DISCUSSION PAPER

NO 356

# Loss Aversion in Social Image Concerns


Vasilisa Petrishcheva
Gerhard Riener
Hannah Schildberg-Hörisch


November 2020

## IMPRINT

## DICE DISCUSSION PAPER

**Published by:**
Heinrich-Heine-University Düsseldorf,
Düsseldorf Institute for Competition Economics (DICE),
Universitätsstraße 1, 40225 Düsseldorf, Germany
www.dice.hhu.de

**Editor:**
Prof. Dr. Hans-Theo Normann
Düsseldorf Institute for Competition Economics (DICE)
Tel +49 (0) 211-81-15125, E-Mail normann@dice.hhu.de



ISSN 2190-9938 (online) / ISBN 978-3-86304-355-1

# Loss Aversion in Social Image Concerns

**Vasilisa Petrishcheva**
University of Düsseldorf, Düsseldorf Institute for Competition Economics (DICE), Universitätsstr. 1, 40225 Düsseldorf, Germany

**Gerhard Riener**
University of Düsseldorf, Düsseldorf Institute for Competition Economics (DICE), Universitätsstr. 1, 40225 Düsseldorf, Germany

**Hannah Schildberg-Hörisch**
University of Düsseldorf, Düsseldorf Institute for Competition Economics (DICE), Universitätsstr. 1, 40225 Düsseldorf, Germany; IZA, Schaumburg-Lippe-Strasse 5-9, 53113 Bonn, Germany

November 2020

**Abstract**
This paper explores whether loss aversion applies to social image concerns. In a simple model, we combine loss aversion in social image concerns and attitudes towards lying. We then test its predictions in a laboratory experiment. Subjects are first ranked publicly in a social image relevant domain, intelligence. This initial rank serves as within-subject reference point. After inducing an exogenous change in subjects' rank across treatments, subjects are offered scope for lying to improve their final rank. We find evidence for loss aversion in social image concerns. Subjects who face a loss in social image lie more than those experiencing gains if they sufficiently care about social image and have a reputation to lose. Individual-level analyses document a discontinuity in lying behavior when moving from rank losses to gains, indicating a kink in the value function for social image. (JEL: C91; D91)

Keywords: Loss aversion; Social image concerns; Lying behavior; Laboratory experiment.

Acknowledgments: We are grateful to Johannes Abeler, Loukas Balafoutas, Jonas Fooken, Lorenz Götte, Paul Heidhues, Mats Köster, Hans-Theo Normann, Lionel Page, Wojtek Przepiorka, Marta Serra-Garcia, Sigrid Suetens, Agnieszka Tymula, Joël van der Weele, Jana Willrodt, and Florian Zimmermann for helpful comments and suggestions. We also thank Robin Bitter and Vladyslav Zhurakhinskyy for excellent research assistance. Financial support from DFG through the grant for GRK 1974 Competition Economics is gratefully acknowledged. AEA RCT Registry: https://www.socialscienceregistry.org/trials/3422

E-mail: petrishcheva@dice.hhu.de. (Petrishcheva); riener@dice.hhu.de (Riener); schildberg-hoerisch@dice.hhu.de (Schildberg-Hörisch)

## 1. Introduction

Humans care how they are perceived by their fellow humans and go a great length to build up a positive image of themselves (e.g., Bénabou and Tirole (2006); Bursztyn and Jensen (2017); Andreoni and Bernheim (2009); Ariely et al. (2009); Soetevent (2011); Ewers and Zimmermann (2015)). These carefully crafted images are at stake in everyday interaction, and reputation can decline rapidly. Casual observations suggest that when social image is at risk of being lost people engage in lies and denial to maintain it in many domains of economic life. Managers who do not reach expected targets may engage in fraudulent behavior—as happened recently in the manipulation of car emission tests (Aurand et al., 2018). A person losing her job may still leave the house everyday pretending to her family that she still is employed. However, the reference point for status loss does not necessarily have to come from own achievements or calamities, it may also be transmitted through generations as a sense of class entitlement (Alsop, 2008). In the 2019 college admission scandal, affluent parents criminally conspired to influence admission decisions of prestigious colleges (Halleck, 2019; Lovett, 2020). While the special role of losses has been extensively documented in the monetary domain (Kahneman and Tversky, 1979; Camerer, 1998; Wakker, 2010; Barberis, 2013), the causal effect of losses on moral behavior deserves a closer look.

Does trying to shield oneself from a loss in social image generally lead to more morally deviant behavior than striving for a gain in social image? Or is it a particular behavior of those people who are more inclined to immoral decisions that can lead to tragic fall in the first place? Exogenous variation in the loss of social image is hard to imagine in the field and the extent of lying difficult to observe. Hence, we design a parsimonious laboratory experiment to test for the presence of loss aversion in social image concerns.

To fix ideas, we develop a simple model combining loss aversion in social image concerns and attitudes towards lying to derive testable hypotheses. In the experiment, we induce exogenous variation in reference points across treatments such that subjects either experience a potential loss or gain in their social image, while keeping average social image constant across treatments. We then offer subjects scope for improving their social image by lying about their true type. This allows us to test whether—on average—subjects lie more (and are thus willing to incur higher lying costs) when they experience losses than when they experience gains in their social image.

Our results provide evidence for loss aversion in social image concerns. Comparing average behavior across treatments, we find that subjects who sufficiently care about their social image—as measured by an independent survey instrument—and those with high initial social image who have a reputation to lose behave in line with loss aversion in social image concerns. Further individual-level analyses document that, on average, the extent of lying decreases discontinuously when moving from small losses to small gains in social image. This pattern in lying behavior is compatible with loss aversion in social image concerns but not a simple concave utility function for social image.

Our main contribution is thus documenting loss aversion in social image concerns. Importantly, our findings imply that loss aversion can also play a role in the non-material domain of social image. So far, loss aversion is widely documented for money (e.g., Booij and Van de Kuilen (2009); Pennings and Smidts (2003)) and material goods (e.g., Kahneman et al. (1990))[1], but evidence on whether humans have the same inclination when it comes to social image utility is lacking.

Image concerns expand over various domains[2]: People care about being perceived smart and skillful (e.g., Ewers and Zimmermann (2015) and Burks et al. (2013)), prosocial and altruistic (e.g., Carpenter and Myers (2010)), pro-environmental (e.g., Sexton and Sexton (2014)) and supportive of fair trade (Friedrichsen and Engelmann (2018)), trustworthy (Abeler et al. (2019)), promise-keeping (Grubiak (2019)), or wealthy (Leibenstein (1950)).

In our experiment, we induce social image concerns by letting subjects perform an IQ test and reporting its results publicly. However, signaling skillfulness can be a two-sided sword as Austen-Smith and Fryer Jr (2005) show in a two-audience signaling model. For example, high ability students may under-invest in education because such investments lead to rejection by their peer group (Bursztyn et al. (2019)).[3] So it is important to establish that an IQ test is indeed suitable to induce social image that is worth striving for in our university student sample. This is underlined by Ewers and Zimmermann (2015) who document that, in a student sample similar to the one used in this study, subjects misreport their private information on ability in a laboratory context in order to appear more skillful even when strong monetary incentives are given to tell the truth.

While there is plenty of evidence that many people care about social image, recent, both theoretical and empirical work stresses that there is heterogeneity in the extent to which people care about social image and whether they do so at all. For example, Bursztyn and Jensen (2017) expand the model of Bénabou and Tirole (2006) to explicitly account for heterogeneity in social image concerns.[4] Friedrichsen and Engelmann (2018) empirically reject the hypothesis of homogeneous image concerns and show that individuals react

1. See Bleichrodt et al. (2001) for an application to health outcomes.

2. Bursztyn and Jensen (2017) present a detailed overview of the recent literature on social image concerns.

3. Bursztyn et al. (2019) show that students are less likely to sign up for an SAT preparation course and to take an SAT exam itself, if their choices are observable. They therefore forgo educational investment due to possible social stigma.

4. Their theoretical framework distinguishes conformists who experience social pressure to act in a socially desirable way, contrarians who feel pressured to act differently from what is socially desirable, and those who are not subject to social image concerns at all.

differently to image-building opportunities. In our experiment, we will therefore measure each subject's individual extent of social image concerns.

On top of addressing image concerns, this study also contributes to the growing literature on lying behavior, extensively summarized in Abeler et al. (2019).[5] Based on a comprehensive meta-analysis, Abeler et al. (2019) identify two main channels why people prefer to tell the truth, namely, lying costs that increase in the size of a lie and image concerns for being perceived as an honest person. Our theoretical framework and experiment design build on their work. First, our experiment design ensures that lying cannot be detected such that image concerns for being seen as an honest person by others cannot play a role in the context of our experiment. Second, in order to avoid possible interactions between loss aversion in the monetary and social image domain, our design offers subjects a flat payment and uses the extent of lying, i.e., the lying costs subjects are willing to incur, to quantify how much they suffer from losing or gain from improving their social image. Therefore, our finding that subjects who care about their social image and have a reputation to lose are more likely to report more dishonestly than others speaks to situations in which honest reporting of private information is key but not incentive-compatible. Since lying in the laboratory is a predictor of dishonesty and rule violations in real life (Hanna and Wang, 2017; Dai et al., 2018), our findings suggest that monitoring efforts should be targeted at individuals who have a high reputation and care strongly about it.

We also relate to the literature which links the concept of loss aversion (in the monetary domain) to lying behavior. Grolleau et al. (2016) and Schindler and Pfattheicher (2017) compare the extent of lying for individuals who face monetary losses and gains. They find that participants misreport more to avoid a monetary loss than they do to increase their monetary gain. Garbarino et al. (2019) show that the less likely a low monetary payoff is, the more likely individuals lie to avoid it.

The paper proceeds as follows: Section 2 introduces a theoretical framework combining social image concerns and loss aversion. Section 3 describes the experiment design and procedures, before we outline our hypotheses in Section 4. Results are presented in Section 5 and Section 6 concludes.

## 2. Theoretical Framework

Our model integrates three key psychological features that—up to now have been treated separately—into individual utility: (1) agents gain positive utility from social image, (2) agents experience loss aversion in the social image

5. Abeler et al. (2019) provide a web interface where they present recent experiments on lying in great detail.

domain, i.e., losses of social image loom larger than gains of the same size, and (3) agents dislike lying, i.e., they experience costs of misreporting the true state of the world.

Consider a two-period game with $t \in \{1, 2\}$. In the first period, an agent $i$ receives a signal of her type $s_{i1}$ that is communicated to herself and her peers. One can think of it as her social image relevant performance, and we assume that this signal establishes a reference point concerning her true type.

In the second period, she learns about her true type $\tilde{s}_{i2}$, while peers are only going to see a signal of the true type $s_{i2}$. This signal can be actively misrepresented in an unverifiable manner by the agent. In each period $t$, she derives $u(s_{it})$ from the signal of her social image, where $u(\cdot)$ is assumed to be differentiable and weakly increasing in $s_i$.

To model the cost of misrepresenting the true state we follow Abeler et al. (2019) and Khalmetski and Sliwka (2019). We denote the true state of the world as $\omega \in [-\overline{\omega}, \overline{\omega}]$ which is independently and identically distributed (iid) across individuals. The agent's report of the true state is $r \in [-\overline{r}, \overline{r}]$ with $\overline{\omega} = \overline{r} > 0$. In period $t = 2$, her final public signal $s_{i2}$ consists of her actual performance $\tilde{s}_{i2}$ plus her report of the true state $r_i$ ($s_{i2} = \tilde{s}_{i2} + r_i$). Agent $i$ dislikes misreporting the true state and experiences lying costs $c(\omega_i, r_i)$. Lying costs are zero if the state is reported truthfully, i.e., $c(\omega_i, \omega_i) = 0$, and positive otherwise. Lying costs depend on the size of misreporting and are symmetric around $\omega_i$, i.e., $c(\omega_i, \omega_i + a) = c(\omega_i, \omega_i - a)$ for all $a$.[6] In other words, as in Abeler et al. (2019), an agent experiences the same lying costs when misreporting in either direction to the same extent. In contrast to Abeler et al. (2019), we do not model the social image concerns of being seen as a liar as we explicitly rule them out in our experiment design.

The agent has to make a choice in period 2 only. We therefore limit our attention to the utility function in the second period:

$$\varphi_{i2} = \theta_i^{social}[u(\tilde{s}_{i2} + r_i) + v(\tilde{s}_{i2} + r_i - s_{i1})] - \theta_i^{lying} c(\omega_i, r_i),$$

which she maximizes with respect to her report $r_i$. $\theta_i^{social}$ represents the sensitivity to social image that may differ across individuals (Bursztyn and Jensen, 2017). $\theta_i^{lying}$ represents the individual's sensitivity to lying (Gibson et al., 2013). $s_{i1}$ and $\tilde{s}_{i2}$ are parameters[7], hence utility of period 1 is fixed

6. Note that in our model, agents follow teleological moral theory that can be seen as a form of act consequentialism. In contrast, agents who adhere to a deontological normative moral reasoning would never engage in lying as it is considered a moral wrong, independent of the cost structure and the other parameters of the model.

7. In laboratory experiments, subjects tend to exert close to maximal effort in real-effort tasks in general (Araujo et al., 2016; Corgnet et al., 2015; Gächter et al., 2016; Goerg et al., 2019) and in IQ tests like the Raven's Progressive Matrices that we will use specifically (Eckartz et al., 2012). Hence, we assume $s_{i1}$ and $\tilde{s}_{i2}$ to be parameters, not variables.

