpha(\theta)$, for all θ.*

In order to see the behind of this result, consider the FPA ($\alpha = 0$). As above, let $q(x)$ be the probability of winning the object when submitting a bid of x. In case he wins, a bidder pays x which is a loss of x relative to paying nothing (in case he loses). This loss sensation is weighted with the probability of not having to pay, which is $(1 - q(x))$ since the reference point is formed at the interim stage. Since a bidder wins with probability $q(x)$, from an interim perspective, this feeling of loss occurs with probability $q(x)$, and thus there is an interim expected loss of $\eta^m \lambda^m q(x)(1-q(x))x$. Likewise, in the event of losing, a bidder considers the bid saved a gain of x in the money dimension, and the outcome of winning is weighted with $q(x)$ in the reference lottery. Furthermore, from an interim perspective, a bidder expects to lose with probability $(1 - q(x))$ so his interim expected gain is $\eta^m q(x)(1 - q(x))x$. Consequently, the overall expected gain loss sensation is $\eta^m(\lambda^m - 1)q(x)(1 - q(x))x$, and since losses loom larger than gains ($\lambda^m > 1$), this is always negative when there are multiple monetary outcomes. Hence, a bidder's interim benefit of winning is less and hence, in equilibrium, this lowers the bidders' bids relative to the APA, in which payments are certain from an interim perspective. More specifically, the reduction in interim expected payoff consists of three parts, the overall degree of loss aversion, $\eta^m(\lambda^m - 1) = \Lambda^m$, the variance of the Bernoulli distributed outcome of winning or losing the auction, $q(x)(1 - q(x))$, and the wedge between the amount paid when winning and losing. Since the last part is identical to 0 in the APA, gain loss considerations have no effect on the interim expected payoff, whereas for all $\alpha \in [0, 1)$, the payoff is reduced. In a symmetric equilibrium, the interim probability of winning and receiving the good is the same across auction formats, in particular, $q(x) = F^{N-1}(\theta)$, and therefore, does not affect the comparison across auction formats. In order to examine how loss aversion in the good dimension affects the bidding behavior, consider the case in which $1 \geq \Lambda^g > 0$ and $\Lambda^m = 0$. Then, the equilibrium bidding function is given by $b_\alpha(\theta) = (1 - \Lambda^g)b_\alpha^{RN}(\theta) + \Lambda^g \kappa_\alpha(\theta)$, where

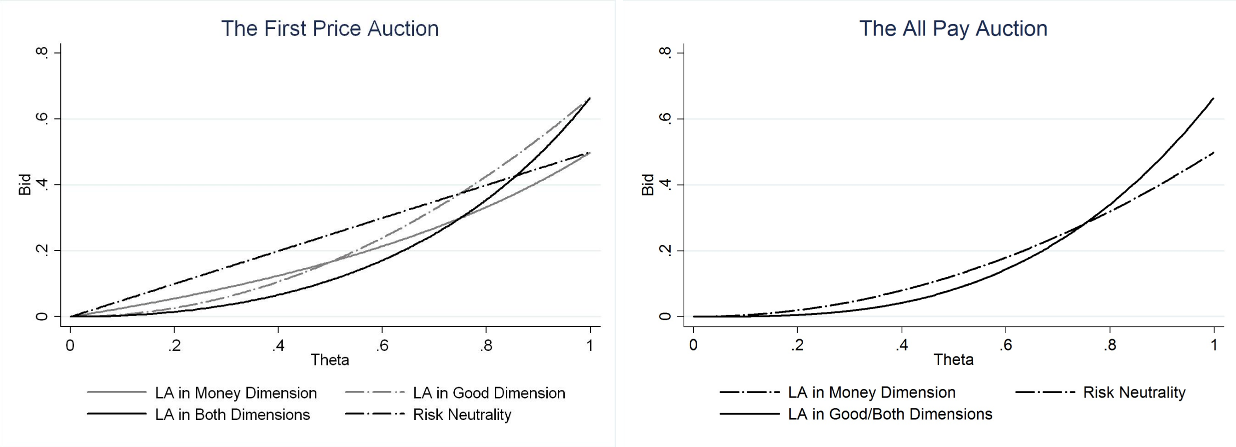
$$\kappa_\alpha(\theta) := \frac{F^{2(N-1)}(\theta)\theta - \int_{\theta_{\min}}^{\theta} F^{N-1}(s)ds}{(1 - \alpha)F^{N-1}(\theta) + \alpha}.$$

The following proposition summarizes the impact of loss aversion in the good dimension on the equilibrium bid.

Proposition 3. *If $0 < \Lambda^g \leq 1$ and $\Lambda^m = 0$, there is a unique interior threshold, $\bar{\theta}$, such that $b_\alpha(\theta)$ is strictly decreasing in Λ^g, for almost all $\theta < \bar{\theta}$, and strictly increasing in Λ^g, for all $\theta > \bar{\theta}$ $\alpha \in [0, 1]$.*

To see the intuition behind this result, again, let $q(x)$ be the interim probability of winning. Then, from an interim perspective, expected gain loss utility in the good dimension is $-\eta^g(\lambda^g - 1)q(x)(1 - q(x))\theta$. As above, the variance of the Bernoulli distributed outcome of winning or losing the auction reduces the interim expected payoff. This implies that the loss is maximized at $q(x) = 1/2$, and minimized at $q(x) = 0$ and $q(x) = 1$. Hence, whenever $q(x)$ is less than $1/2$ a bidder has an incentive to lower $q(x)$ in order to reduce this feeling of loss while whenever $q(x)$ is greater than $1/2$ he has an incentive to increase $q(x)$ in order to lower this feeling of loss. Of course, in equilibrium, the probability of winning for a bidder is the probability that he is the highest type, which is unaffected by loss aversion. Therefore, loss aversion in the good dimension increases the bid of the highest types and reduces the bid of the lowest types through an indirect effect caused by this preference for

Figure 3.1: FPA and APA with $\Lambda^g = 1$, $\Lambda^m = 1$, $N = 2$, and $\theta \sim U[0, 1]$.



certain outcomes. Figure 3.1 depicts the equilibrium bidding functions for $N = 2$ and $\theta \sim U[0, 1]$, compared to the same situation with risk neutral bidders ($\Lambda^g = \Lambda^m = 0$). In the APA, the bidding functions with risk neutrality and loss aversion in the money dimension coincide (the dashed line in the right panel of Figure 3.1).

So far, it has been assumed that assumption 1 is satisfied. As Lange and Ratan (2010) show, if assumption 1 is not met, there are some bidders who choose to not participate in the auction and submit a bid of 0. The argument is that by choosing a bid of 0, a bidder can secure himself a payoff of 0. However, if a set of types of strictly positive measure bid 0, then these types tie at 0 and win with positive probability. Taking into account the ties at 0, the implications of a violation of assumption 1 are the following.

