DISCUSSION PAPER SERIES

IZA DP No. 6425

Image and Misreporting

Mara Ewers
Florian Zimmermann

March 2012

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Image and Misreporting

Mara Ewers
University of Bonn
and IZA

Florian Zimmermann
University of Bonn
and IZA

Discussion Paper No. 6425
March 2012

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 6425
March 2012

ABSTRACT

Image and Misreporting*

In this paper we ask if reports of private information about skills, abilities or achievements are affected by image concerns. We develop 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 lab experiment. In the experiment, all subjects go through a series of quiz questions and subsequently report a performance measure. 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.

JEL Classification: C91, D03, D82

Keywords: image concerns, self-assessment, overconfidence, experiment

Corresponding author:

Florian Zimmermann
University of Bonn
Adenauerallee 24
53113 Bonn
Germany
E-mail: florian.zimmermann@uni-bonn.de

* We thank Steffen Altmann, Steve Burks, Gary Charness, Lorenz Götte, Matthias Kräkel, Sebastian Kube, Ernesto Reuben, Paul Schempp, Joel Sobel, Matthias Wibral, and especially Armin Falk for insightful comments and discussions. Valuable comments were also received from participants of the Bonn "retreat", the 7th International Meeting on Experimental and Behavioral Economics in Barcelona (2011), and the International ESA conference in Chicago (2011). Financial support from the Bonn Graduate School of Economics (BGSE) is gratefully acknowledged.

1 Introduction

Individuals hold private beliefs about their performance, skills, abilities and achievements. Transmission of this private information is crucial for the efficiency of 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 insurance contexts, the design of efficient insurance plans is difficult when individuals hold private information about their underlying risk. In this paper, we analyze whether individuals' image concerns can lead them to misreport private information in situations, where from a traditional, purely pecuniary perspective, truthful revelation would be optimal. Individuals that 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 illustrate with a simple model how image concerns make people misreport their own performance, skill or ability. In equilibrium, some individuals with low performance will choose to report high performance. Consequently, reports become less informative. Then we provide evidence from a lab experiment documenting 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 direct monetary concerns and an image part, reflecting reputational concerns. The way we model image concerns is a shortcut that captures all potential 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. Apart from sending positive signals about skills and abilities, our model also captures situations where decision makers might want to appear humble or modest in front of others. If modesty is the dominant signaling motive, misreporting might go in

the opposite direction, i.e., decision makers downplay own skills and display underconfidence. 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 has 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. This procedure has been used by Ariely, Bracha, and Meier (2009) to increase image concerns in the context of prosocial decision making. In the private treatment, subjects do not report their assessment to the group. Our data reveals 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. This difference seems to be driven by a stronger response of men to the presence of an audience.

To further assess if subjects' reports are also affected by a desire to appear modest in front of others, we conduct a feedback treatment. The treatment is identical to our audience treatment. The only difference is that after subjects report their self-assessments to the group, the experimenter will also report the true performances to the group. If misreporting in the audience treatment was solely driven by the signaling of skills or ability, we should observe that reported self-assessments do not differ between the private and the feedback treatment, because in the latter, true performances will be revealed. If concerns to appear modest are relevant, we should observe stated self-assessments in the feedback treatment below the level we found in the private treatment. When comparing stated self-assessments between the feedback and private treatment, we find no evidence for modesty concerns on the aggregate level. However, we do find some evidence that subjects with rather low quiz performance want to appear modest in the feedback treatment.

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.[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. Thus, in our approach, overconfident behavior is rather the outcome of a preference, e.g. a desire to signal skills or ability, than a mistaken self-perception. This might explain why people do not "learn" about their overconfidence over time.

Our findings are also 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. Likewise, our findings are 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 provide a solution to this problem.

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. 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

[1] Other empirical studies on overconfidence include for example Camerer and Lovallo (1999) and Hoelzl and Rustichini (2005). For a recent overview, see Benoît and Dubra (forthcoming). 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 or DellaVigna and Malmendier (2006) for overestimation of future gym attendance.

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 stigmatization of welfare recipients.