($\varphi_{i1} = \theta_i^{social} u(s_{i1})$), and we just consider the utility function in period 2 for maximization. We assume the following differentiable value function:

$$v(s_i): \quad v(\Delta_i) < -v(-\Delta_i),$$

where $\Delta_i$ is the difference between the true type and the first signal ($\tilde{s}_{i2} - s_{i1}$). The value function satisfies the standard assumptions of prospect theory (Kahneman and Tversky, 1979). Negative deviations from the reference points $s_{i1}$ have a larger absolute impact on utility than equally sized positive deviations, i.e., $v'(\Delta_i) < v'(-\Delta_i)$. Additionally, the value function is concave for gains ($v''(\Delta_i) < 0$ for $\Delta_i > 0$) and convex for losses ($v''(\Delta_i) > 0$ for $\Delta_i < 0$). The first result follows directly:

Proposition 1. *Individuals without social image concerns never misreport the true state.*

*Proof.* If $\theta_i^{social} = 0$, agent's utility in period 2 is reduced to

$$\varphi_{i2} = -\theta_i^{lying} c(\omega_i, r_i),$$

which reaches its maximum when lying costs are minimized, i.e., in the absence of lying. Hence, an agent who does not care about her social image will always report truthfully: $r_i = \omega_i$. □

The utility derived from the social image is weakly increasing when the agent's report increases ($\partial u(\tilde{s}_{i2} + r_i)/\partial r_i \geq 0$) because the agent obtains a non-negative marginal utility when the signal gets better. $\partial v(\tilde{s}_{i2} + r_i - s_{i1})/\partial r_i > 0$ is independent of whether an individual is in the loss or gain domain (or shifts from the loss to the gain domain). Lying costs are positive whenever the true state is misreported and $\partial c(\omega_i, r_i)/\partial r_i > 0$ if $\omega_i < r_i$ and $\partial c(\omega_i, r_i)/\partial r_i < 0$ if $\omega_i > r_i$. The next result is straightforward:

Proposition 2. *Individuals never under-report the true state.*

*Proof.*
Given the true state $\omega_i$, an agent always strictly prefers to report $r_i = \omega_i$ to any $\tilde{\omega}_i$ such that $\tilde{\omega}_i < \omega_i$ because under-reporting lowers utility due to three factors. First, an individual obtains weakly lower utility derived from the social image: $u(\tilde{s}_{i2} + \tilde{\omega}_i) \leq u(\tilde{s}_{i2} + \omega_i)$. Second, the level of value function is lower at $\tilde{\omega}_i$ than at $\omega_i$ for any value of $\Delta_i$, i.e., $v(\Delta_i + \tilde{\omega}_i) < v(\Delta_i + \omega_i)$, because $\partial v(\tilde{s}_{i2} + r_i - s_{i1})/\partial r_i > 0$. Third, reporting $r_i = \omega_i$ yields zero lying costs while reporting $\tilde{\omega}_i$ misreports the true state, which is costly, i.e., $c(\omega_i, \tilde{\omega}_i) > c(\omega_i, \omega_i)$. □

Additionally, if $\omega_i = \overline{\omega}$ and an agent does not under-report, it directly follows that the agent will report truthfully (i.e., $r_i = \omega_i$) which leads to Lemma 1.

Lemma 1. *Individuals always report truthfully if* $\omega_i = \overline{\omega}$.

In the following, we assume that $\theta_i^{social} \neq 0$. We then denote $\theta_i \equiv \theta_i^{lying}/\theta_i^{social}$ which expresses individual's relative sensitivity to lying.

*Partial and full lying.* We study the conditions under which an agent engages in full and partial lying. We refer to "full lying" whenever an agent reports $r_i = \overline{\omega} > \omega_i$ and to "partial lying" whenever an agent reports $\omega_i < r_i < \overline{\omega}$. Lying costs are non-differentiable at zero with their first order derivative being strictly positive from the right and strictly negative from the left.[8] We believe it is plausible to assume lying costs which are non-differentiable at zero: Otherwise, all individuals, even those with extremely high relative lying sensitivity, would necessarily engage in over-reporting.[9]

Proposition 3. *Individuals report truthfully if* $\theta_i \geq \theta_i^{true}$ *or* $\omega_i = \overline{\omega}$.

*Proof.*

Non-differentiability of lying costs at zero implies that individuals with high sensitivity to lying will report truthfully if:

$$\theta_i \geq \theta_i^{true} = \left( \frac{\partial c(\omega_i, r_i)}{\partial r_i} \Bigg|_{r_i=\omega_i}^{+} \right)^{-1} \left( \frac{\partial u(\tilde{s}_{i2}+r_i)}{\partial r_i} \Bigg|_{r_i=\omega_i} + \frac{\partial v(\tilde{s}_{i2}+r_i-s_{i1})}{\partial r_i} \Bigg|_{r_i=\omega_i} \right). \qquad \square$$

A relative lying sensitivity $\theta_i$ lower than $\theta_i^{true}$ increases misreporting. Which factors determine the extent of lying?

*Social image sensitivity* $\theta_i^{social}$*:* If the agent has very strong image concerns, i.e., $\theta_i^{social}$ is very high, she might misreport the true state up to $\overline{\omega}$. If, on the contrary, the agent does not value social image so highly, she might only lie partially. If the valuation of her social image is particularly low, or absent, she will not engage in lying at all.

*Lying sensitivity* $\theta_i^{lying}$*:* If the agent is very insensitive to lying, she might engage in full lying. However, if her sensitivity parameter is relatively high (but not that high to report truthfully), she chooses to lie partially.

*True state* $\omega_i$*:* If the true state $\omega_i$ is small enough, the difference between the true state $\omega_i$ and the best state $\overline{\omega}$ is large. A large difference offers a lot of scope for lying but also means that lying costs may potentially get very high. Therefore, partial lying is more likely in the bad true states. If, on the contrary, the true state is very good, the lying costs to reach $\overline{\omega}$ are quite small, so lying to the full extent is more likely.

*Curvature of lying cost function* $\partial^2 c(\omega_i, r_i)/\partial^2 r_i$*:* If marginal costs of lying increase steeply, the agent is more likely engage to in partial lying. If,

8. $\frac{\partial c(\omega_i, r_i)}{\partial r_i}\Big|_{r_i=\omega_i}^{+} > 0$ and $\frac{\partial c(\omega_i, r_i)}{\partial r_i}\Big|_{r_i=\omega_i}^{-} < 0$, respectively.

9. A detailed proof is provided in Appendix A.

on the contrary, $\partial c(\omega_i, r_i)/\partial r_i$ is increasing in $r_i$ rather slowly, individuals are more likely to choose to lie all the way up to $\overline{\omega}$.

*Marginal utility from social image:* If the agent cares a lot about social image, and every additional score point brings her a lot of utility, she is more likely to engage in misreporting all the way up to $\overline{\omega}$.

*Marginal value of misreporting:* If increasing the gain, reducing the loss or shifting from loss to gain in the social image domain has a higher marginal value, the agent has more incentive to misreport.

PROPOSITION 4. *Individuals lie fully if $\theta_i \leq \theta_i^{full}$ and $\omega_i < \overline{\omega}$.*

*Proof.*

The agent chooses to lie fully if reporting $r_i = \overline{r} = \overline{\omega}$ yields marginal costs ($MC$) that are the same or lower than the marginal benefits ($MB$) of lying:

$$\underbrace{\frac{\partial v(\tilde{s}_{i2} + \overline{\omega} - s_{i1})}{\partial r_i} + \frac{\partial u(\tilde{s}_{i2} + \overline{\omega})}{\partial r_i}}_{MB\ of\ lying} \geq \underbrace{\theta_i \frac{\partial c(\omega_i, \overline{\omega})}{\partial r_i}}_{MC\ of\ lying}.$$

By rearranging with respect to $\theta_i$ we get

$$\theta_i \leq \theta_i^{full} = \left( \frac{\partial c(\omega_i, r_i)}{\partial r_i}\bigg|_{r_i = \overline{\omega}} \right)^{-1} \left( \frac{\partial v(\tilde{s}_{i2} + r_i - s_{i1})}{\partial r_i}\bigg|_{r_i = \overline{\omega}} + \frac{\partial u(\tilde{s}_{i2} + r_i)}{\partial r_i}\bigg|_{r_i = \overline{\omega}} \right).$$

Therefore, if the agent is sufficiently insensitive to lying, i.e., $\theta_i \leq \theta_i^{full}$, she will always lie fully ($r_i = \overline{\omega}$). □

LEMMA 2. *Individuals lie partially if $\theta_i \in (\theta_i^{full}, \theta_i^{true})$ and $\omega_i < \overline{\omega}$.*

If the agent's lying sensitivity is high enough not to engage in full lying but still not high enough to be willing to report truthfully, she will engage in partial lying. She will report a state which is between the true state and the best possible state, i.e., $r_i \in (\omega_i, \overline{\omega})$.

We proceed by analyzing behavior with respect to gains and losses in social image concerns, our key interest.

PROPOSITION 5. *There is more incentive to lie if an agent experiences a loss in social image than a gain in social image of the same size.*

*Proof.*

We compare cases denoted $(\Delta_i + \omega_i)^+$ and $(\Delta_i + \omega_i)^-$ in which $(\Delta_i + \omega_i)^+ = -(\Delta_i + \omega_i)^-$. Those cases are driven by changes in $s_{i1}$ or $\omega_i$, i.e., holding $\tilde{s}_{i2}$ constant, and they both imply zero lying costs and symmetry. We assume that an individual makes a lying decision after observing a true state $\omega_i$. We illustrate the proof in Figure 1. It follows from Proposition 2 that individuals will not lie downwards and therefore we only consider the case of $r_i \geq \omega_i$. We know that for $r_i = \omega_i$ the following holds:

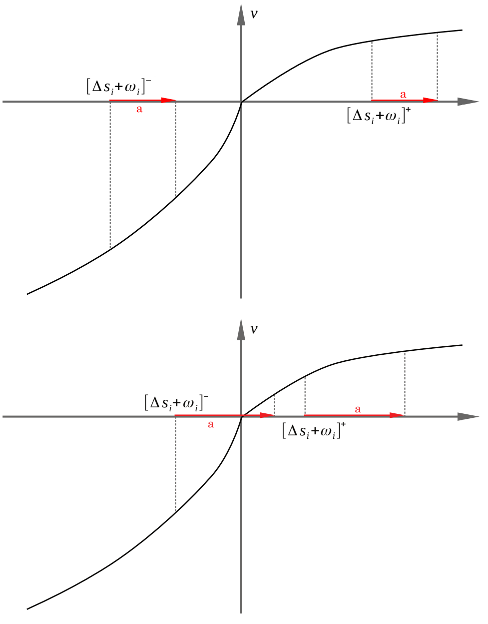



FIGURE 1. Illustration of a value function

*Note:* We display a value function that is in line with standard assumptions of prospect theory (Kahneman and Tversky, 1979) to illustrate the intuition of the proof of Proposition 5. In the top figure, we show the case of sufficiently small $a$, such that an agent in the loss domain remains in the loss domain after reporting $r_i = \omega_i + a$. In the bottom figure, we present a case of a sufficiently large $a$: In that case, an agent in the loss domain who reports $r_i = \omega_i + a$ switches to the gain domain.

$$v'\left((\Delta_i + \omega_i)^+\right) < v'\left((\Delta_i + \omega_i)^-\right). \tag{1}$$

Moreover, the value function is convex for losses, i.e., for any $a > 0$ it is true that

$$v'\left((\Delta_i + \omega_i)^-\right) < v'\left((\Delta_i + \omega_i + a)^-\right),$$

and concave for gains, such that

$$v'\left((\Delta_i + \omega_i)^+\right) > v'\left((\Delta_i + \omega_i + a)^+\right).$$

Then Condition 1 also holds for $r_i = \omega_i + a$:

$$v'\left((\Delta_i + r_i)^+\right) < v'\left((\Delta_i + r_i)^-\right) \tag{2}$$

and therefore reporting $r_i = \omega_i + a > \omega_i$ is more attractive if an individual is in the loss domain than the gain domain. Note that if $a$ is sufficiently large, $v\left((\Delta_i + \omega_i)^-\right) < 0$ but $v\left((\Delta_i + r_i)^-\right) > 0$ which means that the agent has been in the loss domain before reporting but has entered the gain domain by over-reporting. Condition 2 still holds in this case. $\square$

*Does loss aversion depend on the location of the reference point?* Our model also captures the idea of "having more to lose", as ceteris paribus, a higher first public signal—$s_{i1}$—causes a higher incentive to lie. This is because varying $s_{i1}$ changes the utility in the second period only through the difference $\Delta_i$, but does not change lying costs and social image utility per se.

PROPOSITION 6. *Given a fixed $\tilde{s}_{i2}$ and a fixed $\omega_i$, agents with a high reference point have more incentive to lie than agents with a low reference point for any level of $\Delta_i$.*

To summarize, our model predicts that individuals never under-report the true state and those without social image concerns never misreport the true state. If an agent cares about her social image and the true state is not the best possible one, she might engage in misreporting. Importantly, an agent has more incentives to misreport her true state if she experiences a loss in social image than a gain in social image. Moreover, an agent's incentive to misreport is stronger if her reference point is better.

## 3. Experiment design

*General setup.* Our experiment consists of two stages. Stage 1 is designed to establish a personal reference point for social image utility—a publicly reported rank in an intelligence test—against which subjects can fall short of or improve their image in Stage 2. In the second stage, we induce an exogenous change of the rank across two treatments. Subjects are then informed about their true rank and offered scope to manipulate the reporting of their rank to their peers. We test whether (i) subjects in the treatment in which subjects' average rank deteriorates—who experience a loss in social image—misreport their rank more strongly than those in another treatment who, on average, experience an increase in their rank and (ii) whether lying depends on the reference point from the first stage—namely the question of how high is a potential fall?

We create social image concerns through reporting a subject's ranking in a standardized test of fluid intelligence—Raven's Progressive Matrices test (1983)—to two randomly selected peers. Fluid intelligence encompasses logical reasoning and abstract thinking and constitutes an image providing trait for university students.[10] Public reporting of results shall hence create social image utility. In order to strengthen this link we explicitly mention in the instructions that the matrices (labeled as quizzes) are designed to measure fluid intelligence,

10. Our approach is similar to Falk and Szech (2020), Ewers and Zimmermann (2015), and Burks et al. (2013) who also use reporting of the performance in IQ or knowledge tests to induce image concerns.

that fluid IQ is an important part of an individual's overall IQ, and that such or related tasks are often employed in recruitment processes.

At the beginning of each session, two subjects per session are randomly assigned the role of peer observers. We randomly draw one observer from all male subjects and the other from all female subjects. This avoids possible gender-specific observer effects. After the observers have been determined, they stand up in front of the other subjects and announce "I am one of the two observers". The other subjects are randomly assigned to one of two treatments that vary the sequence of the quizzes over the two stages of the experiment. We label the treatments as *HardEasy* and *EasyHard*. In treatment *HardEasy* subjects work on a *Hard* quiz in Stage 1 and an *Easy* quiz in Stage 2 and in *EasyHard* on an *Easy* quiz in Stage 1 and a *Hard* quiz in Stage 2. At the end of the experiment, all subjects in both treatment groups have worked on the exactly same 48 matrices. All subjects—including the observers—received the same instructions. Then subjects performed two quizzes (consisting of 24 matrices each) and after each quiz report their relative performance (rank) to the observers. In the second stage, subjects have the possibility to lie in order to improve their rank before reporting it. Figure 2 illustrates the timeline of the experiment that we explain in detail below.