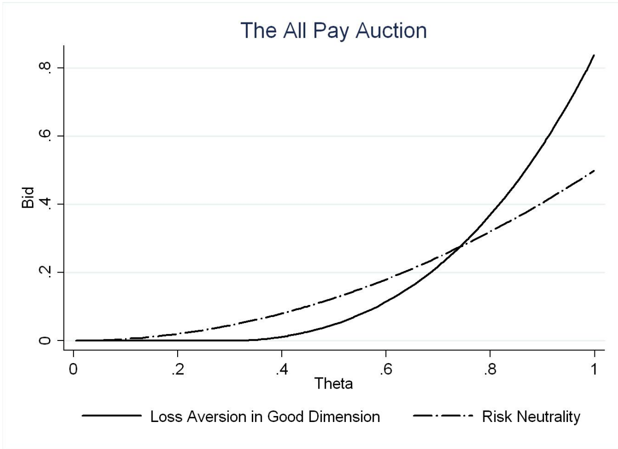
Proposition 4. *Suppose assumption 1 does not hold, i.e.* $\Lambda^g > 1$*. Then, in the unique symmetric pure strategy CPE, there is a unique interior threshold,* $\hat{\theta} \in (\theta_{\min}, \theta_{\max})$*, given by* $F^{N-1}(\hat{\theta}) = (\Lambda^g - 1)/\Lambda^g$*, such that for all* $\theta \geq \hat{\theta}$*,*

$$b_\alpha(\theta) = \frac{\beta(\theta)\theta - \int_{\hat{\theta}}^{\theta} \beta(s)ds}{(1-\alpha)F^{N-1}(\theta)\Delta(\theta) + \alpha},$$

and for all for all $\theta < \hat{\theta}$*,* $b_\alpha(\theta) = 0$*, for all* $\alpha \in [0, 1]$*. Additionally,* $\hat{\theta}$ *is strictly increasing in* Λ^g *and the number of bidders,* N*.*

This result indicates that when loss aversion in the good dimension is too pronounced, there is a set of types of strictly positive measure, for which it is not optimal

Figure 3.2: APA with $\Lambda^g = 2$, $N = 2$, and $\theta \sim U[0, 1]$.



to submit a positive bid. The cut off point, $\hat{\theta}$, is identical across all auction formats, $\alpha \in [0, 1]$. If bidders are given the option of not participating in the auction, with a certain payoff of 0, the cut off in the above proposition changes, but the analysis remains the same. In fact, the cut off with the option of non participation is obtained as the cut off in proposition 4, as $N \to \infty$, since the probability of winning with a winning tie at 0 goes to 0 as the number of bidders increases, and when not participating, the probability of winning is 0, as well. As argued above, the variance of the Bernoulli distributed outcome of winning or losing the auction reduces the equilibrium payoff. The above result says that, depending on the value of Λ^g, this reduction can be too pronounced to make bidding a positive amount worth while for the lowest types, since they have the lowest information rents to start with. If loss aversion is very pronounced ($\Lambda^g > 1$), it is not profitable for the bidders at the bottom of the distribution to take the risk of submitting a positive bid. Loss averse bidders prefer certain outcomes. If gain loss utility dominates intrinsic utility ($\Lambda^g > 1$), then the lowest types have to be compensated for taking the risk associated with participating in the auction, which translates into the non negativity constraint on the submitted bid to be binding for these types. Figure 3.2 depicts the equilibrium bidding function for the APA in the setting of the previous example in Figure 3.1 if assumption 1 is violated ($\Lambda^g = \Lambda^m = 2$).

3.3.2 One Dimension

Similar to the considerations in the previous subsection, with only one dimension, a bidder of type θ solves

$$\begin{aligned} V(\theta) &:= \max_{x\in\mathbb{R}_+}\Big\{q(x)\left(\theta - x + \eta\left(1-q(x)\right)\left(\theta - x(1-\alpha)\right)\right) \\ &\qquad\qquad + \left(1-q(x)\right)\left(-\alpha x - \eta\lambda q(x)(\theta-(1-\alpha)x)\right)\Big\} \\ &= F^{N-1}(\theta)\left(1-\Lambda\left(1-F^{N-1}(\theta)\right)\right)\theta \\ &\quad -F^{N-1}(\theta)\left(1-\Lambda\left(1-F^{N-1}(\theta)\right)\right)(1-\alpha)b_\alpha(\theta) - \alpha b_\alpha(\theta). \qquad (3.2) \end{aligned}$$

Application of the envelope theorem yields the following expression for the symmetric CPE bidding function.

$$b_\alpha(\theta) = \frac{\beta(\theta)\theta - \int_{\theta_{\min}}^{\theta}\beta(s)ds}{(1-\alpha)\beta(\theta)+\alpha},$$

where $\beta(\theta)$ is as defined above. Analogous to the case with two dimensions, the following results hold for one dimensional reference dependence.

Proposition 5. *Suppose assumption 1 holds. Then, $b_\alpha(\theta)$ is strictly increasing, for almost all θ and constitutes the unique symmetric pure strategy CPE bidding function.*

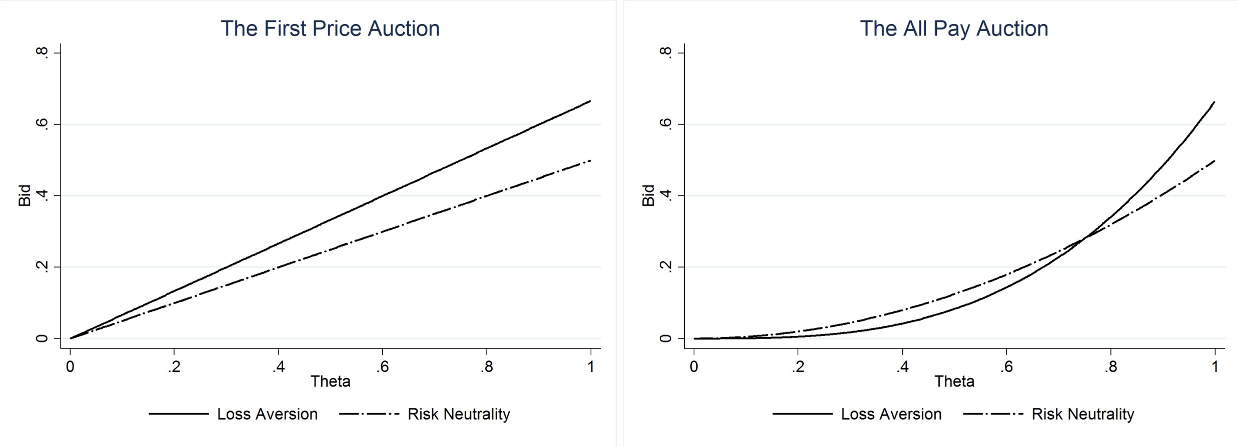
Proposition 6. *Suppose assumption 1 does not hold, i.e. $\Lambda > 1$. Then, in the unique symmetric pure strategy CPE, there is a unique interior threshold, $\hat{\theta} \in (\theta_{\min}, \theta_{\max})$, given by $F^{N-1}(\hat{\theta}) = (\Lambda-1)/\Lambda$, such that for all $\theta \geq \hat{\theta}$,*

$$b_\alpha(\theta) = \frac{\beta(\theta)\theta - \int_{\hat{\theta}}^{\theta}\beta(s)ds - \beta(\hat{\theta})\hat{\theta}}{(1-\alpha)\beta(\theta)+\alpha},$$

and for all $\theta < \hat{\theta}$, $b_\alpha(\theta) = 0$, for all $\alpha \in (0,1]$, and $b_\alpha(\theta) = \theta$, for $\alpha = 0$. Additionally, $\hat{\theta}$ is strictly increasing in Λ and the number of bidders, N.

For all auctions which are not an FPA, the above result is essentially identical to the two dimensional model, and the intuition from above applies. For the FPA, the intuition is now different. Whereas in the two dimensional model and for all auctions with $\alpha \in (0,1]$, there is no possibility for the bidders to secure themselves a payoff

Figure 3.3: FPA and APA with $\Lambda = 1$, $N = 2$, and $\theta \sim U[0, 1]$.



of 0, there is in the FPA, which can be achieved by always bidding the value, θ, so that the ex post payoff is 0, and therefore, also the expected payoff. Regarding the comparative statics properties with respect to the parameter, Λ, the following result holds, as an analogue to propositions 2 and 3.

Proposition 7. *If $\alpha = 1$ and $0 < \Lambda \leq 1$, there is a unique interior threshold, $\bar{\theta}$, such that $b_\alpha(\theta)$ is strictly decreasing in Λ, for almost all $\theta < \bar{\theta}$, and strictly increasing in Λ, for all $\theta > \bar{\theta}$. If $\alpha = 0$, $b_\alpha(\theta)$ is strictly increasing in Λ, for almost all θ.*

As argued above, loss averse bidders exhibit an aversion to the variance and the wedge between the payoff when winning and losing, $\theta - x$ and $-\alpha x$. In the FPA, the wedge between the payoff when winning and losing decreases if the bid increases. This effect drives bidders to increase their bid when the degree of loss aversion increases. Since the CPE bidding function is continuous in α, the above result implies that for auctions with a low enough all pay component, the higher the degree of loss aversion, the higher the bid. Overbidding behavior is observed in the experimental literature (e.g. Filiz-Ozbay and Ozbay (2007) and Noussair and Silver (2006)). In the APA, the bid is always paid for sure, and due to the linearity of the Kahneman and Tversky (1979) value function, does not enter the bidders' gain loss considerations, so that the intuition from proposition 3 applies. Figure 3.3 shows the CPE bidding function in the same environment as the previous examples.