This paper relates to a few recent papers that considered the social signaling component of biases in self-assessments. Burks, Carpenter, Goette, and Rustichini (2010) compare different explanations for 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, Rey-Biel, Sapienza, and Zingales (forthcoming) find that subjects exaggerate past performance in order to become a group leader.

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 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

[2]Eil and Rao (2011) and Möbius, Niederle, Niehaus, and Rosenblat (2011) also provide evidence for a consumption value from overconfidence. Bénabou and Tirole (2002) provide a theoretical argument for a value of overconfidence as it can increase motivation of individuals with imperfect willpower. Other models have assumed a value of self-confidence and show how overconfident self-assessments can be produced by selectively choosing information or by asymmetrically processing information (see Brunnermeier and Parker (2005), Kőszegi (2006) or Möbius, Niederle, Niehaus, and Rosenblat (2011)).

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 3 presents the experimental design, section 4 the results from our experiment and section 5 concludes.

2 The Model

We provide a simple framework that allows illustrating how image concerns can influence reports of private information. The next two sections introduce the model, assuming that decision makers have perfect knowledge about their performance. In section 2.3 we relax the assumption of perfect knowledge. In section 2.4 we show how a desire to appear modest could be captured with our model framework.

2.1 Set-Up

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 section 2.3 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 non-monetary 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 imperfect recall. Since actions are easier to recall than signals, decision makers base their self-evaluation on past stated reports.

[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%. Note that throughout the paper we focus on direct monetary utility, but always mean to include non-monetary interpretations such as costs of lying.

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

2.2 Equilibrium

We now show under which conditions there exists a unique Perfect Bayesian Equilibrium where decision-makers misreport their private information. In the absence of image concerns, D's behavior would be straightforward. Decision makers choose $x = Yes$, if their performance, skill or ability p is higher than $\overline{r}$ and $x = No$ otherwise. In the presence of image concerns however, there exists a trade-off between stating a truthful report and gaining social approval. In equilibrium, all decision makers with $p > \overline{r}$ will choose $x = Yes$. For decision makers with $p < \overline{r}$ there exists a threshold type α^* such that all D with $\alpha > \alpha^*$ will choose $x = Yes$ and all D with $\alpha < \alpha^*$ will choose $x = No$. This is stated formally in the following Proposition:

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* $x = No$.

Next, we verify that behavior described above is indeed an equilibrium and show that if $\overline{\alpha}$ (the highest possible realization of α) is sufficiently large, such an equilibrium always exists. In Appendix A we show that this equilibrium is unique.

First, we state precisely what we mean by $\overline{\alpha}$ being sufficiently large. We assume that there exist decision makers with image concerns large enough, such that they choose $x = Yes$ if $p < \overline{r}$ and all other decision makers simply maximize monetary outcomes. More precisely, $\overline{\alpha}$ is large enough, such that the 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.

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:

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

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 leads the following:

$$\alpha^*\beta\left[\frac{1}{\sum_{p>\overline{r}} f(p) + \int_{\alpha^*}^{\overline{\alpha}} g(z)dz \sum_{p\leq\overline{r}} f(p)} \left(\sum_{p>\overline{r}} f(p)p + \int_{\alpha^*}^{\overline{\alpha}} g(z)dz \sum_{p\leq\overline{r}} f(p)p\right) - \sum_{p\leq\overline{r}} \frac{f(p)p}{\sum_{p\leq\overline{r}} f(p)}\right] = y. \tag{2}$$

One can see from equation (2) that decision makers with $\alpha < \alpha^*$ and $p < \overline{r}$ optimally choose $x = No$. As the expression in square brackets (gains in reputation) remains unchanged, but the strength of image concerns is smaller ($\alpha\beta < \alpha^*\beta$), image gains in total weigh less in utility terms 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 (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 (2) holds.

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. Considering equation (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 $x = Yes$. Thus, our model predicts that an exogenous increase in image concerns increases the number of decision makers that misreport information. Consequently, reports become less informative.

2.3 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 $\hat{p} \in \{0, 1, ..., \overline{p}\}$ and $\hat{p}$ is (potentially) different from p.[5] D's choice x is again to report whether p is larger than some value $\overline{r}$, 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 $\overline{r}$.