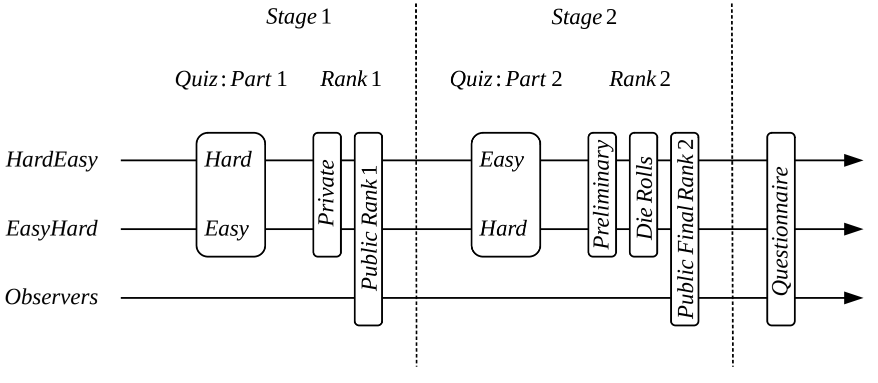



Figure 2. Timeline

*Treatment setup.* The original Raven's Progressive Matrices test (RPM) consists of 60 matrices that are divided into 5 equally sized sets (A to E) which increase in difficulty. Figure 3 provides an example of a Raven's Progressive Matrix. Subjects have to choose that box below the picture puzzle which is the best logical fit to the empty box within the picture. Progressive means that the matrices are increasing in difficulty. In our design, we do not use the 12 matrices of the easiest set A since we expect our student subjects to solve them all correctly. We split the remaining 48 matrices in two parts consisting of 24 matrices each that we will use for the quizzes. One quiz is easier

(*Easy*), while the other is harder (*Hard*). Both quizzes contain tasks from sets B to E. We calibrated the two sets such that *Hard* has a higher likelihood to contain matrices that have been solved by fewer subjects in a reference sample. The reference sample includes 413 observations (students) from a previous experiment which took place at the same lab in 2014. Subjects of the reference group solved exactly the same overall 48 matrices as our subjects.[11] In both quizzes, the difficulty of the tasks is gradually increasing over time. Matrices in quiz *Easy* and *Hard* do not repeat or overlap.

Subjects have 30 seconds to work on each matrix. The time limit ensures that performance is comparable across subjects: both within our experiment and with respect to the reference sample we use, in which subjects also had 30 seconds to work on each matrix. On average, it took subjects 11.5 seconds to answer a matrix. 2.7% of answers were provided in the last five seconds and in only 0.7% of cases subjects ran out of time, which suggests that the time limit was not restrictively binding. For each correctly solved matrix, subjects get one point. Wrong answers or no answer within the 30 seconds time limit do not give any points.

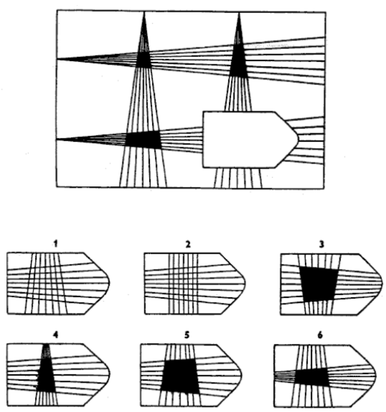



FIGURE 3. Example of a Raven's progressive matrix

*Stage 1.* After completing the treatment specific Raven's Matrices, subjects received private feedback on their relative performance (i.e., Rank 1) on their screen telling them that "$X$ % of the participants of the reference group have a higher rank than you in Quiz 1". A higher rank (lower $X$) implies better relative performance. The instructions provide several examples how individual rank is calculated and how to interpret it.[12]

11. The *Easy* quiz consists of the following matrices: B1, B5, B6, B7, B8, B9, B10, B11, B12, C1, C2, C3, C7, C8, C9, C10, C12, D2, D3, D5, D7, E2, E6, and E11. The *Hard* quiz contains the following matrices: B2, B3, B4, C4, C5, C6, C11, D1, D4, D6, D8, D9, D10, D11, D12, E1, E3, E4, E5, E7, E8, E9, E10, and E12.

12. We explicitly explain in instructions:

To determine the rank, we compare the share of correctly solved matrices among the first 24 matrices to the distribution of the share of correctly solved matrices among all 48 matrices of the reference sample. Our calibration of the matrix distribution between *Easy* and *Hard* ensures that subjects in treatment *EasyHard* will on average rank better than subjects in treatment *HardEasy* in Quiz 1 since both groups are compared to the same reference sample but the first 24 matrices are easier for subjects in treatment *EasyHard* than in *HardEasy*.

Subjects report their rank in the first stage to the observers. This establishes the individual Rank 1 as a personal reference point for social image concerns. Since subjects are randomized into treatments, their initial reference points before the feedback on Rank 1 are the same on average (given skill, ability, etc.). We give both subjects and observers detailed instructions on the reporting procedure to control the reporting process using the same protocol for all sessions. We instruct subjects to fill in report sheets named “Rank 1” and “Rank 2” in Stages 1 and 2, respectively, and to present these sheets to observers who verify the report. No further verbal communication between subjects and observers is allowed, i.e., the entire reporting procedure happens in silence. Report sheets contain two pieces of information: a 4-digit individual code and a rank. After each Stage, observers see a table on their computer screen in which each individual code corresponds to a rank, and thus can compare the report sheet to the true information from the table. If the reported rank matches the true rank, observers stamp the report sheet to verify it.[13] We organized our laboratory setup in a way that subjects cannot see observers' computer screens while reporting their rank. Additionally, to assure anonymity, we use 4-digit individual codes instead of cubicle numbers which, in the unlikely case of a subject seeing the table on the observer's screen, makes it uninformative.

*Stage 2.* Subjects work on the remaining 24 matrices. For subjects in treatment *EasyHard*, Stage 2 is more complicated than Stage 1. In expectation, they rank worse than in Part 1. For subjects in treatment *HardEasy*, rank improves in expectation. We construct a Preliminary Rank 2 by comparing the overall individual correctly solved number of matrices to their distribution in the reference group. Consequently, we do not expect the average Preliminary Rank 2 to differ across the two treatments. After completing the task in this

“For example, the statement “9% of participants of the reference group have a higher rank than you in part 1” implies that “9% performed better than you (i.e., they solved a higher share of the overall 48 matrices from part 1 and 2 correctly than you) and 90% worse (i.e., they solved a lower share of the matrices correctly than you). That means you belong to the 10% of best performers in solving the matrices that were designed to measure fluid IQ.”

13. Examples of filled in and verified report sheets (in German) as well as their translations to English are shown in Appendix Figures C.1 and C.2 for Ranks 1 and 2, respectively.

stage, both Rank 1 and the Preliminary Rank 2 are displayed privately to each subject, so that subjects can compare their ranking in the two stages. While average Preliminary Rank 2 does not differ systematically across treatments, subjects' average reference point (Rank 1) will be better in treatment *EasyHard* than *HardEasy*.

*Die reports.* After learning about their ranks, subjects are asked to throw a die twice and report the rolled numbers. The first reported number is then added to the number of correctly solved matrices in the reference group. The second reported number is added to a subject's own number of correctly solved matrices, giving the subjects two ways of cheating on the final reported rank that bear exactly the same consequences for their social image.

We use a modified version of the die roll task by Fischbacher and Föllmi-Heusi (2013).[14] Each subject rolls the die in private in the cubical so that no one, including the experimenters, can observe the actually rolled numbers.[15] Building on the work of Abeler et al. (2019), we use lying costs which increase in the size of the lie to quantify utility changes due to changes of social image. Importantly, lying cannot be detected at the individual level in the die roll task. However, the underlying distribution of true die roll outcomes is known such that it can be observed whether and how much subjects lie on average within a treatment group. Hence, we will conduct our main analysis on the treatment level.

Including two die rolls instead of only one has several advantages. First, a subject's Final Rank 2 can either be better or worse than the Preliminary Rank 2. Adding a smaller number to the reference sample's score than to the own score will improve a subject's Final Rank 2 compared to Preliminary Rank 2, and vice versa. Second, if subjects have a preference for telling the truth, two die rolls help to satisfy various preferences for truth-telling: Subjects can, for example, tell the truth about the die rolls to the experimenters by reporting the actual numbers they have rolled. Alternatively, subjects can tell observers the truth about their Preliminary Rank 2 by reporting the same number for both die rolls such that the Final Rank 2 is exactly the same as their Preliminary Rank 2. This option would not be available with only one die roll.

Additionally, if we based the ranking system on comparing subjects only within the current experiment (for example, ranking them from best to worst

14. In Fischbacher and Föllmi-Heusi (2013), subjects roll a die once, report on the rolled number (which does not necessarily need to be the truly rolled number), and are paid according to the reported number (i.e., higher numbers give a higher payoff except for 6, which pays zero). We build on the original die roll task but adjust it for our purposes in two aspects. First, instead of using monetary payoffs, we reward subjects with additional points which add up to the number of correctly solved matrices. Thus, lying enables subjects to improve their rank. Second, our subjects are told to throw the die twice.

15. According to Gneezy et al. (2018), the fact that the experimenter cannot observe participants' true outcomes facilitates lying.

score), there would be an incentive to add a higher number to the own score if subjects expect others to add a high number to their score. So, subjects' lying behavior would depend on their beliefs on others' lying. In order to avoid this and to be able to interpret lying as a reflection of image concerns independent of individual beliefs, it is important to construct a ranking system which compares subjects to a predetermined reference group one by one.

*Further remarks.* Introducing observers instead of allowing subjects to report their rank to each other has two major advantages. First, our subjects do not get feedback on others' rank which could affect their perception of their own social image. Second, observers only know about the existence of a "further task" on top of the second quiz in Stage 2 and that the score in this task will feed into a subject's Final Rank 2. Observers are not informed about the exact nature of the die roll task, do not know how and to which extent the further task influences final ranks, and this is common knowledge to all subjects.[16] Consequently, subjects do not risk loosing social image because of possible reputation cost of being seen as a liar. The remaining subjects receive the instructions regarding the die roll task on their computer screen after they have worked on Part 2 of the quiz.

Once the reported die rolls have been added and Final Rank 2 calculated, subjects go to observers again and report their Final Rank 2. After Stage 2, observers' information tables include, for each subject, the individual code, Final Rank 2, Rank 1 and the difference between Final Rank 2 and Rank 1. This is common knowledge for all subjects. Reporting procedures are the same as in Stage 1.

*Procedural details and implementation.* Our experiment design and hypotheses are preregistered on AEA RCT Registry.[17] We conducted our experiment using zTree (Fischbacher, 2007). After two pilot sessions as a prerequisite for power calculations, we run 19 main sessions in the DICE Lab, University of Düsseldorf between November 2018 and November 2019.

383 subjects participated, 38 as observers. 177 subjects (51%) were randomly assigned to treatment *HardEasy* and the remaining 168 (49%) to treatment *EasyHard*. We randomized within each session in order to balance the two treatments with respect to possible confounding factors such as day of the week, time of the day, or weather. Our sample mainly consists of a student population and was recruited using ORSEE (Greiner, 2015). 142 subjects were male (67 in treatment *HardEasy* and 75 in *EasyHard*), 203 were female (110 in treatment *HardEasy* and 93 in *EasyHard*). Age varied between 18 and 63 years with a median age of 23 years and 95% of subjects being younger than 33 years. No particular exclusion criteria applied.

16. The role of observers is passive: They are not allowed to communicate with subjects.

17. Petrishcheva, Vasilisa, Gerhard Riener, and Hannah Schildberg-Hörisch. 2019. "Loss Aversion in Social Image Concerns." AEA RCT Registry. April 09. https://doi.org/10.1257/rct.3422-5.0.

All participants received a flat payment of 12 Euro, but no additional performance-contingent payment for correctly solving the matrices, which was clearly communicated to the subjects. This enables us to test whether solving matrices is indeed an image-relevant task: Subjects' behavior indicates image concerns if they exert effort to solve the matrices correctly, even if this does not increase their monetary reward. On average, subjects earned €12.65, which includes the €12 flat payment plus one lottery outcome (as described below). In total, the experiment lasted about 90 minutes (including payment).

*Post-experimental questionnaire.* The questionnaire provides information on socio-economic and demographic characteristics (age, gender, high school GPA, last math grade at school, student status and field of study, previous participation in experiments). It also assessed subjects' general willingness to take risks, based on a question from the German Socio-Economic Panel (GSOEP) questionnaire as well as the importance of social image, using the following question (similar to the one used by Ewers and Zimmermann (2015)): "How important is the opinion that others hold about you to you?". Additionally, following Gächter et al. (2007) and Fehr and Goette (2007), we measure loss aversion in the monetary domain using a set of incentivized lotteries which subjects can choose to accept or decline. Appendix E provides the exact wording of the entire questionnaire.

## 4. Hypotheses

First, we test that our RPM-based task is indeed image-relevant for our subjects. Since their payment is unrelated to performance, exerting effort on solving the matrices will provide evidence for the relevance of either social and/or self-image concerns in our experiment design. This leads us to Hypothesis 1(a). Moreover, if subjects have social image concerns they will over-report as shown in Proposition 2, leading to Hypothesis 1(b).

HYPOTHESIS 1. (Image relevance of task)

(a) *On average, subjects will exert substantial effort on solving the matrices.*
(b) *On average, subjects will over-report their score.*

In our experiment design, over-reporting implies that subjects report higher die rolls for themselves than for the reference group to be able to report a better Final Rank 2 to the observers. This behavior establishes the relevance of social image concerns for our subjects as a whole.

HYPOTHESIS 2. (Unconditional loss aversion in social image concerns) *Subjects with sufficiently strong social image concerns report higher die roll differences in treatment* EasyHard *than in treatment* HardEasy.

We hypothesize that subjects in treatment *EasyHard* (who on average experience a loss in social image since their rank deteriorates from Stage 1 to Stage 2) lie more than subjects in treatment *HardEasy* (who on average experience a gain in social image since their rank improves from Part 1 to Part 2). We compare the average difference in die roll reports (average reported number to be added to own performance minus average reported number to be added to the reference group's performance) from treatments *HardEasy* and *EasyHard*. If this difference is significantly higher in treatment *EasyHard* than in treatment *HardEasy*, this provides evidence for loss aversion in social image concerns unconditional on reference point because it implies that subjects who risk losing social image are ready to lie more than those with social image gains.