3.3.3 Relationship between CPE and UPE

As mentioned above, in the auction setting under consideration, CPE and UPE are equivalent. For convenience, we first state the definition of UPE.

Definition 3. *(Kőszegi and Rabin (2006)) Conditional on the realization of the type, θ, for any choice set, D, $H \in D$ is an interim CPE if $U(H|H,\theta) \geq U(H'|H,\theta)$, for all $H' \in D$.*

In CPE, the bidder picks a lottery which maximizes his expected payoff taking into account that his reference lottery adjusts accordingly; in UPE, given a reference lottery, the bidder needs to be willing to pick this very lottery. While the interpretation of CPE and UPE is different, every CPE can be supported as a UPE in the auction setting under consideration. In order to see this, suppose that each bidder has some reference lottery, say H. Given this reference lottery, each bidder maximizes his expected payoff by submitting a bid. The essence of the argument involves noticing that given the reference lottery, maximizing the expected payoff over the submitted bid leads to a probability of winning of $F^{N-1}(\theta)$, just as in CPE. More specifically, for any reference lottery, each bidder's payoff maximizing probability of winning is $F^{N-1}(\theta)$. In UPE, both lotteries have to coincide, so that $H = F^{N-1}(\theta)$.

3.3.4 Revenue Non Equivalence

In this section, we compare the expected ex ante revenue for the auctioneer across auction formats. Since the results depend on whether the two or one dimensional model is applied, the revenue properties are discussed separately.

Two Dimensions

As seen above, the interim expected payoff of a bidder of type θ is identical across auction formats. Hence, gain loss considerations in the good dimension leave the interim expected payoff unaffected across auction formats. By proposition 2, each bidder's bid is reduced by loss aversion in the money dimension if $\alpha < 1$; if $\alpha = 1$, then loss aversion in the money dimension has no effect on the equilibrium bid. Hence, for

$\alpha = 1$, bidders bid as if they are risk neutral in the money dimension. Consequently, the interim expected payment of each bidder is reduced by loss aversion in the money dimension if $\alpha < 1$, so revenue equivalence breaks down, as summarized by the following propositions.

Proposition 8. *If bidders are loss averse in the money dimension ($\Lambda^m > 0$), the expected revenue for the auctioneer is strictly increasing in α.*

Gain loss considerations in the money dimension distort the equilibrium bid downwards. By requiring bidders to pay their bid regardless of whether they win the object or not, gain loss distortions in the money dimension are minimized. If $\alpha < 1$, loss averse bidders realize gains in the money dimension if they lose, and losses if they win. Since, under loss aversion, losses loom larger than gains, bidders bid more hesitantly in any auction with $\alpha < 1$ than in the APA. Therefore, among all auctions with fixed all pay component, α, the APA maximizes the auctioneer's expected revenue. In addition, payoff equivalence implies that loss aversion in the good dimension is irrelevant for the revenue ranking across auction formats, as summarized in the following proposition.

Proposition 9. *If bidders are loss averse in the good dimension and risk neutral in the money dimension ($\Lambda^g > 0, \Lambda^m = 0$), the expected revenue for the auctioneer is identical, for all $\alpha \in [0, 1]$.*

This result confirms that the revenue ranking across auction formats is solely driven by loss aversion in the money dimension. An immediate implication is that the revenue ranking is unaffected by whether assumption 1 is met or not. Furthermore, as the bidders' payoff depends solely on the allocation rule, the above results imply the inefficiency of any auction that is not an APA if $\Lambda^m > 0$, since when increasing α bidders remain indifferent, but the auctioneer strictly gains.

One Dimension

As in the two dimensional model, the interim expected payoff for each bidder is identical across auction formats. Consider the payoff from (3.2), which can be rewritten as

$$V(\theta) = F^{N-1}(\theta)\theta - \Big(F^{N-1}(\theta)(1-\alpha) + \alpha\Big) - \Lambda F^{N-1}(\theta)\left(1 - F^{N-1}(\theta)\right)\left(\theta - (1-\alpha)b_\alpha(\theta)\right).$$

Similar to the case with two dimensions, the bidder's payoff is reduced by the variance of the Bernoulli distributed outcome of winning and losing the auction, the degree of loss aversion, and the wedge between the payoff when winning and losing, $\theta-(1-\alpha)b_\alpha(\theta)$. The bidder's objective function satisfies strictly increasing differences in $(\alpha, -x)$, so that $b_\alpha(\theta)$ is non increasing in α, for all θ. Increasing α lowers the bidder's bid, since an increased fraction is paid for sure. In addition, keeping the (non negative) bid fixed, increasing α leads to an increase in the wedge between the payoff when winning and losing. Both of these effects reduce the bidder's payoff. Since the payoff is determined by the allocation rule alone, this payoff reduction is compensated for by bidding less aggressively in the APA compared to the FPA, so that the following result holds.

Proposition 10. *If $\Lambda > 0$, the expected revenue for the auctioneer is strictly decreasing in α.*

3.3.5 Risk Aversion or Loss Aversion?

A natural question to ask is whether the results derived above are driven by risk aversion rather than loss aversion. Auctions with risk averse bidders are studied in Riley and Samuelson (1981), Maskin and Riley (1984), and Matthews (1987), where bidders' preferences take the form $u(\theta, -x)$, and u is strictly increasing and strictly concave in both arguments. As a special case of this formulation, which is studied in Fibich, Gavious, and Sela (2006), bidders' preferences take the form, $u(\theta - x)$, where u is strictly increasing and strictly concave. Fibich, Gavious, and Sela (2006) compare the expected revenue in the APA and the FPA. Their finding is that the revenue ranking is ambiguous in the sense that there are utility functions and distributions for which either the APA or the FPA yields higher expected revenue for the auctioneer. Maskin and Riley (1984) study optimal auctions with risk averse bidders. They find that a *perfect insurance auction* is optimal with homogeneously risk averse bidders, who differ only in their type, θ. A perfect insurance auction is an auction with two payment schemes, one for bidders who win the auction, x^W, and one for bidders who lose the auction, x^L, that depend on the reported type, but are deterministic otherwise, and have the property that for highest type, the marginal utility of money is identical in each state. The APA is nested in the class of perfect insurance auctions, for $x^W = x^L$, and the FPA is nested for $x^L = 0$. The results in Maskin and Riley (1984) imply that a necessary condition for the APA ($x^W = x^L$) to yield the highest expected revenue for the auctioneer is that the marginal utility of money is independent of the valuation, θ, e.g. $u(\theta, -x) = \theta - m(x)$. Furthermore, the insights obtained by Maskin and Riley (1984) rationalize the ambiguous revenue ranking between the APA and the FPA reported in Fibich, Gavious, and Sela (2006). Furthermore, every risk averse bidder with the above preferences is locally risk neutral, which implies that every risk averse bidder participates in the auction and submits a positive bid, because he obtains non negative expected payoff from doing so. As seen above, this is not necessarily the case if bidders are loss averse. This raises the question whether the limited participation results derived above for high degrees of loss aversion in the good dimension can be explained by first order risk aversion. If bidders have rank dependent expected utility preferences as in Yaari (1987), which allow for first order risk aversion, the revenue ranking and the participation is as with additively separable risk aversion[1], i.e. the APA yields the highest revenue for the auctioneer.