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

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

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 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.[6] In Appendix A we state formally the requirement that $\overline{\alpha}$ is sufficiently large. 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.

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

2.4 Modesty

In addition to signaling skill or ability, there might exist other signaling motives. Decision makers may want to appear humble or modest in front of others, i.e. they downplay own skills or achievements when reporting them to others. In our model, there is a simple way to capture parts of this additional signaling motive. We model modesty in the context of self-assessments as a reluctant

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

view about oneself. This can be signaled towards others by reporting low self-assessments.

We assume that decision makers hold a point belief $\hat{p} \in \{0, 1, ..., \bar{p}\}$ and $\hat{p}$ is potentially different from p. $\hat{p}$ is D's private knowledge, but is commonly known to be distributed according to a probability function g defined over P. We now consider mistaken point beliefs in more detail. First, take decision makers with $\hat{p} < p$. Instead of being interpreted as a simple consequence of receiving imperfect signals on p, downward biases in self-evaluation are now being interpreted as a character trait, namely having a modest and reluctant view about oneself. By the same argument, biases upwards ($\hat{p} > p$) could capture traits such as overconfidence. We now want to allow for the signaling of modesty or (not) signaling overconfidence. Utility is given by

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

The fact that D's image or approval stems from the public's belief about $\hat{p}$ captures the additional signaling motive. Depending on the strength of the correlation between $\hat{p}$ and actual performance p, image concerns still reflect the desire to appear skilled and able to the public. In addition, now decision-makers might also want to signal modesty or not being overconfident. The sign of common image utility β determines which signaling-motives dominate. Positive values of β capture dominance of the desire to signal skill and ability. Negative values of β might reflect situations where social approval stems from modesty.

The structure of this game is identical to that in the model with perfect knowledge. Consequently, we can state the following proposition.

PROPOSITION 5: *If $\overline{\alpha}$ is sufficiently large, i.e. $\mid \overline{\alpha}\beta \left[\sum_{\hat{p}>\overline{r}} \frac{g(\hat{p})\hat{p}}{\sum_{\hat{p}>\overline{r}} g(\hat{p})} - \sum_{\hat{p}\leq\overline{r}} \frac{g(\hat{p})\hat{p}}{\sum_{\hat{p}\leq\overline{r}} g(\hat{p})}\right] \mid > y$, there exists a unique Perfect Bayesian Equilibrium. If $\beta > 0$, 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 $x = No$. If $\beta < 0$, decision makers with $\hat{p} > \overline{r}$ and $\alpha > \alpha^*$ choose $x = No$. Decision makers with $\hat{p} < \overline{r}$ choose $x = No$ and those with $\hat{p} > \overline{r}$ and $\alpha < \alpha^*$ choose $x = Yes$.*

Proposition 5 shows how image concerns for skill, ability or modesty affect stated self-assessments. The sign of common image utility β determines which signaling motive dominates. If $\beta < 0$, decision-

makers in equilibrium underreport their private information in order to appear humble or modest. If $\beta > 0$, decision makers overreport due to a desire to signal ability or skill. Which of the motives dominates in reality is ultimately an empirical question and our experiment can be viewed as an attempt to answer it.

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 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.[7] 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 such as 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.[8] We asked them to assess whether their own performance was better or worse than the average quiz-performance of a group of other participants.[9] The group of other participants consisted of 95 different subjects who had also performed the quiz. Subjects received 5 euros for a correct self-assessment. The only difference between our two treatments was the following: In the

[7]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.

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

[9]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.

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 their 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 and report their self-assessment to the group.[10] 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.

The experiment ended with a questionnaire. We elicited subjects' risk preferences, image concerns and several socio-demographic characteristics like gender and age. We measured subjects' risk preference by using a question from the German Socio-Economic Panel Study (GSOEP): "How do you see yourself: Are you generally a person who is fully prepared to take risks or do you try to avoid taking risks? Please tick a box on the scale, where the value 0 means: 'completely unwilling to take risks' and the value 10 means: 'completely willing to take risks'." Dohmen, Falk, Huffman, Sunde, Schupp, and Wagner (2009) show that this survey question is very well suited for analyzing risk preferences, because it is highly correlated with incentivized lottery choice measures. In addition, we used the question "How important is the opinion of others to you?" as a survey-based measure of image concerns. Subjects could choose between five answers from 'not at all' to 'very important'. To gain further insights on whether a high or low stated self-assessment is associated with social (dis)approval, we asked two questions. 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'. In addition, we wanted to learn more about subjects' perception of others' self-assessments. We asked: "When stating their self-assessments, do you think the other participants overestimated, underestimated or correctly estimated their performance?"