We also test for loss aversion in social image concerns conditional on the reference point: We hypothesize that subjects who performed well in Stage 1 resulting in a better Rank 1 (reference point), i.e., those who have reputation to lose, are more strongly loss averse than those who have less reputation to lose, as summarized in Proposition 6.

Hypothesis 3. (Loss aversion in social image concerns conditional on reference point) *Subjects with sufficiently strong social image concerns and a good Rank 1 report higher die roll differences in treatment* EasyHard *than in treatment* HardEasy.

## 5. Results

First, we will establish that the matrix task is a source of image-concerns. We will then proceed by analyzing how subjects react to losses as opposed to gains in social image, both by exploiting our exogenous treatment variation and by providing descriptive analyses of lying behavior at the individual level.

### 5.1. Relevance of social image

We start the discussion comparing the performance on the matrix task between the treatments. As intended by our design, subjects in treatment *HardEasy* performed worse in Part 1 than subjects in treatment *EasyHard*. On average, they solved 2.6 matrices less in Stage 1 than subjects in treatment *EasyHard*. This then also reflects in the average Rank 1 in treatment *HardEasy* that was 61.5%, while it was 29.7% in treatment *EasyHard*. Figure 4 displays the kernel densities of Rank 1 (left) and 2 (right) by treatment. The difference of Rank 1 distributions between treatments is highly significant (Mann-Whitney

$U$ test, MWU, $p < 0.0001$).[18] Thus, the exogenous manipulation of Rank 1, the reference point for social image concerns, worked as expected.

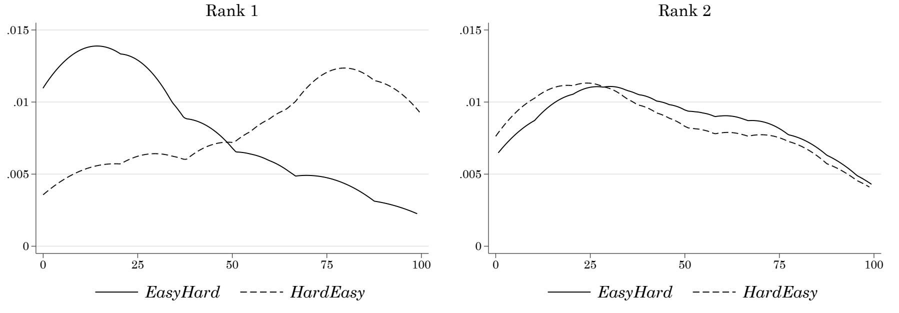



FIGURE 4. Distributions of Rank 1 and Preliminary Rank 2 by treatment

*Note:* The best possible rank is 1, while the worst is 100. Densities are estimated using Epanechnikov kernels with a bandwidth of 15.

The total number of solved matrices after completion of Part 2 is similar across treatments. On average, subjects in treatment *HardEasy* and *EasyHard* solve 39.2 and 38.5 matrices, respectively, which results only in a small, average difference in Preliminary Rank 2 of 3.7 percentage points between treatments. The difference in distributions of Preliminary Rank 2 between treatments is not significant (MWU test, $p = 0.2027$). This ensures that possible differences in average lying across treatments do not reflect differences in Preliminary Rank 2 but only in the reference point for social image concerns, Rank 1.

Moreover, the numbers above underline that subjects exerted substantial effort on the quizzes. They solved an average of 38.8 out of all 48 matrices correctly. No subject solved less than 20 matrices, and more than 90% of subjects gave 34 or more correct answers. Note that the cumulative probability of correctly solving 20 or more matrices by guessing is close to zero ($p < 0.0001$). Since correct answers are not incentivized monetarily, substantial effort provision suggests that image concerns are a likely driving force behind solving the matrices.

*What did subjects report about their die rolls?* They reported two values: The variable $DieSubject$ which is added to their own score and the variable $DieSample$ which is added to the scores of all subjects in the reference sample. In the absence of lying, die roll reports for each of the variables should follow a discrete uniform distribution with the support $\{1, \dots, 6\}$ and an average of 3.5. Figure 5 displays histograms of $DieSubject$ (left) and $DieSample$ (right) as well as the probability density function of the uniform distribution (red line). The average of $DieSubject$ is 4.03 and we reject the null hypothesis for the

18. Throughout the paper, we report two-sided tests and refer to results as (weakly/highly) significant if the two-tailed test's $p$-value is smaller than 0.05 (0.10/0.01).

point prediction ($t$-test, $H_0$: $DieSubject = 3.5$, $p < 0.0001$). The distribution of $DieSubject$ is also highly significantly different from the discrete uniform distribution (Pearson's $\chi^2$-test, $p < 0.0001$) and left-skewed. In contrast, the average of $DieSample$ is 3.43 which is not significantly different from 3.5 ($t$-test, $p = 0.4614$). Moreover, the distribution of $DieSample$ does not differ significantly from the discrete uniform distribution (Pearson's $\chi^2$-test, $p = 0.881$).

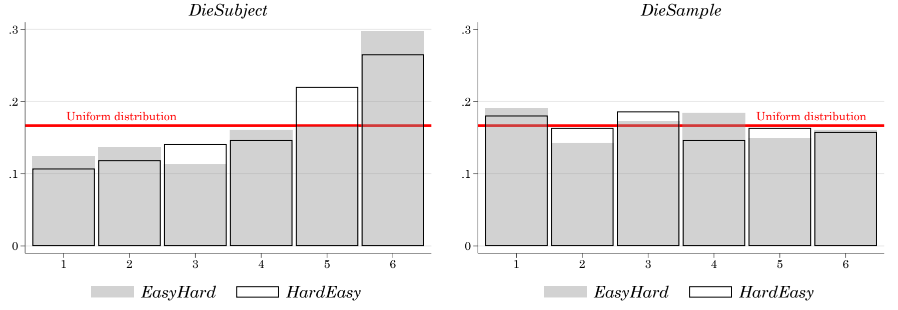



FIGURE 5. Distributions of *DieSubject* and *DieSample* by treatment

*Note:* Figures illustrate histograms of *DieSubject* (left) and *DieSample* (right). Horizontal axis indicates reported die rolls (from 1 to 6). Vertical axis indicates the fraction of subjects who reported the respective die rolls. Absent misreporting, die rolls should follow uniform distributions (red lines).

Subtracting $DieSample$ from $DieSubject$ results in the die roll difference, $DieDiff$, which indicates whether subjects improve or worsen their Final Rank 2 through reporting. The higher $DieDiff$, the better becomes *Final Rank 2*. In principle, $DieDiff$ can vary between $-5$ and 5, and, in the absence of lying, follows a discrete binomial distribution with zero mean. Pooling the data from both treatments, our subjects report an average die roll difference of 0.59 which is highly significantly different from zero ($t$-test, $p < 0.0001$). As illustrated in Figure 6, the values of 4 and 5 are significantly over-reported (binomial probability tests, two-sided $p = 0.0253$ and $p < 0.0001$ for the values of 4 and 5, respectively). Thus, subjects lie both fully (maximal over-reporting) and partially (less than maximal over-reporting) which is in line with our theoretical predictions and experimental evidence of Gneezy et al. (2018) and Fischbacher and Föllmi-Heusi (2013). Over-reporting high values of $DieDiff$ provides further evidence that subjects perceive our matrices task as image-relevant and additionally shows that *social* image concerns matter: as all subjects know their Preliminary Rank 2, over-reporting their own score is unlikely to improve their self-image.

RESULT 1. *(a) Subjects solve more matrices correctly than expected by simple guessing.*

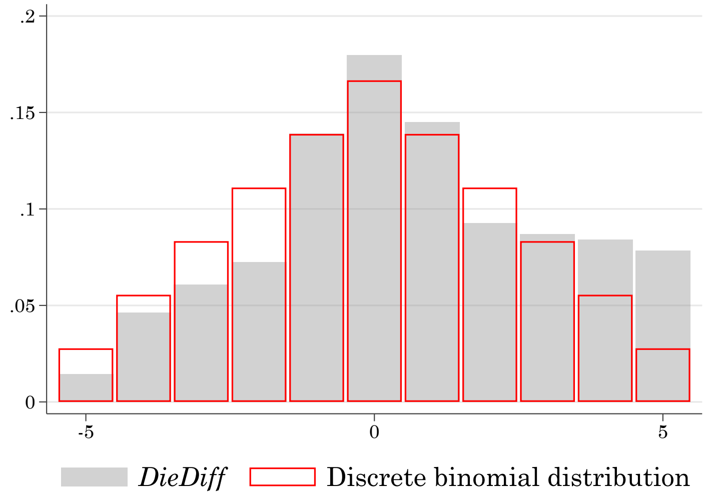



FIGURE 6. Reported die roll difference

*Note:* Figure illustrates histogram of $DieDiff$. Horizontal axis indicates a reported die roll difference (from $-5$ to 5, higher $DieDiff$ means adding more to one's own score). Vertical axis indicates the fraction of subjects who reported the respective die roll difference. Absent misreporting, die roll difference should follow the discrete binomial distribution (red outlines).

*(b) Subjects report higher die rolls to be added to their own score than expected by rolling a fair die.*

This first set of results suggests that on average, public reporting of own performance in the Raven's matrices induces social image concerns and that subjects engage in lying in order to report higher ranks to the observers.

### 5.2. Gains and losses in social image

*5.2.1. Treatment comparison.* We now turn to the effect of our treatments on reporting behavior. Obviously, loss aversion in social image can only be observed for those subjects who indeed care about their social image and do so sufficiently to bear the lying costs involved. While we have shown above that many of our subjects do over-report, it is also well documented that people are heterogeneous in the degree of social image concerns (see Bursztyn and Jensen, 2017; Friedrichsen and Engelmann, 2018) and lying costs (Abeler et al., 2019). This is also true in our sample, as Figure C.3 in Appendix C shows.

According to Hypotheses 2 and 3, we are particularly interested in testing whether subjects with social image concerns are loss averse in social image. Therefore, we present two sets of results: (a) evidence from subjects with social image concerns and (b) evidence for our sample as a whole. We classify subjects based on a median sample split on social image concerns as measured

Table 1. Regression analysis: die roll difference (two-limit Tobit)

| | Social image concerns | | Whole sample | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| EasyHard | 0.356 | 1.522** | -0.050 | 0.392 |
| | (0.529) | (0.684) | (0.388) | (0.366) |
| | [0.528] | [0.026] | [0.908] | [0.298] |
| EasyHard×Rank 1 | | -0.024* | | -0.008 |
| | | (0.011) | | (0.009) |
| | | [0.050] | | [0.387] |
| Rank 1 | | 0.016** | | 0.006 |
| | | (0.007) | | (0.005) |
| | | [0.016] | | [0.362] |
| Constant | 0.425 | -0.528 | 0.695*** | 0.319 |
| | (0.353) | (0.470) | (0.242) | (0.313) |
| Number of obs. | 173 | 173 | 345 | 345 |

*Note:* *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$ based on score-bootstrap $p$-values. In columns (1) and (2), we split our sample based on the median of the importance of social image concerns that subjects reported on a 11-point Likert-scale. Individuals who reported 6 or higher are categorized as "Social image concerns". Rank 1 ranges from 0 to 100, with lower values indicating better rank. All columns display two-limit tobit estimates with score bootstrap clustering at the session level. Standard errors clustered at the session level in parentheses. Score-bootstrap $p$-values in brackets.

at the individual level through our survey instrument: "How important is the opinion that others hold about you to you?" (11-point Likert scale, social image concerns if answer 6 or higher).[19]

Table 1 displays our main results that are all based on two-limit tobit models to account for the censored nature of the dependent variable $DieDiff$. Columns (1) and (3) provide a first test for unconditional loss aversion in social image concerns for subjects with social image concerns and the sample as a whole, respectively; columns (2) and (4) for loss aversion in social image concerns conditional on the reference point, Rank 1. Standard errors are clustered at the session level in order to account for possible within-session correlations generated by observers. We use score bootstrap clustering with null imposed, Rademacher weights and 999 replications in order to account for the relatively small number of clusters (19 sessions) and to obtain a conservative estimate of the standard errors.

Regarding Hypothesis 2, there is no evidence for loss aversion in social image concerns pooled over all reference points. While the constant terms in columns (1) and (3) indicate that subjects over-report their own score, subjects in treatments *HardEasy* and *EasyHard* do not differ significantly in their average over-reporting. As expected, subjects with social image concerns in treatment *EasyHard* tend to overreport more strongly; however, this effect

19. Social image concerns do not differ significantly across treatments (MWU test, $p = 0.1514$).

is not significant. Similarly, MWU tests confirm that subjects' overall lying behavior does not differ across treatments. They yield $p = 0.9108$ for $DieDiff$, $p = 0.8970$ for $DieSubject$ and $p = 0.9232$ for $DieSample$, respectively.

RESULT 2. *On average, subjects with social image concerns do not over-report significantly more in treatment* EasyHard *than* HardEasy *irrespective of Rank 1.*

We continue by investigating Hypothesis 3 that postulates the existence of loss aversion in social image concerns conditional on having a high reference point (low Rank 1) and caring about social image in columns (2) and (4) of Table 1. Intuitively, subjects who perform well in Part 1 and therefore have a reputation to lose in Part 2 may be more averse to losing social image than subjects who have no reputation to lose since they are ranked less favorably in Part 1. This is indeed what we find for subjects with social image concerns.

In particular, estimation results in column (2) imply that for high initial performance in Part 1 (high reference point which means low Rank 1) subjects in Treatment *EasyHard* over-report substantially more than subjects in Treatment *HardEasy* who on average have lower reference points. For the best possible Rank 1 of zero, the difference in over-reporting to one's own advantage is substantial: it amounts to 1.5 units on the die roll difference scale from -5 to 5. For each 10 percentage point increase in Rank 1 (decrease in reference point), e.g., moving from 10 to 20 in Rank 1, subjects in Treatment *EasyHard* lie 0.08 units less ($= (-0.016 + 0.024) \times 10$). Thus, in Treatment *EasyHard*, subjects with higher reference points in social image over-report more, while those who have less reputation to lose do so much less. We show that these findings are robust to different social image splits in Appendix B.

Interestingly, we observe the opposite pattern for subjects with social image concerns in Treatment *HardEasy* who typically start with worse initial reputation: they over-report more in Part 2, the worse their initial reputation in Part 1, i.e., the higher Rank 1. This explains why we do not observe a significant treatment difference in column (1), which tests for loss aversion in social image unconditional on the reference point.