3.4 Experiment

As seen above, the revenue ranking between the FPA and the APA is opposite in the one and two dimensional model. In this section, we describe the experiments, which are designed to test the theoretical results derived above and which contribute to a better understaning of methodological robustness concerning laboratory and field experiments. A common method in experimental economics to analyze IPV auctions is the induced value (IV) method, where money is auctioned. Bidders are assigned a randomly drawn valuation, and if they win, they receive a monetary payoff equal to their valuation. This is in contrast to a real object (RO) auction, where actual goods

[1]The results can be obtained from the authors upon request.

are auctioned. Since money is auctioned in the IV method, only the one dimensional model is applicable in this setting; for RO auctions, the two dimensional model is more plausible. Therefore, we examine revenue equivalence between the APA and the FPA for RO and IV auctions separately and compare the results of the IV and RO method. Based on the theoretical results above, we seek to test the following hypotheses.

Hypothesis 1. *IV Auction: The expected revenue is higher for the FPA than for the APA (Proposition 10).*

Hypothesis 2. *IV Auction: Almost every bidder submits a positive bid in the FPA, but not necessarily in the APA (Proposition 6).*

Hypothesis 3. *RO Auction: The expected revenue is higher for the APA than for the FPA (Proposition 8).*

Hypothesis 4. *RO Auction: The fraction of bidders submitting zero bids is identical in the APA and in the FPA (Proposition 4).*

Two of the above hypotheses can be tested with revenue data alone; the other two require data on the bidders' behavior. The ability to test using revenue data alone is due to the diverging revenue rankings in the one and two dimensional model of reference dependence for the APA and the FPA. Comparing the FPA to the Vickrey auction, Lange and Ratan (2010) find that the theoretical revenue ranking is identical in the one and two dimensional model, so one can only reject (or not reject) both models together when using revenue data.

3.4.1 Experimental Design

In order to experimentally test our hypotheses, we employ two different methodological approaches. Both approaches are run in the laboratory instead of in the field to contribute to the analysis of auctions with experimental methods. For the IV method, subjects were anonymously matched into groups of three. Each subject was given an endowment of 700 points, where 100 points are worth 1€, which was \$1.42 at that time, to submit a bid in the auction. Subjects' valuations in the auction were independently drawn from the uniform distribution on $\{0, 1, 2, \ldots, 299, 300\}$, which was made common knowledge to the subjects. The maximum valuation is therefore 300 points and lower than the endowment. Bids were allowed to have up to two decimal points, e.g. 2.99, as we did not want bidders with low valuations to floor their bids down to zero, which would inhibit the testability of our hypothesis on limited participation. The important part of the design of the induced value method is the following. Subjects were provided with a list of ten different valuations, such that a subject participated in ten auctions. The subjects had to bid for each valuation, however only one of the ten auctions was payoff relevant, and each auction was equally likely to be payoff relevant. We adapted this procedure from Filiz-Ozbay and Ozbay (2007). Figure 3.6 in the Appendix 3.B shows an example of such a list. In the FPA, the subject with the highest bid in the group of three won the auction and received the valuation plus the endowment minus the respective bid as his or her payoff. If more than one subject was the highest bidder, the computer chose the winner with equal probability. The other subjects lost the auction and received only the endowment. The rules of the APA were exactly the same, except for the losers' payoff, which was the endowment minus the bid. Finally, to secure the understanding of the game, we asked several control questions (see Appendix 3.B). In the RO auction, we auctioned a real good. The good was chosen, such that subjects valuations do not largely differ from the induced value auction, and are plausibly independent and private. Therefore, we decided to auction a blackboard cup with a piece of chalk. The cup has a blackboard sheathing, on which can be written with chalk (see Figure 3.5 in Appendix 3.B). Every subject had the possibility to have a look at the cup before bidding in the auction. The buying price of the cup was 1.75 euros, which was not revealed to the subjects. As in the IV auction, subjects received an initial endowment of 700 points before the auction began. In contrast to the IV auction, each subject can only take part in one RO auction. The rules of the auction formats are the same as in the IV method. The IV auction and the RO auction were played on different days with different subjects in the BonnEconLab at the University of Bonn with 192 participants from various fields of study, recruited via ORSEE Greiner (2004), out of which 96 subjects participated in the IV auction and 96 in the RO auction. The experiment was programmed and conducted with the software z-Tree by Fischbacher

(2007).

3.4.2 Results

Result 1. *The average revenue for the FPA is significantly greater than for the APA, for both the IV and the RO auction.*

A group of three subjects bids in one auction. Since 24 subjects are in one session, we have eight groups per session. In the induced value method, each subject had to submit ten bids for ten valuations. Thus, each group performs ten auctions and we receive ten revenues per group. We take the average of the respective ten revenues of one group as an independent observation and therefore use 8 independent revenue observations per session for the data analysis. The average revenue in the IV method for the FPA is 170 points and 152 points in the APA. Summary statistics of the data collected are shown in Table 3.1. We will use t-tests to show whether this sizeable difference in means is also significant. Performing a one sided t-test, we reject the hypothesis that the average revenues are equal in favor of the alternative hypothesis that the average revenue of the FPA is greater than the average revenue of the APA with $p = 0.04$, which confirms our Hypothesis 1, that the revenue of the FPA is larger than of the APA.[2] The results are similar in the RO auction, with an even larger difference in average revenue, which is 263 points for the FPA and 150 points for the APA. Therefore, our data does not confirm Hypothesis 2, that the opposite should occur when bidders think in two dimensions. Using a two sided t-test rejects the hypothesis that average revenues are equal with a $p - value = 0.03$. For the IV auction, the results are consistent with the model of one dimensional reference dependence. In the RO auction, the results reject the model of two dimensional reference dependence, yet are consistent with the one dimensional model. One possible explanation for this is that although a real object is auctioned, subjects behave according to the one dimensional model of reference dependence.

[2] The p-value of a two sided t-test is 0.08.

Table 3.1: Summary statistics

	Induced Value Auction				Real Object Auction			
	Mean	Std. Dev.	N	Zero Bids	Mean	Std. Dev.	N	Zero Bids
FPA	170	29.8	16	2.5 %	263	150	16	8.3 %
APA	152	26.5	16	27.7 %	150	129	16	39.6 %

N is the number of independent valuations

Result 2. *The fraction of bidders submitting zero bids is significantly greater in the APA than in the FPA, for both the IV and the RO auction.*

A strikingly large amount of subjects submits a bid of zero in the APA, 27.7% in the IV auction and 39.6 % in the RO and only 2.5% and 8.3% respectively for the FPA. Using a t-test, we find that zero bids occur significant more often in the APA compared to the FPA with $p < 0.01$. As with the revenue data, these findings are consistent with the one dimensional model fo reference dependence, but not with the two dimensional one.

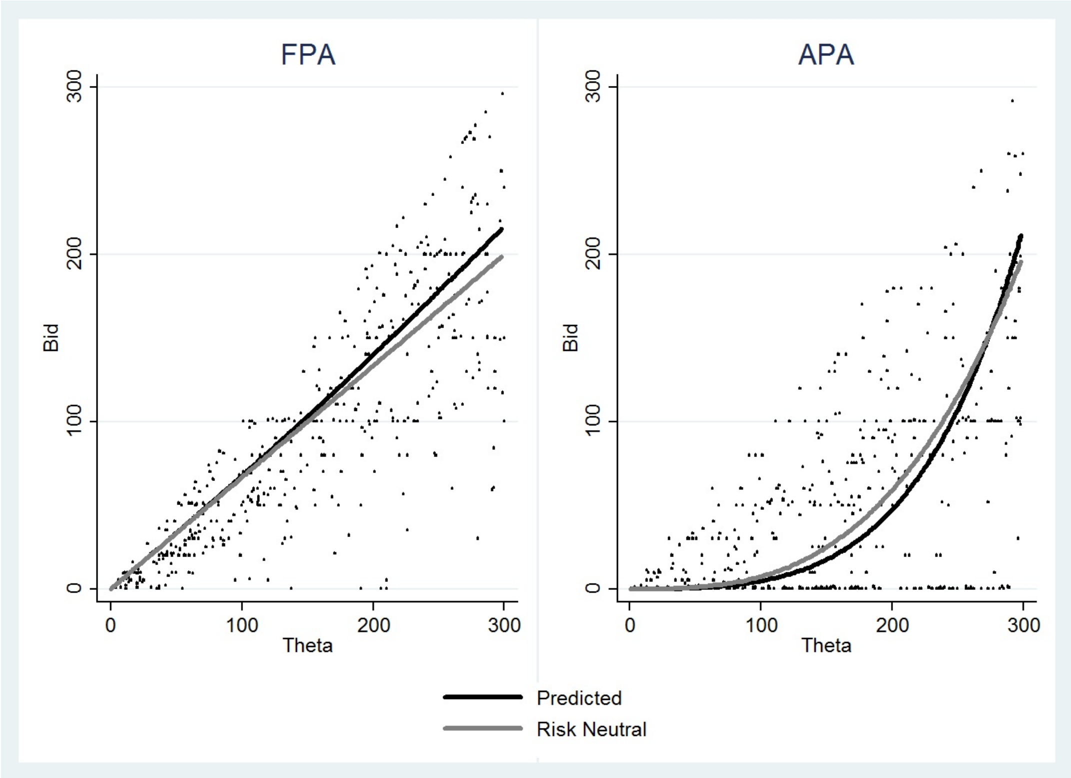
Structural Estimation For the Induced Value Method