The version of our model we presented in section 2.4 suggests that subjects might signal more than just ability through their choice. The stated self-assessment can also be informative about traits

[10] 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.

like modesty. While we suspected that the desire to signal skills or ability would dominate, we still wanted to analyze if concerns to appear modest affected stated self-assessments as well. Therefore, we ran a *feedback treatment* where we controlled 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, the experimenter informed the public whether the assessment was correct or not. In this situation, the public learns the true relative performance and therefore subjects can no longer signal ability. However, they could signal modesty by reporting to be worse than average. Thus, if misreporting in the audience treatment is solely driven by the signaling of skills or ability, we should observe that reported self-assessments should not differ between the private and the feedback treatment. However, if concerns to appear modest are relevant, stated self-assessments in the feedback treatment should be below the level of the private treatment.

Table 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? - 5euros / correct self-assessment	1. Private: no further action 2. Audience: reporting self-assessment to an audience 3. Feedback: public feedback after reported self-assessment	- Risk - Survey questions - Demographics

3.1 Experimental Procedures

A total of 143 subjects participated in the experiment, 47 in the private treatment, 48 in the audience treatment, and 48 in the feedback treatment. We were interested in potential gender differences and therefore invited an equal amount of women and men to each session. All six 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[11] on average. We distributed the instructions for stage one and two immediately before the stage started and they were read aloud.[12]

3.2 Hypotheses

In the experiment, we systematically increase image concerns of subjects by introducing an audience. When comparing the private and the audience treatment, we hypothesize that signaling ability is the dominant signaling motive. Thus, by Propositions 2 and 4 of our model, reported self-assessments should be higher in the audience treatment compared to the private treatment.

HYPOTHESIS 1: *Subjects choose "better than average" more frequently in the audience treatment than in the private treatment.*

Our feedback treatment allows us to control for the signaling of ability. Thus, if a desire for appearing modest is present, stated self-assessments should be lower in the feedback treatment compared to the private treatment.

HYPOTHESIS 2: *Subjects choose "better than average" less frequently in the feedback treatment than in the private treatment.*

4 Results

We start with the analysis of our two main treatments. In section 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 analyze individuals' perceptions of others' stated self-assessments. In section 4.2, we present results from the feedback treatment.

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 treat-*

[11] 1 euro was worth about 1.4 Dollars at the time.
[12] Instructions are available from the authors upon request.

ment.

Table 2: Determinants of stated self-assessment

Dependent variable: Relative self-assessment= $\begin{cases} 1 & \text{if better} \\ 0 & \text{if worse} \end{cases}$					
	(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 the survey based risk measure, image concerns, age, and relationship status.

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 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".[13] Column 1 of Table 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, gender and several additional controls. In column 2, we take the number of correctly solved quiz

[13]Columns 4 and 5 of Table 2 are discussed later.