We further illustrate results from Table 1 in Figure 7(a-b) which depicts how subjects with different reference points differ in their reported die roll difference and between treatments. Figure 7(a) focuses on participants with social image concerns. According to MWU tests, subjects with Rank 1 better than average lie weakly significantly more in treatment *EasyHard* than in *HardEasy* ($p = 0.0609$), while those with Rank 1 worse than average do not differ in their misreporting behavior across treatments ($p = 0.7345$).

*5.2.2. Individual-level analysis.* After comparing behavior across treatments, we now turn to an individual level analysis to study reactions to losses as opposed to gains in social image. This approach allows accounting for subjects

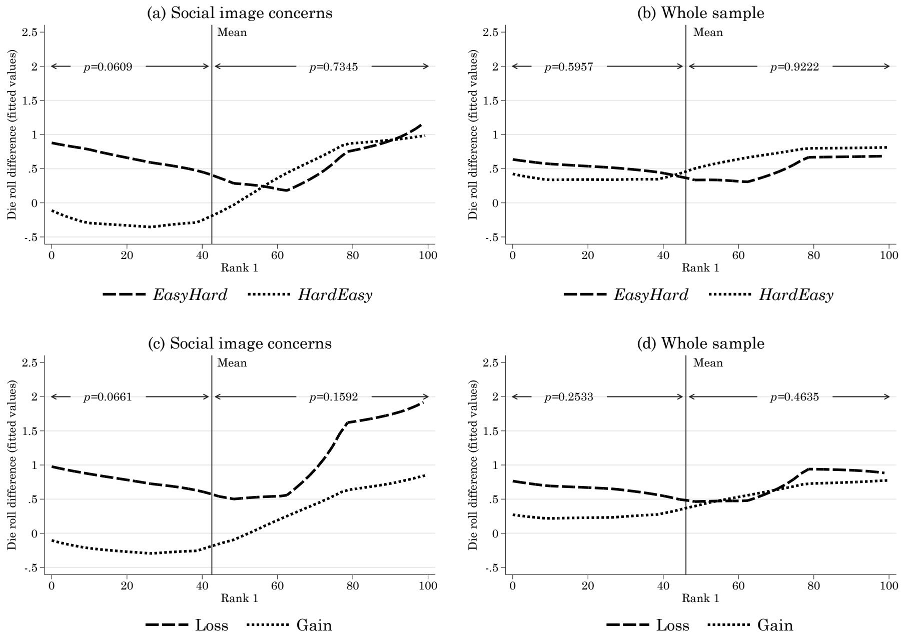



FIGURE 7. Die roll difference by Rank 1

*Note:* Figures (a)-(d) illustrate the dynamics of die roll differences over Rank 1 for a variety of sub-samples: (a) and (c) show subjects who reported the importance of social image to be 6 or higher, (b) and (d) show the whole sample. Figures (a)-(b) display the difference between treatments *HardEasy* and *EasyHard*; Figures (c)-(d) display the difference between subjects who experience actual losses versus actual gains in social image (i.e., positive rank differences in contrast to negative rank differences). Fitted values are estimated using Epanechnikov kernel with a bandwidth of 20. Reported $p$-values indicate MWU test results for die roll differences between treatments in (a)-(b) or for gains versus losses in (c)-(d). Displayed test results refer to comparisons of either above or below mean in Rank 1.

who were, for example, assigned to treatment *HardEasy* and thus on average expected to experience a gain in social image but who performed extraordinarily well in Part 1 relative to Part 2, such that they actually experienced a loss in social image.

In Figures 7(c) and 7(d), we consider subjects who experience actual gains and losses in social image. Negative rank differences are labeled “Loss” and positive rank differences “Gain”. For *any* reference point, subjects with image concerns who experience an actual loss in social image misreport more than those who experience an actual gain (MWU test, $p = 0.0754$), which is in line with Hypothesis 2. In line with Hypothesis 3, this effect is especially pronounced for those with better than average Rank 1 (MWU test, $p = 0.0661$).

We further explore subjects’ lying behavior using an instrumental variable approach, which again allows for correctly assigning each individual to gains and losses in social image. We define a dummy variable *Loss* which equals one

TABLE 2. Regression analysis: die roll difference, instrumental variable approach

| | Social image concerns | | Whole sample | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Loss | 0.449 | 1.710** | -0.066 | 0.494 |
| | (0.651) | (0.729) | (0.518) | (0.444) |
| | [0.537] | [0.030] | [0.908] | [0.289] |
| Loss×Rank 1 | | -0.026* | | -0.011 |
| | | 0.014 | | 0.015 |
| | | [0.072] | | [0.446] |
| Rank 1 | | 0.016 | | 0.006 |
| | | (0.007) | | (0.005) |
| | | [0.147] | | [0.274] |
| Constant | 0.419 | -0.554 | 0.696*** | 0.304 |
| | (0.356) | (0.491) | (0.249) | (0.329) |
| F-statistic (first stage, Loss) | 288.93 | 330.34 | 395.46 | 242.55 |
| F-statistic (first stage, Loss×Rank 1) | | 35.33 | | 73.28 |
| F-statistic (first stages, joint) | | 35.29 | | 76.00 |
| Number of obs. | 173 | 173 | 345 | 345 |

*Note:* *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$ based on wild-bootstrap $p$-values. Columns (1) and (2) contain estimations for the sub-sample of subjects who reported the importance of social image concerns to be 6 or higher, while columns (3) and (4) show estimations for the whole sample. Rank 1 ranges from 0 to 100, with lower values indicating better rank. Standard errors clustered at session level in parentheses. Wild bootstrap $p$-values in brackets. All first-stage F-statistics are cluster-robust Kleibergen-Paap rk Wald F-statistics.

if the true rank difference at the individual level, i.e., Rank 1 minus Preliminary Rank 2, is negative and zero otherwise.

Table 2 summarizes the IV estimates. In the first stage, we instrument an actual loss in social image with our exogenously assigned treatment dummy which allows us to correctly account for those subjects who were assigned to *HardEasy*, but had an actual loss in social image as well as those assigned to *EasyHard* but who nevertheless experienced a gain in social image. More specifically, we estimate two first-stage regressions: We instrument the loss in social image concerns with the exogenously assigned treatment *EasyHard* as well as its interaction with Rank 1. We document separate F-statistics of 330.34 and 35.33, respectively, as well as a joint cluster-robust F-statistic of 35.29, which confirms the relevance of the chosen instruments.

In the second stage, we rely on models (1)–(4) from Table 1 and estimate how the instrumented loss in social image concerns as well as its instrumented interaction with Rank 1 influence misreporting behavior. The results confirm our previous findings, namely, that subjects with social image concerns lie more to their favor when experiencing a loss in social image compared to those who experience a gain, and they tend to lie more if their reference point for reputation is higher. In particular, best-ranked subjects who experience a loss in their social image report a 1.71 higher die roll difference than those who experience a gain in social image. As Rank 1 increases (i.e., gets worse) the difference in misreporting between those in loss and gain in social image gets smaller. Unsurprisingly, the effect in Table 2 is of a similar magnitude as the one in Table 1: Our treatments provided a good exogenous variation in rank differences, such that only 19 of 173 subjects with social image concerns are reassigned in the IV approach.[20]

In sum, the IV approach provides individual-level, causal evidence of differences in misreporting behavior between subjects who experience gains and losses in social image concerns.

Result 3. *Subjects with social image concerns and a high initial reference point report significantly higher die roll differences in case of image losses than for image gains.*

### 5.3. Concave utility function or loss aversion: Is there a discontinuity in misreporting behavior when moving from gains to losses in social image?

We have established that, for subjects who sufficiently care about their social image, misreporting behavior differs systematically when experiencing gains and losses in social image—a pattern that is predicted if subjects are loss averse in social image concerns. However, an alternative explanation for such a pattern is a simple concave utility function over changes in social image, which also implies that losses in social image induce stronger changes in utility than equally sized gains in social image. For an illustration of a standard concave utility function, see the solid, black line in Figure 8. In contrast, the dashed line depicts a value function that is compatible with the assumption of loss aversion.

In order to differentiate between both possible explanations of misreporting behavior, we present results from a regression discontinuity design in Table 3. The regression discontinuity specification maps the first derivative of the value function $v$ which is commonly assumed to be larger for losses than for gains around zero and discontinuous at zero. Allowing for a discontinuity (RD) at a rank difference of zero (i.e., at the intersection of both axes in Figure 8) allows exploring whether subjects report systematically different die roll differences when moving from the loss to the gain domain in social image. If we find such a significant discontinuity in the derivative of the value function at the rank difference of zero, the empirical approximation of the value function has a kink—as is generally assumed in prospect theory. In contrast, such a kink would not be compatible with a standard concave utility function $v'$ for changes in social image.

20. In a whole sample, there are 10 subjects with a rank difference of zero and 3 subjects with a negative rank difference in *HardEasy*. In *EasyHard*, 20 subjects have a zero rank difference and 19 subjects a positive rank difference. Subjects with a rank difference of zero are assigned to the "Gain" category. By introducing "Gain" and "Loss", we reassign those overall 42 of 345 individuals to the intended category.

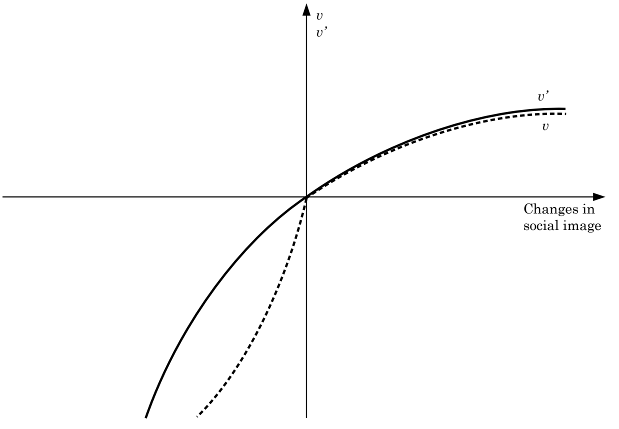



FIGURE 8. Illustration of potential value functions

TABLE 3. Regression discontinuity design

| | Social image concerns | | Whole sample | |
|---|---|---|---|---|
| | Cross-validation | CCT | Cross-validation | CCT |
| | (1) | (2) | (3) | (4) |
| RD estimates | -1.818*** | -1.914* | -1.267*** | -1.444* |
| Conventional Std. Err. | 0.696 | 1.014 | 0.488 | 0.759 |
| Conventional $p$-value | 0.009 | | 0.009 | |
| Robust $p$-value | | 0.099 | | 0.099 |
| Number of obs. | 160 | 81 | 318 | 149 |

*Note:* *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$ based on conventional $p$-values for cross-validation in columns (1) and (3) and robust $p$-values for CCT estimates in columns (2) and (4). We use local-linear estimators around a rank difference of zero with Epanechnikov kernels and two different optimal bandwidth selection criteria: the cross-validation procedure and the MSE-optimal bandwidth selection criterion (Calonico et al. (2014), CCT). Estimations are either based on the full sub-sample of subjects with social image concerns (173 subjects who reported the importance of social image concerns to be 6 or higher) in columns (1) and (2) or on the whole sample (345 subjects) in columns (3) and (4) and the reported number of observations indicates how many observations were actually used given a particular bandwidth selection criterion.

Overall, we find a significant discontinuity at the rank difference of zero for subjects with social image concerns as well as for the whole sample. On average, subjects below the threshold who experience a small loss in social image report 1.3–1.4 higher die roll differences than those above who experience a small gain in social image. For subjects with social image concerns the discontinuity is even more pronounced: those below the threshold report on average 1.8–1.9 higher die roll differences than those above. These findings are robust in two different specifications: (i) conventional RD estimates with an optimal bandwidth selected by a cross-validation procedure (columns (1) and (3)) and (ii) the robust procedure of Calonico et al. (2014) (CCT), employing the MSE-optimal bandwidth selection criterion (columns (2) and (4)). Conventional RD estimates with the cross-validation selection criterion suffer from potentially biased standard errors, while CCT uses debiased standard errors, allowing for correct inference on the treatment effect (Calonico et al., 2014).

Result 4. *We observe a significant discontinuity in lying behavior at the rank difference of zero, indicating a kink in the value function for social image as predicted by loss aversion.*

## 6. Conclusion

Does loss aversion apply to social image concerns? In sum, we observe loss aversion in social image concerns for those individuals who care about their reputation and have a reputation to lose. When taking a closer look at subjects' behavior when moving from losses to gains in social image, we find a sharp decrease in lying—providing evidence for social image concerns irrespective of initial reputation and extent of social image concerns.

More generally, our findings underline that loss aversion can also play a role in the non-material domain. While loss aversion is a well-established phenomenon for money and material goods (Kahneman et al., 1991), our findings take a first step in a new line of research investigating the relevance of loss aversion to non-material sources of utility such as various drivers of reputation or self-image.

Since our experimental paradigm quantifies utility changes due to changes in social image by the amount of lying that individuals are willing to engage in, our findings also speak to the manifold situations in which honest reporting of private information is of great importance but not necessarily incentive-compatible. Dai et al. (2018) have shown that dishonesty in the lab can predict fraud and rule violation in real life. Our results reveal that individuals who care about their social image and have a high reputation to lose are likely to report more dishonestly than others. Thus, monitoring efforts should be targeted at those groups. Moreover, one should try to make it harder to lie while keeping a good reputation, e.g., via transparency, naming-and-shaming, or reputation systems (see also Abeler et al., 2019).

Finally, we find that the way social image evolves over time affects behavior. While making a decision, this reference-dependence implies that individuals may not only take present or discounted future reputation into consideration, but also account for the history of their social image. Two otherwise identical individuals may thus take opposite actions only due to differences in their social image in the past.

**References**


Abeler, Johannes, Daniele Nosenzo, and Collin Raymond (2019). "Preferences for Truth-Telling." *Econometrica*, 87(4), 1115–1153.

Alsop, Ron (2008). *The Trophy Kids Grow Up: How the Millennial Generation is Shaking Up the Workplace*. John Wiley & Sons.

Andreoni, James and B Douglas Bernheim (2009). "Social image and the 50–50 norm: A theoretical and experimental analysis of audience effects." *Econometrica*, 77(5), 1607–1636.

Araujo, Felipe A, Erin Carbone, Lynn Conell-Price, Marli W Dunietz, Ania Jaroszewicz, Rachel Landsman, Diego Lamé, Lise Vesterlund, Stephanie W Wang, and Alistair J Wilson (2016). "The slider task: an example of restricted inference on incentive effects." *Journal of the Economic Science Association*, 2(1), 1–12.

Ariely, Dan, Anat Bracha, and Stephan Meier (2009). "Doing good or doing well? Image motivation and monetary incentives in behaving prosocially." *American Economic Review*, 99(1), 544–55.