The above results consider the revenue data alone. While we do not have data on the valuations in the real object auctions, we have the induced valuations and the submitted bids in the induced value auctions, which enables to structurally estimate the parameter, Λ. By doing this, we can compare our estimate to the ones reported in the literature and obtain an internal consistency check. We employ the Generalized Method of Moments (GMM) using the moment conditions,

$$E[b_0] = E[b_0(\theta|\Lambda)] \qquad \text{and} \qquad E[b_1] = E[b_1(\theta|\Lambda)].$$

We estimate $\hat{\Lambda} = 0.42$ with a standard error of 0.16, which is statistically different from 0 and 1 at all conventional significance levels using a Wald test. Since we have

Figure 3.4: Bids with GMM estimate, $\hat{\Lambda} = 0.42$.



two moment conditions and only one parameter to estimate, we perform a J-test for overidentifying restrictions, which does not reject the null that the model is valid at all conventional significance levels ($\chi^2(1) = 2.31$). The following figure depicts the predicted bids compared to the risk neutral benchmark. Recall that $\Lambda = \eta(\lambda - 1)$, so that η and λ are not identified. Once we normalize $\eta = 1$, we can identify $\hat{\lambda} = 1.42$.

3.5 Conclusion

While the above analysis is robust to considering more general, non linear specifications of the Kahneman and Tversky (1979) value function in the two dimensional model, the analysis becomes analytically intractable for non linear functions in the one dimensional model. By continuity and the revenue ranking being strict, it follows that introducing small amounts of non linearities will not change any of the above results. However, as in Fibich, Gavious, and Sela (2006), no closed form expression for the equilibrium bidding function can be obtained, and one has to resort to approximate perturbation methods. As the above analysis shows, the implications of bidders with reference dependent preferences in auction environments differ depending on which specification of preferences (one or two dimensional) is assumed. In a working paper preceding their published articles, Köszegi and Rabin (2004) discuss the distinction between consumption and hedonic dimensions in their model of reference dependent preferences. The dimensions across which gambles are considered separately are not necessarily equal to physical consumption dimensions. They give the example of peanut butter, jelly, and bread, where one is plausibly interested only in the total number of sandwiches produced. Which of the two specifications studied above is more appropriate in which context demands further research. The separable model is commonly used in applications, e.g. Köszegi and Rabin (2006), Heidhues and Köszegi (2008), Herweg, Müller, and Weinschenk (2010). As the above analysis illustrates, the theoretical implications may differ drastically, depending on which specification of reference dependent preferences is employed. Using experiments, we can reject the two dimensional model, but not the one dimensional model.

3.6 Appendix

Appendix 3.A

Proof of Proposition 1: If assumption 1 is satisfied, then

$$F^{N-1}(\theta)\left(1-\Lambda^g(1-F^{N-1}(\theta))\right) = F^{N-1}(\theta)(1-\Lambda^g) + F^{2(N-1)}(\theta)\Lambda^g,$$

is strictly increasing in θ, so the payoff is strictly increasing in the type. Differentiating $b_\alpha(\theta)$ with respect to θ gives

$$\frac{\partial}{\partial\theta}b_\alpha(\theta) > 0 \Longleftrightarrow (1-\Lambda^g+2F^{N-1}(\theta)\Lambda^g)\theta - b_\alpha(\theta)(1-\alpha)(1+\Lambda^m-2\Lambda^m F^{N-1}(\theta)) > 0 \quad (3.3)$$

(3.3) is equivalent to

$$\begin{aligned}
& (1-\Lambda^g+2F^{N-1}(\theta)\Lambda^g)F^{N-1}(\theta)\theta - b_\alpha(\theta)(1-\alpha)F^{N-1}(\theta)(1+\Lambda^m-2\Lambda^m F^{N-1}(\theta)) > 0 \\
\Longleftrightarrow\quad & \beta(\theta)\theta + \Lambda^g F^{N-1}(\theta) - b_\alpha(\theta)(1-\alpha)F^{N-1}(\theta)(1+\Lambda^m-\Lambda^m F^{N-1}(\theta)) + b_\alpha(\theta)\Lambda^m F^{N-1}(\theta) > 0 \\
\Longleftrightarrow\quad & V(\theta) + \Lambda^g F^{N-1}(\theta) + b_\alpha(\theta)\Lambda^m F^{N-1}(\theta) > 0
\end{aligned}$$

which is true since, the expected payoff is positive. Additionally, since b_α is strictly increasing, the envelope representation of the bidder's payoff applies, the type space is convex and has no mass points, uniqueness follows from Myerson (1981). □

Proof of Proposition 2: Immediate by inspection. □

Proof of Proposition 3:

$$\begin{aligned}
& \frac{\partial}{\partial\Lambda^g}b_\alpha(\theta) = -b_\alpha^{RN}(\theta) + \kappa_\alpha(\theta) \leq 0 \Longleftrightarrow \kappa_\alpha(\theta) \leq b_\alpha^{RN}(\theta) \\
\Longleftrightarrow\quad & F^{2(N-1)}(\theta)\theta - \int_{\theta_{\min}}^{\theta} F^{2(N-1)}(s)ds \leq F^{N-1}(\theta)\theta - \int_{\theta_{\min}}^{\theta} F^{N-1}(s)ds.
\end{aligned}$$

For $\theta = \theta_{\min}$, both the LHS and the RHS of the above expression are equal to 0. The derivative of the expression on the RHS is greater than the derivative of the expression on the LHS if and only if $F^{N-1}(\theta) \leq 1/2$, which implies that the bid of the lowest types is always reduced by an increase in Λ^g. For $\theta = \theta_{\max}$, the bidding function is increased by an increase in Λ^g. $\square$