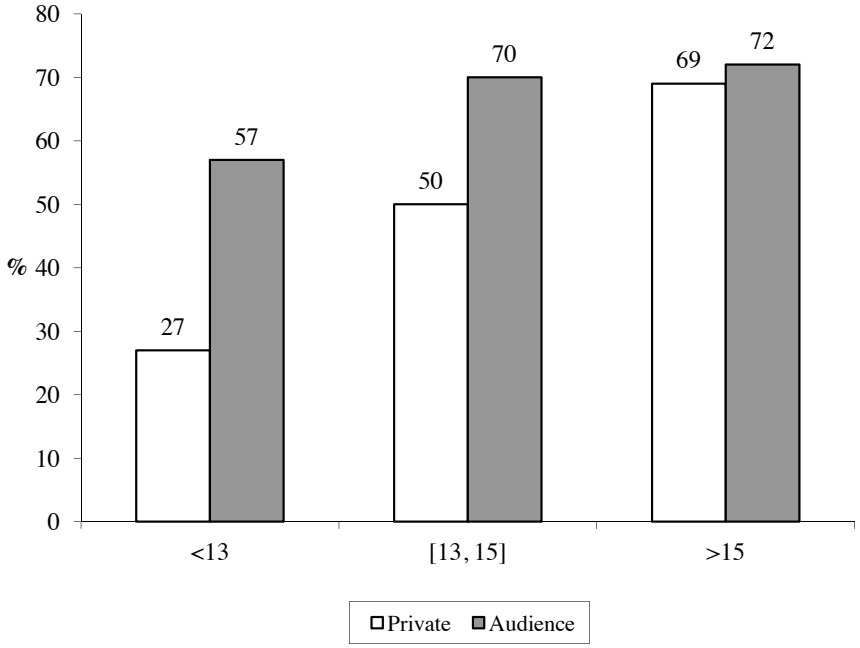
questions as a control for quiz performance. In column 3, we use a different measure: we create a performance dummy, taking the value one if performance was actually better than average and zero otherwise. In both regressions, the treatment effect remains significant at the 5% level.[14] In addition, 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 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. In line with section 2.3 of our model, we assume that 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 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. First, 64% of our subjects stated that they enjoyed the quiz or enjoyed it very much. Only 10% indicated they did not like the quiz. Second, when subjects were asked to imagine they owned a firm and had the opportunity to hire new workers, none of the

[14]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.

Figure 1: Percentage of "better than average" reports for high, low and close to average quiz performance, for subjects in the audience treatment and the private treatment.



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.

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 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 found in many studies and provides a possible explanation for documented gender differences 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 2 show separate Probit regressions for the private and the audience treatment. The data indicates that the gender effect is mostly driven by more men overreporting in the audience treatment. While men report to be "better than average" significantly

Table 3: Subjects' beliefs about the other participants' self-assessments

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

more often than women in the audience treatment, the effect is not significant in the private treatment. This finding might provide a possible explanation for gender differences in overconfidence. It suggests, that men feel a stronger desire to signal skills or abilities towards others which results in overconfident appearance.[15]

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 in the experiment. Table 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.[16]

[15]Note, however, that this interpretation should be taken with caution. In Appendix 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.

[16]Ludwig and Nafziger (2010) 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.

4.2 Feedback Treatment

Result 4 *Comparing the feedback treatment to the private treatment, we do not find evidence for strong modesty concerns on the aggregate level.*

We now analyze results from our feedback treatment. The purpose of the treatment was to identify whether subjects' reports are also affected by concerns for appearing modest in front of others. 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 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. When asking our subjects about their perception of other subjects' reports, 81% indicated that they think others chose a correct self-assessment. Thus, on an aggregate level, we do not find evidence for strong modesty concerns.

Note, however, that we do find evidence for modesty concerns for subjects with low quiz performance. Figure 2 depicts the percentages of subjects in the feedback and the private treatment who report to be better than average for the same intervals of actual quiz performance as in section 4.1 (Figure 1). Among subjects with rather low quiz performance (less than 13 correctly solved questions) 27 % report to be better than average in the private treatment, while only 10 % do so in the feedback treatment. However, this desire to appear modest seems to counteract with a desire to appear confident in own skills or performance for high performance subjects. Similar to a desire to appear modest in front of others, some subjects might want to display confidence in their own performance. Figure 2 shows that for subjects with high quiz performance (more than 15 correctly solved questions), the treatment effect goes in the opposite direction. 69 % report to be better than average in the private treatment, compared to 93 % in the feedback treatment. Thus, it seems that the feedback treatment has a differential impact on stated self-assessments, depending on actual quiz performance. Knowing that the experimenter will subsequently provide feedback to the audience, low performance subjects want to appear modest, while high performance subjects want to signal confidence in their own performance.