Aurand, Timothy, Wayne Finley, Vijaykumar Krishnan, Ursula Sullivan, Jackson Abresch, Jordyn Bowen, Michael Rackauskas, Rage Thomas, and Jakob Willkomm (2018). "The VW Diesel Scandal: A Case of Corporate Commissioned Greenwashing." *Journal of Organizational Psychology*, 18(1), 23–32.

Austen-Smith, David and Roland G Fryer Jr (2005). "An economic analysis of "acting white"." *The Quarterly Journal of Economics*, 120(2), 551–583.

Barberis, Nicholas C (2013). "Thirty years of prospect theory in economics: A review and assessment." *Journal of Economic Perspectives*, 27(1), 173–96.

Bénabou, Roland and Jean Tirole (2006). "Incentives and prosocial behavior." *American Economic Review*, 96(5), 1652–1678.

Bleichrodt, Han, Jose Luis Pinto, and Peter P Wakker (2001). "Making descriptive use of prospect theory to improve the prescriptive use of expected utility." *Management Science*, 47(11), 1498–1514.

Booij, Adam S and Gijs Van de Kuilen (2009). "A parameter-free analysis of the utility of money for the general population under prospect theory." *Journal of Economic Psychology*, 30(4), 651–666.

Burks, Stephen V, Jeffrey P Carpenter, Lorenz Goette, and Aldo Rustichini (2013). "Overconfidence and social signalling." *The Review of Economic Studies*, 80(3), 949–983.

Bursztyn, Leonardo, Georgy Egorov, and Robert Jensen (2019). "Cool to be smart or smart to be cool? Understanding peer pressure in education." *The Review of Economic Studies*, 86(4), 1487–1526.

Bursztyn, Leonardo and Robert Jensen (2017). "Social image and economic behavior in the field: Identifying, understanding, and shaping social pressure." *Annual Review of Economics*, 9, 131–153.

Calonico, Sebastian, Matias D Cattaneo, and Rocio Titiunik (2014). “Robust nonparametric confidence intervals for regression-discontinuity designs.” *Econometrica*, 82(6), 2295–2326.

Camerer, Colin F (1998). “Prospect theory in the wild: Evidence from the field.” Tech. rep., California Institute of Technology.

Carpenter, Jeffrey and Caitlin Knowles Myers (2010). “Why volunteer? Evidence on the role of altruism, image, and incentives.” *Journal of Public Economics*, 94(11-12), 911–920.

Corgnet, Brice, Roberto Hernán-González, and Eric Schniter (2015). “Why real leisure really matters: Incentive effects on real effort in the laboratory.” *Experimental Economics*, 18(2), 284–301.

Dai, Zhixin, Fabio Galeotti, and Marie Claire Villeval (2018). “Cheating in the lab predicts fraud in the field: An experiment in public transportation.” *Management Science*, 64(3), 1081–1100.

Eckartz, Katharina, Oliver Kirchkamp, and Daniel Schunk (2012). “How do incentives affect creativity?” *Jena Economic Research Papers, 6(68).*

Ewers, Mara and Florian Zimmermann (2015). “Image and misreporting.” *Journal of the European Economic Association*, 13(2), 363–380.

Falk, Armin and Nora Szech (2020). “Competing Image Concerns: Pleasures of Skills and Moral Values.” Working paper.

Fehr, Ernst and Lorenz Goette (2007). “Do workers work more if wages are high? Evidence from a randomized field experiment.” *American Economic Review*, 97(1), 298–317.

Fischbacher, Urs (2007). “z-Tree: Zurich toolbox for ready-made economic experiments.” *Experimental Economics*, 10(2), 171–178.

Fischbacher, Urs and Franziska Föllmi-Heusi (2013). “Lies in disguise—an experimental study on cheating.” *Journal of the European Economic Association*, 11(3), 525–547.

Friedrichsen, Jana and Dirk Engelmann (2018). “Who cares about social image?” *European Economic Review*, 110, 61–77.

Gächter, Simon, Lingbo Huang, and Martin Sefton (2016). “Combining “real effort” with induced effort costs: the ball-catching task.” *Experimental Economics*, 19(4), 687–712.

Gächter, Simon, Eric J Johnson, and Andreas Herrmann (2007). “Individual-level loss aversion in riskless and risky choices.” *IZA Discussion Paper.*

Garbarino, Ellen, Robert Slonim, and Marie Claire Villeval (2019). “Loss aversion and lying behavior.” *Journal of Economic Behavior and Organization*, 158, 379–393.

Gibson, Rajna, Carmen Tanner, and Alexander F Wagner (2013). “Preferences for truthfulness: Heterogeneity among and within individuals.” *American Economic Review*, 103(1), 532–48.

Gneezy, Uri, Agne Kajackaite, and Joel Sobel (2018). “Lying Aversion and the Size of the Lie.” *American Economic Review*, 108(2), 419–53.

Goerg, Sebastian J, Sebastian Kube, and Jonas Radbruch (2019). “The effectiveness of incentive schemes in the presence of implicit effort costs.” *Management Science*, 65(9), 4063–4078.

Greiner, Ben (2015). “Subject Pool Recruitment Procedures: Organizing Experiments with ORSEE.” *Journal of the Economic Science Association*, 1(1), 114–125.

Grolleau, Gilles, Martin G Kocher, and Angela Sutan (2016). “Cheating and loss aversion: do people cheat more to avoid a loss?” *Management Science*, 62(12), 3428–3438.

Grubiak, Kevin (2019). “Exploring Image Motivation in Promise Keeping: An Experimental Investigation.” Tech. rep., School of Economics, University of East Anglia, Norwich, UK.

Halleck, Rebecca (2019). “Who’s Been Charged in the College Admissions Cheating Scandal? Here’s the Full List.” *N.Y. Times, 2019, March 12*.

Hanna, Rema and Shing-Yi Wang (2017). “Dishonesty and Selection into Public Service: Evidence from India.” *American Economic Journal: Economic Policy*, 9(3), 262–290.

Kahneman, Daniel, Jack L Knetsch, and Richard H Thaler (1990). “Experimental tests of the endowment effect and the Coase theorem.” *Journal of Political Economy*, 98(6), 1325–1348.

Kahneman, Daniel, Jack L Knetsch, and Richard H Thaler (1991). “Anomalies: The endowment effect, loss aversion, and status quo bias.” *Journal of Economic Perspectives*, 5(1), 193–206.

Kahneman, Daniel and Amos Tversky (1979). “Prospect Theory: An Analysis of Decision under Risk.” *Econometrica*, 47(2), 263–292.

Khalmetski, Kiryl and Dirk Sliwka (2019). “Disguising lies—Image concerns and partial lying in cheating games.” *American Economic Journal: Microeconomics*, 11(4), 79–110.

Leibenstein, Harvey (1950). “Bandwagon, snob, and Veblen effects in the theory of consumers’ demand.” *The Quarterly Journal of Economics*, 64(2), 183–207.

Lovett, Benjamin J (2020). “Disability Identification and Educational Accommodations: Lessons From the 2019 Admissions Scandal.” *Educational Researcher*, 49(2), 125–129.

Pennings, Joost ME and Ale Smidts (2003). “The shape of utility functions and organizational behavior.” *Management Science*, 49(9), 1251–1263.

Raven, John C (1983). “Manual for Raven’s progressive matrices and vocabulary scales.” *Standard Progressive Matrices*.

Schindler, Simon and Stefan Pfattheicher (2017). “The frame of the game: Loss-framing increases dishonest behavior.” *Journal of Experimental Social Psychology*, 69, 172–177.

Sexton, Steven E and Alison L Sexton (2014). “Conspicuous conservation: The Prius halo and willingness to pay for environmental bona fides.” *Journal of Environmental Economics and Management*, 67(3), 303–317.

Soetevent, Adriaan R (2011). “Payment choice, image motivation and contributions to charity: Evidence from a field experiment.” *American Economic Journal: Economic Policy*, 3(1), 180–205.

Wakker, Peter P (2010). *Prospect theory: For risk and ambiguity*. Cambridge University Press.

*All material from here on should go into an online appendix.*

## Appendix A: Proofs

PROPOSITION A.1. ***Under the assumption of differentiable lying costs, individuals never report truthfully unless*** $\omega_i = \overline{\omega}$.

*Proof.*

The agent faces a trade-off. On the one hand, she can report $r_i = \omega_i$ which results in zero lying costs, but no benefit regarding the social image component and value function. On the other hand, she can report $r_i > \omega_i$ which is going to increase her social image and the value function but imposes lying costs. If the lying cost function is convex with a minimum at $r_i = \omega_i$, that implies

$$\left. \frac{dc(\omega_i, r_i)}{dr_i} \right|_{r_i = \omega_i} = 0.$$

Therefore no one reports truthfully if $\omega_i < \overline{\omega}$ because the first order condition should be set to zero in order to maximize utility, which is only possible if $r_i > \omega_i$. □

# Appendix B: Robustness checks

## Social image split

In our main regression analysis in Table 1, we split the sample according to the median of the importance of social image concerns that subjects reported on a 11-point Likert-scale. Only individuals who reported 6 or higher are categorized as "Social image concerns". In the following, we analyze how robust our findings are to alterations of this threshold.

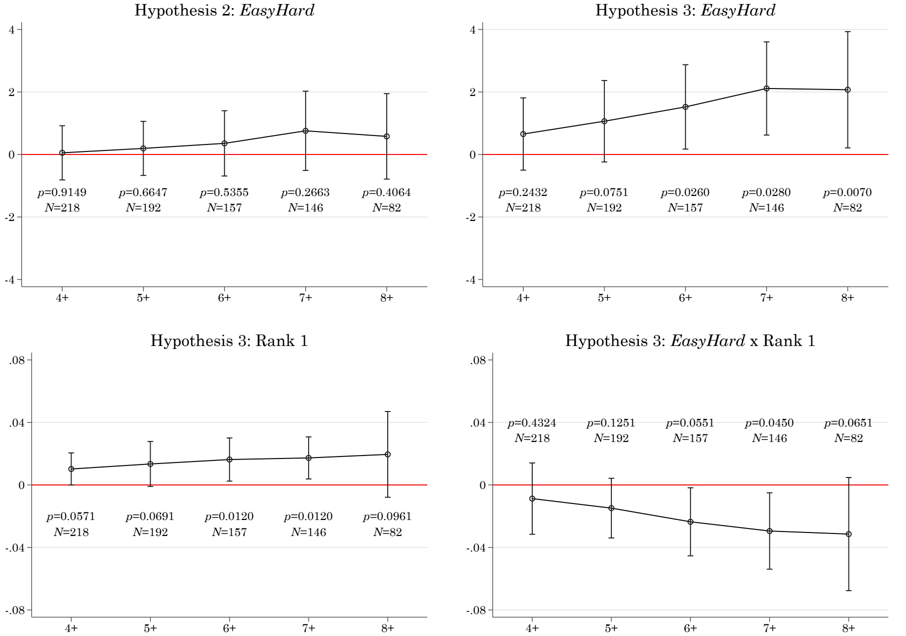



FIGURE B.1. Robustness check: varying the threshold for social image concerns

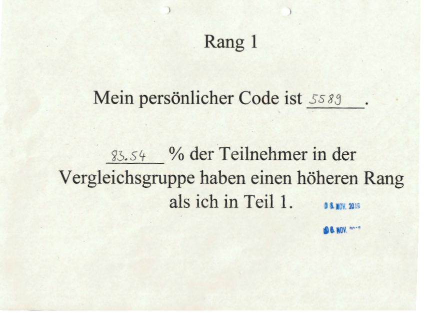
*Note:* The vertical axis displays the magnitude of each respective coefficient along with a 95% confidence interval. Number of observations and score bootstrap *p*-values correspond to each estimated regression. The horizontal axis indicates the sub-sample used for the estimation. "4+" ("5+"/"6+"/"7+"/"8+") includes subjects who reported the importance of social image to be 4 (5/6/7/8) or higher. Each reported coefficient and its 95% confidence interval is estimated using two-limit Tobit models with standard errors clustered at the session level. We label a robustness check "Hypothesis 2" if its underlying regression corresponds to columns (1) and (3) in Table 1 and "Hypothesis 3" if it follows the specification of columns (2) and (4).

Figure B.1 shows the coefficients for the treatment dummy $EasyHard$, Rank 1 and their interaction for an increasingly restricted sample. We indicate the sample as "4+" if it includes subjects who reported the importance of social image to be 4 or higher, "5+" for those who reported it to be 5 or higher, etc. The "8+" sample includes only subjects who are extremely concerned with their social image and reported its importance to be 8, 9 or 10.[21] For each of those sub-samples, we estimate the same two-limit Tobit models with standard errors clustered at the session level as reported in Table 1. We label a robustness check "Hypothesis 2" if its underlying regression corresponds to columns (1) and (3) in Table 1 testing for unconditional loss aversion in social image concerns, and "Hypothesis 3" if it follows the specification of columns (2) and (4) and tests for loss aversion in social image concerns conditional on one's reference point represented by Rank 1.

Our findings are robust for a variety of sub-samples. In particular, we do not find significant evidence for unconditional loss aversion irrespective of the threshold we apply for social image concerns (see upper left panel of Figure B.1). However, conditional loss aversion is strongly supported by our robustness checks: subjects who have a high reference point (low Rank 1) lie significantly more on average in treatment $EasyHard$ than in $HardEasy$, and the effect gets more and more pronounced for stronger image concerns (see upper right panel of Figure B.1). Expanding our sample by adding subjects who report the importance of social image to be 7 and 6 as opposed to 8 or higher gradually reduces the average treatment effect (which, however, stays high and significant). After including those with image concerns of 5 and 4, the effect remains positive but becomes smaller and insignificant as we have previously documented for our sample as a whole. The estimated coefficient of Rank 1 remains rather stable for the various thresholds (see lower left panel of Figure B.1). Finally, the estimated coefficient of the interaction term becomes more negative the stronger social image concerns are (see lower right panel of Figure B.1), gradually offsetting the larger level effect of $EasyHard$ that is displayed in the upper right panel of Figure B.1.

21. Note that we do not consider sub-samples of those who reported the importance of social image to be 9 or higher, or 10 since only 30 subjects in our data reported the importance of social image to be 9 or 10.

## Appendix C: Additional Figures


Rang 1

Mein persönlicher Code ist 5589.