Proof of Proposition 4: If $\Lambda^g > 1$, then $\beta(\theta)$ is minimized at $F^{N-1}(\theta) = (\Lambda^g - 1)/(2\Lambda^g)$, since $\beta(\theta)$ is quadratic in $F^{N-1}(\theta)$, so always increasing above its minimum and decreasing below it. Therefore, $\beta(\theta) < 0$, for $F^{N-1}(\theta) < (\Lambda^g - 1)/\Lambda^g$. These two facts will be used in the remainder of the proof. If all $\theta \leq \hat{\theta}$ bid 0, then they face a probability of winning of $F^{N-1}(\hat{\theta})/N$, since $F^{N-1}(\hat{\theta})$ is the probability of winning, in which case the object is allocated among the winning bidders with equal probability. Hence, one candidate symmetric equilibrium bidding function is $b_\alpha(\theta) = 0$, if $\theta < \hat{\theta}$ and

$$b_\alpha(\theta) = \frac{\beta(\theta)\theta - \int_{\hat{\theta}}^{\theta} \beta(s)ds - (\hat{\theta} - \theta_{\min})\beta(\hat{\theta}) - V(\theta_{\min})}{\alpha + (1-\alpha)F^{N-1}(\theta)\Delta(\theta)} = \frac{\beta(\theta)\theta - \int_{\hat{\theta}}^{\theta} \beta(s)ds - \beta(\hat{\theta})\hat{\theta}}{\alpha + (1-\alpha)F^{N-1}(\theta)\Delta(\theta)},$$

for all $\theta \geq \hat{\theta}$, where the last equality follows since $V(\theta_{\min}) = \beta(\hat{\theta})\theta_{\min}$. Following this strategy, all $\theta < \hat{\theta}$ receive an interim expected payoff of $V(\theta) = \beta(\hat{\theta})\theta$. Since $N \geq 2$ and $\beta(\theta)$ is quadratic in $F^{N-1}(\theta)$, so always decreasing below its minimum, $V(\theta) = \beta(\hat{\theta})\theta < 0$, for all $\theta < \hat{\theta}$. Consider a deviation of a type $\theta < \hat{\theta}$ to any bid $b_\alpha(\theta_{\max}) > b > 0$. Then there is a $\theta_{\max} > \theta^* > \hat{\theta}$ with $b_\alpha(\theta^*) = b$. Hence, the payoff from the deviation is

$$\begin{aligned} & \beta(\theta^*)(\theta - \theta^*) + \int_{\theta_{\min}}^{\theta^*} \beta(s)ds + \beta(\hat{\theta})\theta_{\min} = \beta(\theta^*)(\theta - \theta^*) + \int_{\hat{\theta}}^{\theta^*} \beta(s)ds + \beta(\hat{\theta})\hat{\theta} && (3.4)\\ < \; & \beta(\theta^*)(\theta - \theta^*) + (\theta^* - \hat{\theta})\beta(\theta^*) + \beta(\hat{\theta})\hat{\theta} = \beta(\theta^*)(\theta - \hat{\theta}) + \beta(\hat{\theta})\hat{\theta} && (3.5)\\ = \; & \beta(\theta^*)(\theta - \hat{\theta}) + \beta(\hat{\theta})\hat{\theta} + \beta(\hat{\theta})\theta - \beta(\hat{\theta})\theta = (\beta(\theta^*) - \beta(\hat{\theta}))(\theta - \hat{\theta}) + \beta(\hat{\theta})\theta, && (3.6) \end{aligned}$$

where the above inequality follows from $\beta(\theta)$ being quadratic and increasing above its minimum and $N \geq 2$. The payoff from deviating is greater than the payoff from following the strategy only if $(\beta(\theta^*) - \beta(\hat{\theta}))(\theta - \hat{\theta}) \geq 0$, which is never satisfied, since $\theta < \hat{\theta}$ and $\beta(\theta^*) > \beta(\hat{\theta})$. If types above $\tilde{\theta}$ deviate and submit a bid of 0, they earn a payoff of $\beta(\tilde{\theta})\theta$, and if they stick to the above candidate CPE strategy, they earn a payoff of $\beta(\tilde{\theta})\tilde{\theta} + \int_{\tilde{\theta}}^{\theta} \beta(s)ds$. Since $\beta(\theta) > 0$, for all $\theta > \tilde{\theta}$, a deviation is profitable only if $\beta(\tilde{\theta})\tilde{\theta} < \beta(\tilde{\theta})\theta$, which is never true since $\beta(\tilde{\theta}) < 0$, and $\theta > \tilde{\theta}$. Finally, consider the threshold type, $\tilde{\theta}$. This type is indifferent between deviating downwards or upwards, so that there is no profitable deviation. Hence, the second candidate CPE bidding function constitutes a symmetric pure stategy CPE. Since $V(\theta) \geq 0$, for all types who submit a positive bid, the argument in the proof of proposition 2 shows that $b_\alpha(\theta)$ is strictly increasing for all $\theta \geq \hat{\theta}$, so that for all $\theta > \tilde{\theta}$, Myerson (1981)'s condition implies that the symmetric CPE bidding function is unique. Suppose that there is another cut off with the equality defining $\hat{\theta}$ as a strict inequality in either direction. Then, either the types slightly above or slightly below this cut off can earn a higher payoff by deviating to either bidding 0 or a slightly positive amount or the bidding function is not strictly increasing, which cannot be part of a symmetric CPE. □

Proof of Proposition 5: Exactly the same as the proof of proposition 1. □

Proof of Proposition 6: In the one dimensional model, the analysis is different for $\alpha \in (0, 1]$ and $\alpha = 0$. Consider first all auction except the FPA and the the following candidate symmetric equilibrium bidding function

$$b_\alpha(\theta) = \frac{\beta(\theta)\theta - \int_{\hat{\theta}}^{\theta} \beta(s)ds - (\hat{\theta} - \theta_{\min})\beta(\hat{\theta}) - V(\theta_{\min})}{\alpha + (1-\alpha)F^{N-1}(\theta)\Delta(\theta)} = \frac{\beta(\theta)\theta - \int_{\hat{\theta}}^{\theta} \beta(s)ds - \beta(\hat{\theta})\hat{\theta}}{\alpha + (1-\alpha)\beta(\theta)},$$

and $b_\alpha(\theta) = 0$, for all $\theta < \hat{\theta}$. The steps proving that this candidate CPE bidding function constitutes the unique symmetric COPE biding function is exactly the same as in proposition 4. Consider now the FPA. If all types below $\hat{\theta}$ bid 0, they w=tie and win with strictly positive probability, which yields a negative expected payoff, since $\beta(\theta) < 0$, for all $\theta < \hat{\theta}$. Instead of bidding 0, these types can secure an expected payoff of 0 if they bid their valuation rather than 0. Therefore, we have the following candidate CPE bidding function,

$$b_0(\theta) = \theta - \frac{\int_{\hat{\theta}}^{\theta} \beta(s)ds}{\beta(\theta)},$$

for all $\theta \geq \hat{\theta}$, and $b_0(\theta) = \theta$, for all $\theta < \hat{\theta}$. By L'Hopital's rule, $\lim_{\theta \downarrow th\hat{e}ta} = \theta$, so that this candidate CPE bidding function is continuous and strictly increasing. Uniqueness follows from Myerson (1981)'s condition. □

Proof of Proposition 7: The first part of the proof is exactly the same as the proof of proposition 3, the second part follows from differentiation of the bidding function for $\alpha = 0$ with respect to Λ. This expression is equal to 0 for $\theta = \theta_{\min}$, and strictly negative for $\theta = \theta_{\max}$, in addition to being strictly decreasing in θ. □

Proof of Proposition 8: This result follows from the interim expected payoff only depending only on the probability of winning, which is identical for a given type in all auctions with the same allocation rule. □

Proof of Proposition 9: The expected payment, $p_\alpha(\theta)$, of a bidder of type, $\theta \geq \hat{\theta}$, conditional on the other bidders' behavior, is $p_\alpha(\theta) = \alpha b_\alpha(\theta) + F^{N-1}(\theta)(1-\alpha)b_\alpha(\theta)$, i.e. αb_α with certainty, and $(1-\alpha)b_\alpha$ only if he wins, which happens with probability $F^{N-1}(\theta)$. Differentiating the above expression with respect to α yields

$$\frac{\partial}{\partial \alpha} p_\alpha(\theta) = \frac{F^{N-1}(\theta)(\Delta(\theta) - 1)}{(1-\alpha)F^{N-1}(\theta)\Delta(\theta) + \alpha} b_\alpha(\theta),$$

which is strictly positive, for all $\theta > \hat{\theta}$, since $\Delta(\theta) > 1$, for all $\theta > \hat{\theta}$. For $\theta < \hat{\theta}$, the interim expected payment is 0 and remains unchanged in α. Since the interim expected payment is non decreasing for all types and strictly increasing for a set of types of strictly positive measure, this implies that the ex ante expected revenue for the auctioneer, $N \int p_\alpha(\theta) dF(\theta)$, is strictly increasing in α. □

Proof of Proposition 10: Replace Δ by $\beta(\theta)/F^{N-1}(\theta)$ in the proof of proposition 8. □

Appendix 3.B

Instructions, translated into English. General instructions were identical across treatments. Instructions for the real object auction and induced value auction differed.