Table 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

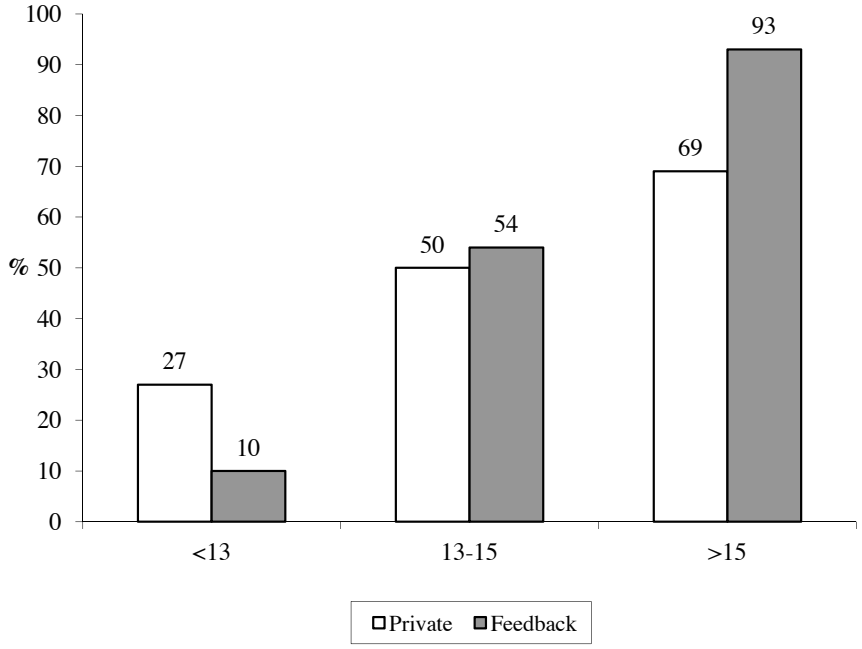
Notes: Probit estimates. Marginal effects (evaluated at the mean of independent variables) reported; robust standard errors are in parentheses. Significance at the 1, 5, and 10 percent level is denoted by ***, **, and *, respectively. *Dummy treatment* =1 if feedback treatment and 0 if private treatment. *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.

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

Figure 2: Percentage of "better than average" reports for high, low and close to average quiz performance subjects in feedback treatment and private treatment.



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. 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, Niederle, Niehaus, and Rosenblat (2011)). Recently, several experimental papers have provided support for biases in information-processing and information demand (see Eil and Rao (2011), Möbius, Niederle, Niehaus, and Rosenblat (2011) and Charness, Rustichini, and van de Ven (2011)). Benoît and Dubra (forthcoming) 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.

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. We could have implemented an instrumental value of appearing skilled or able 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.

Finally, 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)). 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, Niederle, Niehaus, and Rosenblat (2011)) but from self-evaluation based on biased choices.

References

Andreoni, J., and B. D. 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(7-8), 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.

Bem, D. (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 (forthcoming): "Apparent Overconfidence?," *Econometrica*.

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

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

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*.

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.

Dohmen, T., A. Falk, D. Huffman, U. Sunde, J. Schupp, and G. G. Wagner (2009): "Individual Risk Attitudes: Measurement, Determinants and Behavioral Consequences," *Working Paper*.

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.

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

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

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 August," *Quarterly Journal of Economics*, pp. 1049–1074.

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

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

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

Ludwig, S., and J. Nafziger (2010): "Beliefs about Overconfidence," *Theory and Decision*.

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

Möbius, M. M., M. Niederle, P. Niehaus, and T. S. 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.

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.

Rege, and 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 (forthcoming): "The Emergence of Male Leadership in Competitive Environments," *Journal of Economic Behavior and Organization*.

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

Appendix A

Uniqueness of equilibrium described in Proposition 1: 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 (2) we see that α^* and consequently the equilibrium described above is unique. We summarize our results in the following proposition.

Clarification Proposition 3: We assume that there exist decision makers with image concerns large enough such that they would choose $x = Yes$ if $\hat{p} < \overline{r}$ and if all other decision makers would simply maximize monetary outcomes. More precisely, $\overline{\alpha}$ is 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. Note that this condition is more demanding than that in the case of perfect information. The reason 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.

Appendix B

Table 5: Determinants with interactions of relative self-assessment in the private and audience treatment

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.