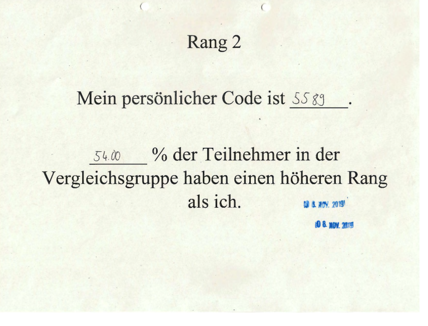
83.54 % der Teilnehmer in der Vergleichsgruppe haben einen höheren Rang als ich in Teil 1.

Rank 1

My individual code is ______.

______ % of the participants of the reference group have a higher rank than me in part 1.

FIGURE C.1. Rank 1 report sheet (original in German and translated to English)


Rang 2

Mein persönlicher Code ist 5589.

54.00 % der Teilnehmer in der Vergleichsgruppe haben einen höheren Rang als ich.

Rank 2

My individual code is ______.

______ % of the participants of the reference group have a higher rank than me.

FIGURE C.2. Rank 2 report sheet (original in German and translated to English)

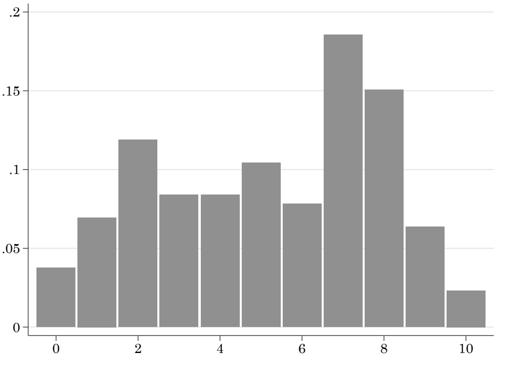



FIGURE C.3. Self-reported importance of social image

Note: Importance of social image concerns is measured on a 11-point Likert scale based on the question "How important is the opinion that others hold about you to you?".

## Appendix D: Instructions of the Experiment

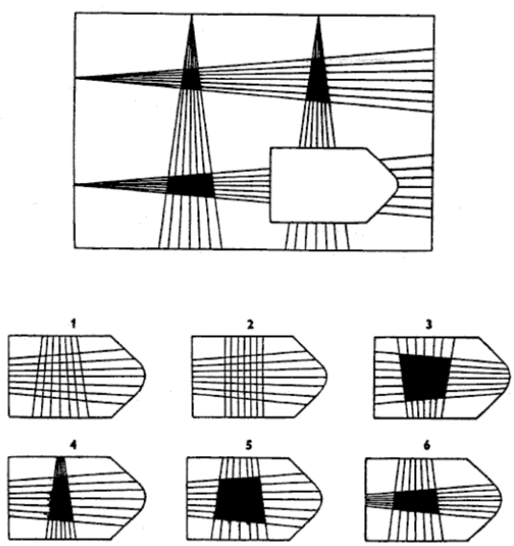
### D.1. English

**General Instructions**

We warmly welcome you to this economic experiment. Please read the following instructions carefully! If you have any questions, please raise your hand from the cubicle—we will then come to your seat. It is not allowed to talk to other participants of the experiment, use mobile phones or start other programs on the computer during the experiment. Non-compliance with these rules will result in exclusion from the experiment and all payments. You will receive a fixed payment of €12 for participating in this experiment, which will be paid in cash at the end of the experiment. On the following pages we describe the exact procedure of the experiment.

**Part 1 of the Experiment**

Parts 1 and 2 consist each of 24 tasks, which are often used to measure so-called fluid intelligence of a person. The fluid intelligence is an important part of the general intelligence of humans. These or similar tasks are also often used by companies in the context of recruitment procedures. Each task corresponds to a picture puzzle. Here you can see an example:




Each picture puzzle shows in its upper part a pattern in a box, in which a "piece of the puzzle" in the lower right corner is left out. Your task is to select one of the puzzle pieces listed below the box, which will logically fill the blank lower right corner of the pattern in the box. Please enter the number of the puzzle piece that you think fits best on the screen. The number of a puzzle piece is stated above each puzzle piece. There is always exactly one piece that fits best.

You have 30 seconds to complete each picture puzzle. For each correctly completed picture puzzle you receive one point. As commonly done with intelligence tests, correct answers are not paid extra. You will receive 0 points for each wrongly answered picture puzzle or if you do not enter the best fitting piece of the puzzle within 30 seconds.

After you have completed all 24 picture puzzles in Part 1, you will first receive a private feedback on your rank on the computer screen, indicating how well you performed in solving the picture puzzles. The feedback has the following form: "X % of the participants of the reference group have a higher rank than you in Part 1". The reference group consists of 413 participants of a previous laboratory experiment conducted in 2014 here at the DICE Lab of the University of Düsseldorf, who have worked on the same picture puzzles as you do in the course of this experiment. So the feedback "9% of the participants of the reference group have a higher rank than you in Part 1" means that 9% performed better than you (i.e. solved a higher percentage of the total 48 picture puzzles from Parts 1 and 2 correctly than you) and 90% performed worse (i.e. solved a lower percentage of picture puzzles correctly than you). So you belong to the 10% of the best at answering the picture puzzles designed to measure individual fluid intelligence. The feedback "83% of the participants of the reference group have a higher rank than you in Part 1" means that 83% performed better and 16% worse than you. So you are among the 17% of the worst in answering the picture puzzles.

Before Part 2 of the experiment starts, you have to inform two so-called "Observers" about your performance in the experiment. Please use the report sheet available in your cubicle. Your cubicle number is already entered. Please enter legibly the number, which you received as feedback on the computer screen, in the sentence "__ % of the participants of the reference group have a higher rank than me in Part 1" in the report sheet "Rank 1". Please enter your personal code, which is also displayed on the screen, in the free field next to it: "My personal code is __". Observers sit in the cubicles number 1 and 2 in the laboratory (directly in front of the entrance door). Please go there with the completed report sheet and show it silently to Observers as soon as your cubicle number is called by the experimenter. This ensures that each participant informs Observers individually without other participants knowing her/his rank. A two-column table will be displayed on the Observers' computer screens, assigning each personal code the corresponding rank in Part 1. Each Observer will silently compare your report sheet with the information in the table and stamp it. Afterwards, please return to your cubicle in silence. Part 2 of the experiment will begin as soon as all participants have informed Observers of their rank.

**The Different Participants in the Experiment**

At the beginning of the experiment, each participant randomly drew a chip with a number indicating his cubicle number. The cubicle numbers have the following additional meaning: The participants who have randomly drawn cubicle numbers 1 and 2 have the role of "Observers" described above. Since the chips with even numbers were reserved for female participants and the chips with odd numbers for male participants, there is always one male Observer and one female Observer. These will introduce themselves to you shortly before the actual experiment begins by standing up and saying "I am one of the two Observers". Observers—just like all other participants—will receive this printed explanation of the rules of the experiment, which you are reading, for information about the experiment.

All other participants in the experiment with cabin numbers 3 or higher solve the picture puzzles described above. Each participant is randomly assigned to one of two groups: Group A or Group B. Throughout the whole experiment, all participants of both groups will solve exactly the same 48 picture puzzles, 24 in Part 1 and 24 in Part 2. The further task in part 2 of the experiment is also exactly the same for both groups. Only the order in which the picture puzzles are processed differs between group A and B. The group membership has no further meaning. In Parts 1 and 2 you belong to the same group.

**Part 2 of the Experiment**

Part 2 of the experiment is very similar to Part 1. First you work on 24 more picture puzzles following the same rules (30 seconds time per puzzle, 1 point for correct answers, 0 points otherwise, etc.). After you have completed remaining 24 picture puzzles in Part 2, you will receive a private feedback on your preliminary rank in Part 2 on the computer screen, indicating how well you have done in the 48 picture puzzles in Parts 1 and 2. The feedback again has the following form: "X% of the participants of the reference group have a higher rank than you". The reference group is again the 413 participants of a previous lab experiment here in the DICE Lab of the HHU from 2014, who have solved the same 48 picture puzzles as you. In addition, the rank you had in Part 1 of the experiment is displayed as a reminder.

The only difference to Part 1 is that you have one more task, which is also used to calculate your final rank in Part 2. You will then receive a private feedback on your final rank in Part 2, which is calculated based on the 48 picture puzzles in Parts 1 and 2 and your score in the further task in Part 2. Details of the further task and how exactly it is included in the calculation of the final rank in Part 2 will be explained on the computer screen during the course of the experiment. For calculation of your final rank the same reference group is used again as for your rank in Part 1 and the preliminary rank in Part 2. The detailed explanations of the further task in Part 2 are given only to the participants, but not to the two Observers.

Just like at the end of Part 1, you still have to inform the two Observers about your performance, i.e. your final rank, in Part 2. Please use the report sheet which is available in your cubicle. In addition, under "Rank 2", please enter legibly in the sentence "__ % of the participants of the reference group have a higher rank than me", which you have received as feedback on your final rank on the computer screen. Please go to two Observers with the completed report sheet and show it to them in silence as soon as your cubicle number is called up by an experimenter. This again ensures that each participant informs Observers individually without the other participants knowing her/his rank. A table with four columns is now displayed to Observers on your computer screen, which assigns to each personal code the corresponding rank in Part 1, the final rank and the difference in rank between the rank in Part 1 and the final rank.

The observers will, also in silence, compare your report sheet with the information in the table and stamp it. Afterwards, please return to your cabin in silence.

**End and Payment of the Experiment**

After Part 2 of today's experiment, there will be some more screens with questions before we proceed to the payment of €12. We will call you individually by cubicle number for payment. If you have any questions now, please raise your hand out of the cubicle. Experiment supervisor will then come to your seat to answer your questions. Do not ask questions out loud!

### D.2. German (original)

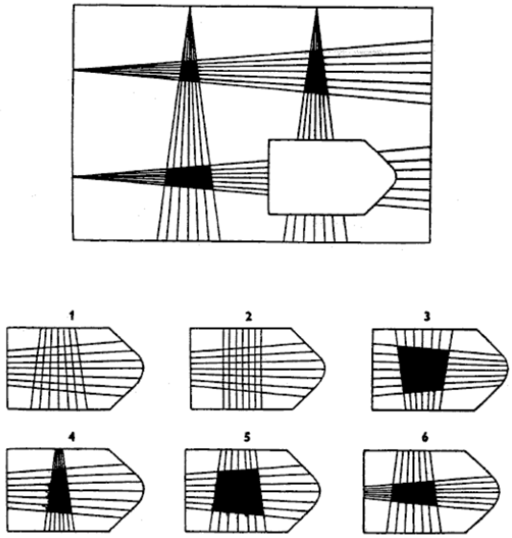
**Allgemeine Erklärungen**

Wir begrüßen Sie herzlich zu diesem wirtschaftswissenschaftlichen Experiment. Lesen Sie die folgenden Erklärungen bitte gründlich durch! Wenn Sie Fragen haben, strecken Sie bitte Ihre Hand aus der Kabine – wir kommen dann zu Ihrem Platz. Während des Experiments ist es nicht erlaubt, mit den anderen Experimentteilnehmern zu sprechen, Mobiltelefone zu benutzen oder andere Programme auf dem Computer zu starten. Die Nichtbeachtung dieser Regeln führt zum Ausschluss aus dem Experiment und von allen Zahlungen. Für die Teilnahme an diesem Experiment erhalten Sie pauschal 12 Euro, die Sie am Ende dieses Experiments bar ausbezahlt bekommen. Auf den nächsten Seiten beschreiben wir den genauen Ablauf des Experiments.

**Teil 1 des Experiments**

In Teil 1 und 2 bearbeiten Sie jeweils 24 Aufgaben, die oft verwendet werden, um die sogenannte fluide Intelligenz eines Menschen zu bestimmen. Die fluide Intelligenz ist ein wichtiger Bestandteil der allgemeinen Intelligenz des Menschen. Oft werden solche oder ähnliche Aufgaben auch im Rahmen von Einstellungsverfahren von Unternehmen verwendet. Jede Aufgabe entspricht einem Bilderrätsel. Hier sehen Sie ein Beispiel:

Jedes Bilderrätsel zeigt in seinem oberen Teil ein Muster in einem Kasten, in dem unten rechts ein "Puzzlestück" ausgelassen ist. Ihre Aufgabe ist es, eines der unterhalb des Kastens aufgeführten Puzzlestücke auszuwählen, das die leere, untere rechte Ecke des Musters im Kasten logisch passend füllt. Bitte geben Sie dazu die Nummer des Puzzlestücks, das Ihrer Meinung nach am besten passt, auf dem Bildschirm ein. Die Nummer eines Puzzlestücks steht




oberhalb jedes Puzzlestücks. Es gibt immer genau ein am besten passendes Puzzlestück.

Für die Bearbeitung eines Bilderrätsels haben Sie jeweils 30 Sekunden Zeit. Für jedes richtig beantwortete Bilderrätsel erhalten Sie einen Punkt. Wie dies bei Intelligenztests üblich ist, werden richtige Antworten nicht extra bezahlt. Sie erhalten 0 Punkte für jedes falsch beantwortete Bilderrätsel oder falls Sie innerhalb der 30 Sekunden keine Eingabe zum Ihrer Meinung nach am besten passenden Puzzlestück machen. Nachdem Sie alle 24 Bilderrätsel in Teil 1 bearbeitet haben, erhalten Sie auf dem Computerbildschirm zunächst ein privates Feedback zu Ihrem Rang, der angibt, wie gut Sie bei den Bilderrätseln abgeschnitten haben. Das Feedback hat die folgende Form: “X % der Teilnehmer in der Vergleichsgruppe haben einen höheren Rang als Sie in Teil 1”. Die Vergleichsgruppe sind dabei 413 Teilnehmer an einem vorherigen Laborexperiment hier im DICE Lab der HHU aus dem Jahr 2014, die dieselben Bilderrätsel bearbeitet haben, wie Sie es im Laufe dieses Experiments tun. Das Feedback “9 % der Teilnehmer in der Vergleichsgruppe haben einen höheren Rang als Sie in Teil 1” bedeutet also, dass 9 % besser abschneiden als Sie (d.h. einen höheren Anteil der gesamten 48 Bilderrätsel aus Teil 1 und 2 korrekt gelöst haben als Sie) und 90 % schlechter (d.h. einen niedrigen Anteil an Bilderrätseln korrekt gelöst haben als Sie). Sie gehören also zu den 10 % der Besten beim Beantworten der Bilderrätsel, die konzipiert wurden, um die individuelle fluide Intelligenz zu messen. Das Feedback “83 % der Teilnehmer in der Vergleichsgruppe haben einen höheren Rang als Sie in Teil 1” bedeutet, dass 83 % besser abschneiden als Sie und 16 % schlechter. Sie gehören also zu den 17 % der Schlechtesten beim Beantworten der Bilderrätsel.