GENERAL INSTRUCTIONS

You are taking part in a decision-making experiment in which you have the opportunity to earn money. The amount of money you earn is paid to you upon completion of the experiment. Please read the instructions carefully. The instructions are identical for all participants. If you have any questions, please raise your hand. The experimenter will answer your question at your place. During the experiment, you have to remain silent. Violation of this rule leads to immediate exclusion from the experiment and all payments.

All monetary units in the experiment are measured in points, and 100 points = 1 Euro.

INSTRUCTIONS *Real object auction*

You now take part in an auction. For this you get an endowment of 700 points.

Task

At the beginning of the auction you will be divided into groups of three. You will not learn who the other participants in your group are. Your task in the three-group is that of a bidder who bids for an item in an auction. For this you can spend an arbitrary amount of your endowment of 700 points. Your bid must have a maximum of two decimal places.

The item

The auctioned item is a chalk-cup with one piece of chalk. The cup can always be rewritten.

Rules

The auction rules for each three-person group are that the participant with the highest bid wins the auction in their group and thus the cup. If several bidders have the same highest bid, we will then toss a coin to determine the winner.

As the winner you will receive the cup, plus the endowment of 700 points minus your bid:

Payoff = cup + 700 - your bid.

Figure 3.5: The chalk-cup for the real object auction

First price auction instructions

If your bid is less than the highest bid, you lose the auction. As a loser, you get the endowment of 700 points:

Payoff = 700.

All pay auction instructions

If your bid is less than the highest bid, you lose the auction. As a loser, you get the endowment of 700 points minus your bid:

Payoff = 700 - bid.

Any questions?

Please enter your cabin number and your bid.

Cabin number:

Bid:

INSTRUCTIONS *Induced value auction*

You now take part in an auction. For this you get an endowment of 700 points.

Task

At the beginning of the auction you will be divided into groups of three. You will not learn who the other participants in your group are. Your task in the three-group is that of a bidder who bids for a fictious item in an auction. For this you can spend an arbitrary amount of your endowment of 700 points. Your bid must have a maximum of two decimal places.

Your value

Before the auction starts, you will see on your computer screen a list of 10 numbers. Each of these numbers is between 0 and 300 points. The numbers are chosen randomly by the computer, where each number can occur with equal probability. Each number represents a possible value for you for the fictious item in the auction. The process of generating the values is identical for all participants. This means that every participant in your group of three got a different list of 10 numbers, where each number is chosen randomly and independently from your numbers from the interval of [0,300].

We ask you to enter a bid for each of your 10 possible values in the column next to the values. For this you can spend an arbitrary amount of your endowment of 700 points.

Thus, you enter bids for ten auctions. However, only one of the ten auctions performed will be payoff relevant. The computer will randomly select one of the ten auctions, where each auction is equally likely. This means that you should enter for each of the ten possible auctions your bid such as if it were the only auction that is conducted. So, for each auction you have an endowment of 700 points and its value on which you can bid is a number between 0 and 300 points.

Rules

The auction rules for each three-person group are that the participant with the highest bid wins the auction in their group and thus the cup. If several bidders have the same highest bid, we will then toss a coin to determine the winner.

Figure 3.6: Example of a screen with an auction list

As the winner you will receive the cup, plus the endowment of 700 points minus your bid:

Payoff = cup + 700 - your bid.

First price auction instructions

If your bid is less than the highest bid, you lose the auction. As a loser, you get the endowment of 700 points:

Payoff = 700.

All pay auction instructions

If your bid is less than the highest bid, you lose the auction. As a loser, you get the endowment of 700 points minus your bid:

Payoff = 700 - bid.

Do you have any questions on this?

After you have entered all 10 bids on the screen, please press the OK button. You are then asked again to confirm your choices and you can once again decide whether you want to make changes.

The auction begins now with several control questions to ensure that all participants understand the rules.
Any questions?

Control questions on screen

Question 1: What is the smallest value you can get?
Question2: What is the highest value you can get?

Example of an auction: Player 1 has a value of 1 and bids 1, player 2 has a value of 20 and bids 2 and player 3 has a value of 30 and bids 3.
Question 3: Who wins the auction?
Question 4: What is the payoff for player 1?
Question 5: What is the payoff for player 2?
Question 6: What is the payoff for player 3?

Bibliography

Andreoni, J. and B. Bernheim. 2009. "Social Image and the 50-50 Norm: A Theoretical and Experimental Analysis of Audience Effects." *Econometrica* 77(5):1607–1636.

Andreoni, J. and R. Petrie. 2004. "Public Goods Experiments Without Confidentiality: A Glimpse into Fundraising." *Journal of Public Economics* 88:1605–1623.

Ariely, D., A. Bracha, and S. Meier. 2009. "Doing Good or Doing Well? Image Motivation and Monetary Incentives in Behaving Prosocially." *American Economic Review* 99(1):544–555.

Baldwin, J. R. 1995. *The Dynamics of Industrial Competition.* Cambridge University Press.

Bartling, B., E. Fehr, M. Marechal, and D. Schunk. 2009. "Egalitarianism and Competitiveness." *American Economic Review* 99(2):93–98.

Bem, D.J. 1972. "Self-Perception Theory." *Advances in Experimental Social Psychology* 6:1–62.

Bénabou, R. and J. Tirole. 2002. "Self-Confidence and Personal Motivation." *Quarterly Journal of Economics* 117(3):871–915.

———. 2004. "Willpower and Personal Rules." *Journal of Political Economy* 112(4):848–886.

———. 2006. "Incentives and Prosocial Behavior." *American Economic Review* 96(5):1652–1678.

Benoît, J.-P. and J. Dubra. 2011. "Apparent Overconfidence?" *Econometrica* 79(5):1591–1625.

Brunnermeier, M. and J. Parker. 2005. "Optimal Expectations." *American Economic Review* 95(4):1092–1118.

Burks, S., J. Carpenter, L. Goette, and A. Rustichini. 2010. "Overconfidence is a Social Signaling Bias." *Working Paper* .

Caliendo, M., A. Kritikos, V. Steiner, and F. Wießner. 2007. "Existenzgründungen - Unterm Strich ein Erfolg." *IAB Kurzbericht* 10.

Camerer, C. and D. Lovallo. 1999. "Overconfidence and Excess Entry: An Experimental Approach." *American Economic Review* 89(1):306–318.

Charness, G., A. Rustichini, and J. van de Ven. 2011. "Self-Confidence and Strategic Deterrence." *Working Paper* .

Choi, J., D. Laibson, B. Madrian, and A. Metrick. 2003. "Optimal Defaults." *American Economic Review* 93:180–185.

Cooper, A., C. Woo, and W. Dunkelberg. 1988. "Entrepreneurs' Perceived Chances for Success." *Journal of Business Venturing* 3(2):97–108.

Cox, J., B. Roberson, and V. Smith. 1988. "Theory and Individual Behavior of First-Price Auctions." *Research in Experimental Economics* 1(1):61–99.

Cox, J., V. Smith, and J. Walker. 1982. "Theory and Behavior of Single-Object Auctions." *Journal of Risk and Uncertainty* 1:537–579.

Daniel, K. D., D. Hirshleifer, and A. Subrahmanyam. 2001. "Overconfidence, Arbitrage, and Equilibrium Asset Pricing,." *Journal of Finance* 56(3):921–965.

DellaVigna, S. and U. Malmendier. 2006. "Paying Not to Go to the Gym." *American Economic Review* 96(3):694–719.

Dohmen, T. and A. Falk. 2011. "Performance Pay and Multi-Dimensional Sorting: Productivity, Preferences and Gender." *American Economic Review* 101 (2):556–590.

Dunning, J., J.A. Meyerowitz, and A.D. Holzberg. 1989. "Ambiguity and Self-Evaluation: The Role of Idiosyncratic Trait Definitions in Self-serving Assessments of Ability." *Journal of Personality and Social Psychology* 57(6):1082–1090.

Eil, D. and J. Rao. 2011. "The Good News-Bad News Effect: Asymmetric Processing of Objective Information about Yourself." *American Economic Journal* 3(2):114–138.

Eisenhuth, R. 2012. "Reference Dependent Mechanism Design." *Working Paper* .

Ellingsen, T. and M. Johannesson. 2008. "Pride and Prejudice: The Human Side of Incentive Theory." *American Economic Review* 98(3):990–1008.

Englmeier, F. 2004. "A Brief Survey on Overconfidence." *Working Paper* .

Ewers, M. and F. Zimmermann. 2011. "Image and Misreporting." *Working Paper* .

Falk, A. and F. Zimmermann. 2011. "Preferences for Consistency." *Working Paper* .

Fibich, G., A. Gavious, and A. Sela. 2006. "All-Pay Auctions with Risk Averse Players." *International Journal of Game Theory* 4:583–599.

Filiz-Ozbay, E. and E. Ozbay. 2007. "Auctions with Anticipated Regret: Theory and Experiment." *American Economic Review* 97:1407–1418.

Fischbacher, U. 2007. "z-Tree: Zurich Toolbox for Ready-made Economic Experiments." *Experimental Economics* 10(2):171–178.

Fischbacher, U. and F. Heusi. 2008. "Lies in Disguise: An Experimental Study on Cheating." *TWI Research Paper Series* 40.

Gächter, S. and E. Fehr. 1999. "Collective Action as a Social Exchange." *Journal of Economic Behavior and Organization* 39(4):341–369.

Gneezy, U. 2005. "Deception: The Role of Consequences." *American Economic Review* 3:384–394.

Gneezy, U., M. Niederle, and A. Rustichini. 2003. "Performance in Competitive Environments: Gender Differences." *Quarterly Journal of Economics* :1049–1074.

Gneezy, U. and A. Rustichini. 2004. "Gender and Competition at a Young Age." *American Economic Review Papers and Proceedings* 94(2):377–381.

Greiner, B. 2004. "The Online Recruitment System ORSEE 2.0 - A Guide for the Organization of Experiments in Economics." *Working Paper Series in Economics* 10.

Healy, P.J. and D.A Moore. 2008a. "Bayesian Overconfidence." *Working Paper* .

———. 2008b. "The Trouble with Overconfidence." *Psychological Review* 115(2):502–517.

Heidhues, P. and B. Köszegi. 2008. "Competition and Price Variation when Consumers are Loss Averse." *American Economic Review* 98:1245–1268.

Herweg, F., D. Müller, and P. Weinschenk. 2010. "Binary Payment Schemes: Moral Hazard and Loss Aversion." *American Economic Review* 100:2451–2477.

Hoelzl, E. and A. Rustichini. 2005. "Overconfident: Do You Put Your Money on It?" *The Economic Journal* 115(4):305–318.

Kahneman, D., J. L. Knetsch, and R. Thaler. 1990. "Experimental Tests of the Endowment Effect and the Coase Theorem." *Journal of Political Economy* 98:1325.

Kahneman, D. and A. Tversky. 1979. "Prospect Theory: An Analysis of Decisions under Risk." *Econometrica* 49:263–291.

Kőszegi, B. 2006. "Ego Utility, Overconfidence, and Task Choice." *Journal of the European Economic Association* 4(4):673–707.

Koellinger, P., M. Minniti, and C. Schade. 2007. "'I think I can, I think I can': Overconfidence and Entrepreneurial Behavior." *Journal of Economic Psychology* 28:507–527.

Köszegi, B. and M. Rabin. 2004. "A Model of Reference-Dependent Preferences." Mimeo, University of California, Berkeley.

———. 2006. "A Model of Reference-Dependent Preferences." *Quarterly Journal of Economics* 121:1133–1166.

———. 2007. "Reference-Dependent Risk Attitudes." *American Economic Review* 97:1047–1073.

Lange, A. and A. Ratan. 2010. "Multi-Dimensional Reference-Dependent Preferences in Sealed-Bid Auctions: How (most) Laboratory Experiments differ from the Field." *Games and Economic Behavior* 68:634–645.

Lucking-Reiley, D. 1999. "Using Field Experiments to Test Equivalence Between Auction Formats: Magic on the Internet." *American Economic Review* 89:1063–1080.

Ludwig, S. and J. Nafziger. 2011. "Beliefs about Overconfidence." *Theory and Decision* 70(4):475–500.

Malmendier, U. and G. Tate. 2008. "Who Makes Acquisitions? CEO Overconfidence and the Market's Reaction." *Journal of Financial Economics* 89 (1):20–43.

Maskin, E. and J. Riley. 1984. "Optimal Auctions with Risk Averse Buyers." *Econometrica* 52:1473–1518.

Matthews, S. 1987. "Comparing Auctions for Risk Averse Buyers: A Buyer's Point of View." *Econometrica* 55:633–646.

Möbius, M., M. Niederle, P. Niehaus, and T. Rosenblat. 2011. "Managing Self-Confidence: Theory and Experimental Evidence." *Working Paper* .

Moffitt, R. 1983. "An Economic Model of Welfare Stigma." *American Economic Review* 73 (5):1023–1035.

Myerson, R. 1981. "Optimal Auction Design." *Mathematics of Operations Research* 6:58–73.

Neale, M.A. and M.H. Bazerman. 1985. "The Effects of Framing and Negotiator Overconfidence on Bargaining Behaviors and Outcomes." *Academy of Management Journal* 28(1):34–49.

Niederle, M. and L. Vesterlund. 2007. "Do Women Shy Away from Competition? Do Men Compete too Much?" *Quarterly Journal of Economics* 122(3):1067–1101.

Noussair, C. and J. Silver. 2006. "Behavior in All-Pay Auctions with Incomplete Information." *Games and Economic Behavior* 55:189–206.

Odean, T. 1998. "Volume, Volatility, Price, and Profit when All Traders are Above Average." *Journal of Finance* 53(6):1887–1934.

Rege, M. and K. Telle. 2004. "The Impact of Social Approval and Framing on Cooperation in Public Good Situations." *Journal of Public Economics* 88(7-8):1625–1644.

Reuben, E., P. Rey-Biel, P. Sapienza, and L. Zingales. 2010. "The Emergence of Male Leadership in Competitive Environments." *IZA Discussion Paper* 5300.

Riley, G. and W. F. Samuelson. 1981. "Optimal Auctions." *American Economic Review* 71:381–392.

Shunda, N. 2010. "Auctions with a Buy Price: The Case of Reference-Dependent Preferences." *Journal of Games and Economic Behavior* 67(2):645–664.

Siegel, R. 2010. "Asymmetric Contests with Conditional Investments." *American Economic Review* 100:2230–2260.

Svenson, O. 1981. "Are We All Less Risky and More Skillful Than Our Fellow Drivers?" *Acta Psychologica* 47:143–148.

Taylor, S. E. and J. D. Brown. 1988. "Illusion and Well-being: A Social Psychological Perspective on Mental Health." *Psychological Bulletin* 103:193–210.

Thaler, R. and C. Sunstein. 2003. "Libertarian Paternalism." *American Economic Review* 93:175–179.

———. 2008. *Nudge: Improving Decisions About Health, Wealth, and Happiness.* Yale University Press.

Weinstein, N. D. 1980. "Unrealistic Optimism about Future Life Events." *Journal of Personality and Social Psychology* 39:806–820.

Wozniak, D., W.T. Harbaugh, and U. Mayr. 2010. "The Menstrual Cycle and Performance Feedback Alter Gender Differences in Competitive Choices." *Working Paper* .

Yaari, M. 1987. "The Dual Theory of Choice under Risk." *Econometrica* 55:95–115.