Bevor Teil 2 des Experiments beginnt, müssen Sie noch zwei sogenannte "Beobachter" über Ihr Abschneiden im Experiment informieren. Bitte verwenden Sie dazu das DIN-A4-Blatt, das in Ihrer Kabine bereitliegt. Ihre Kabinennummer ist bereits eingetragen. Bitte tragen Sie unter "Rang 1" gut leserlich die Zahl in den Satz ein "__ % der Teilnehmer in der Vergleichsgruppe haben einen höheren Rang als ich in Teil 1", die Sie als Feedback auf dem Computerbildschirm erhalten haben. Tragen Sie bitte Ihren persönlichen Code, der ebenfalls auf dem Bildschirm angezeigt wird, daneben in das freie Feld ein: "Mein persönlicher Code ist __". Die Beobachter sitzen in den Kabinen mit Nummer 1 und 2 im Labor (direkt gegenüber der Eingangstür). Bitte gehen Sie mit dem ausgefüllten DIN-A4-Blatt dorthin und zeigen es schweigend den Beobachtern, sobald Ihre Kabinennummer vom Experimentator aufgerufen wird. So wird sichergestellt, dass jeder Teilnehmer die Beobachter einzeln informiert, ohne dass die anderen Teilnehmer seinen Rang erfahren. Den Beobachtern wird auf ihrem Computerbildschirm eine Tabelle mit zwei Spalten angezeigt, die jedem persönlichen Code den entsprechenden Rang in Teil 1 zuordnet. Beide Beobachter werden, ebenfalls schweigend, Ihr DIN-A4-Blatt mit den Angaben in ihrer Tabelle vergleichen und jeweils abstempeln. Bitte begeben Sie sich dann schweigend wieder zurück in Ihre Kabine. Teil 2 des Experiments beginnt, sobald alle Teilnehmer die Beobachter über ihren Rang informiert haben.

**Die verschiedenen Teilnehmer am Experiment**

Zu Beginn des Experiments hat jeder Teilnehmer zufällig einen Chip mit einer Zahl gezogen, die seine Kabinennummer angibt. Die Kabinennummern haben folgende weitere Bedeutung: Die Teilnehmer, die zufällig die Kabinennummern 1 und 2 gezogen haben, haben die Rolle der oben beschriebenen "Beobachter". Da die Chips mit den geraden Zahlen für die Frauen und die Chips mit den ungeraden Zahlen für die Männer reserviert waren, gibt es immer jeweils einen männlichen Beobachter und eine weibliche Beobachterin. Diese werden sich vor Beginn des eigentlichen Experiments kurz bei Ihnen vorstellen, in dem sie aufstehen und sagen "Ich bin eine/r der beiden Beobachter". Die Beobachter erhalten—genau wie die anderen Teilnehmer—diese ausgedruckte Erklärung der Regeln des Experiments, die Sie gerade lesen, zur Information über das Experiment.

Alle anderen Teilnehmer am Experiment mit den Kabinennummern 3 oder höher lösen die oben beschriebenen Bilderrätsel. Dabei wird jeder Teilnehmer zufällig einer von zwei Gruppen zugelost: Gruppe A oder Gruppe B. Im Laufe des gesamten Experiments bearbeiten alle Teilnehmer beider Gruppen exakt dieselben 48 Bilderrätsel, jeweils 24 in Teil 1 und 24 in Teil 2. Auch die weitere Aufgabe in Teil 2 des Experiments ist exakt dieselbe für beide Gruppen. Nur die Reihenfolge, in der die Bilderrätsel bearbeitet werden, unterscheidet sich zwischen Gruppe A und B. Eine weitere Bedeutung hat die Gruppenzugehörigkeit nicht. In Teil 1 und 2 gehören Sie zu derselben Gruppe.

**Teil 2 des Experiments**

Teil 2 des Experiments ist Teil 1 sehr ähnlich. Zunächst bearbeiten Sie 24 weitere Bilderrätsel nach denselben Regeln (30 Sekunden Zeit pro Rätsel, 1 Punkt für richtige Antworten, 0 Punkte sonst etc.). Nachdem Sie die weiteren 24 Bilderrätsel in Teil 2 bearbeitet haben, erhalten Sie auf dem Computerbildschirm zunächst ein privates Feedback zu Ihrem vorläufigen Rang in Teil 2, der angibt, wie gut Sie bei den insgesamt 48 Bilderrätseln in Teil 1 und 2 abgeschnitten haben. Das Feedback hat wieder die folgende Form: "X % der Teilnehmer in der Vergleichsgruppe haben einen höheren Rang als Sie." Die Vergleichsgruppe sind dabei wieder die 413 Teilnehmer an einem vorherigen Laborexperiment hier im DICE Lab der HHU aus dem Jahr 2014, die dieselben 48 Bilderrätsel bearbeitet haben wie Sie. Außerdem wird zur Erinnerung angezeigt, welchen Rang Sie in Teil 1 des Experiments hatten.

Der einzige Unterschied zu Teil 1 ist, dass Sie eine weitere Aufgabe haben, die auch in die Berechnung Ihres finalen Rangs in Teil 2 einfließt. Anschließend erhalten Sie ein privates Feedback zu Ihrem finalen Rang in Teil 2, der auf Grundlage der 48 Bilderrätsel in Teil 1 und 2 und Ihrem Abschneiden in der weiteren Aufgabe in Teil 2 berechnet wird. Details zur weiteren Aufgabe und wie genau sie in die Berechnung des finalen Rangs in Teil 2 einfließt, werden im Verlauf des Experiments auf dem Computerbildschirm erklärt. Zur Berechnung Ihres finalen Rangs wird wieder dieselbe Vergleichsgruppe herangezogen wie für Ihren Rang in Teil 1 und den vorläufigen Rang in Teil 2. Die detaillierten Erklärungen zur weiteren Aufgabe in Teil 2 erhalten nur die Teilnehmer, aber nicht die beiden Beobachter.

Genau wie zum Abschluss von Teil 1 müssen Sie noch die zwei Beobachter über Ihr Abschneiden, also Ihren finalen Rang, in Teil 2 informieren. Bitte verwenden Sie dazu wieder das DIN-A4-Blatt, das in Ihrer Kabine bereitliegt. Bitte tragen Sie nun zusätzlich unter "Rang 2" gut leserlich die Zahl in den Satz ein "__ % der Teilnehmer in der Vergleichsgruppe haben einen höheren Rang als ich", die Sie als Feedback über Ihren finalen Rang auf dem Computerbildschirm erhalten haben. Bitte gehen Sie mit dem ausgefüllten DIN-A4-Blatt zu den beiden Beobachtern und zeigen es ihnen schweigend, sobald Ihre Kabinennummer von einem Experimentator aufgerufen wird. So wird wieder sichergestellt, dass jeder Teilnehmer die Beobachter einzeln informiert, ohne dass die anderen Teilnehmer seinen Rang erfahren. Den Beobachtern wird auf ihrem Computerbildschirm nun eine Tabelle mit vier Spalten angezeigt, die jedem persönlichen Code den entsprechenden Rang in Teil 1, den finalen Rang sowie die Rangdifferenz zwischen Rang in Teil 1 und dem finalen Rang zuordnet.

Die Beobachter werden, ebenfalls schweigend, Ihr DIN-A4-Blatt mit den Angaben in ihrer Tabelle vergleichen und abstempeln. Bitte begeben Sie sich dann schweigend wieder zurück in Ihre Kabine.

**Ende und Auszahlung des Experiments**

Nach Teil 2 des heutigen Experiments folgen dann noch einige Bildschirme mit Fragen u. Ä., bevor wir zur Auszahlung der 12 Euro kommen. Wir werden Sie einzeln nach Kabinennummer zur Auszahlung aufrufen.

Falls Sie jetzt Fragen haben, halten Sie bitte die Hand aus der Kabine. Ein Leiter des Experiments wird dann an Ihren Platz kommen, um Ihre Fragen zu beantworten. Stellen Sie Fragen keinesfalls laut!

## D.3. Additional Instructions on the Computer Screen: Die Roll Task

*D.3.1. English.* There is a die in your cubicle. Please roll the die twice in your cubicle.

Please enter the numbers between 1 and 6 which you rolled on the first and second die rolls on the computer screen. The first number you rolled is added to the number of correctly solved picture puzzles of each participant of the reference group. The second die roll is added to your own number of correctly solved picture puzzles. Your total output is equal to the number of correctly solved picture puzzles in Parts 1 and 2 of the 48 picture puzzles plus the number of points you entered for the second die roll. Your total output is used to calculate your final rank. Your total output is compared with the total output of the peer group. The total output of a participant in the comparison group is equal to the number of correctly solved picture puzzles out of the 48 picture puzzles plus the number of points you entered for the first die roll. Your final rank will be shown to Observers and you will report it to the Observers at the end.

You may, of course, roll the die more often, for example to check that the die is working properly. If you have thrown more than twice, the other throws after the first two do not have any special meaning.

*D.3.2. German (original).* In Ihrer Kabine liegt ein Würfel bereit. Bitte würfeln Sie zwei Mal in Ihrer Kabine.

Bitte geben Sie dann auf dem Computerbildschirm ein, welche Augenzahl zwischen 1 und 6 Sie beim ersten und zweiten Wurf gewürfelt haben. Die erste gewürfelte Augenzahl wird zur Anzahl der korrekt gelösten Bilderrätsel jedes Teilnehmers in der Vergleichsgruppe dazu gezählt. Die zweite gewürfelte Augenzahl wird zur Anzahl der von Ihnen korrekt gelösten Bilderrätsel dazu gezählt. Ihre entstehende Gesamtleistung entspricht also der Anzahl der von Ihnen korrekt gelösten Bilderrätsel in Teil 1 und 2 von den insgesamt 48 Bilderrätseln plus der von Ihnen eingegebenen Augenzahl vom zweiten Würfelwurf. Ihre Gesamtleistung wird verwendet, um Ihren finalen Rang zu berechnen. Dabei wird Ihre Gesamtleistung mit der Gesamtleistung der Vergleichsgruppe verglichen. Die Gesamtleistung eines Teilnehmers der Vergleichsgruppe entspricht der Anzahl der von ihm / ihr korrekt gelösten Bilderrätsel von den insgesamt 48 Bilderrätseln plus die von Ihnen eingegebene Augenzahl vom ersten Würfelwurf. Ihr finaler Rang wird den Beobachtern angezeigt und Sie werden ihn den Beobachtern abschließend berichten.

Natürlich können Sie gerne auch häufiger würfeln, z.B. um festzustellen, dass der Würfel richtig funktioniert. Falls Sie häufiger als zwei Mal gewürfelt haben, haben die weiteren Würfe nach den ersten beiden keine besondere Bedeutung.

## Appendix E: Questionnaire

### E.1. English

Please fill out the following questionnaire now before we proceed to the payment. Please enter the following personal data. If you want to enter decimal numbers, please use a dot (.) instead of a comma (,).

- Age
- Gender (male/female)
- Final grade point average at high school (Abiturnote) (1.0–6.0)
- Last math grade (1.0–6.0)
- Last German grade (1.0–6.0)
- Field of study/job
- How much money do you have available each month (after deducting fixed costs such as rent, insurance, etc.)?
- How much money do you spend each month (after deducting fixed costs such as rent, insurance, etc.)?
- In how many economic science experiments have you (approximately) already participated?
- On a scale of 0 to 10, how would you rate your willingness to take risks? 0 means not willing to take risks at all and 10 means completely willing to take risks.
- How important is the opinion that others hold about you to you? Please answer on a scale of 0 to 10, where 0 is not important at all and 10 is extremely important.
- Have you ever solved similar tasks as the picture puzzles before? (Yes/No)
- If so, how long ago approximately? Please indicate the approximate number of months.

Below, please answer a few more questions about lotteries in which you can earn or lose money in addition to the €12 if you decide to accept the lotteries.

Listed below are 6 different lotteries. For each of the 6 lotteries you can choose whether to accept or decline the lottery. If you choose to decline a lottery, your payout will not change. If you accept a lottery, you will realize either an additional gain or an additional loss based on the €12.

At the end of the experiment, one of the 6 lotteries is randomly selected. So you should make each lottery decision as if it was your only decision. The selected lottery is then drawn to determine whether the additional gain or loss will be realized.

Lottery 1: With 50% probability you lose €2 and with 50% probability you win €6. (accept / reject)
Lottery 2: With 50% probability you lose €3 and with 50% probability you win €6. (accept / reject)

Lottery 3: With 50% probability you lose €4 and with 50% probability you win €6. (accept / reject)
Lottery 4: With 50% probability you lose €5 and with 50% probability you win €6. (accept / reject)
Lottery 5: With 50% probability you lose €6 and with 50% probability you win €6. (accept / reject)
Lottery 6: With 50% probability you lose €7 and with 50% probability you win €6. (accept / reject)

**E.2. German (original)**

Füllen Sie nun bitte die folgenden Fragen aus, bevor wir zur Auszahlung kommen. Bitte geben Sie die folgenden Daten zu Ihrer Person an. Wenn Sie Kommazahlen eingeben möchten, nutzen Sie bitte einen Punkt (.) statt eines Kommas (,).

- Alter
- Geschlecht (männlich/weiblich)
- Abiturdurchschnittsnote (1.0-6.0)
- Letzte Mathenote (1.0-6.0)
- Letzte Deutschnote (1.0-6.0)
- Studienfach/Tätigkeit
- Wie viel Geld haben Sie monatlich (nach Abzug von Fixkosten wie Miete, Versicherungen etc.) zur Verfügung?
- Wie viel Geld geben Sie monatlich aus (nach Abzug von Fixkosten wie Miete, Versicherungen etc.)?
- An wie vielen wirtschaftswissenschaftlichen Experimenten haben Sie (ungefähr) bereits teilgenommen?
- Wie schätzen Sie Ihre Risikobereitschaft auf einer Skala von 0 bis 10 ein? Dabei bedeutet 0 überhaupt nicht risikobereit und 10 vollkommen risikofreudig.
- Wie wichtig ist Ihnen die Meinung, die andere über Sie haben? Bitte antworten Sie auf einer Skala 0 bis 10. Dabei ist 0 überhaupt nicht wichtig und 10 extrem wichtig.
- Haben Sie schon einmal ähnliche Aufgaben wie die Bilderrätsel gelöst? (Ja/Nein)
- Falls ja, wie lange ist das ungefähr her? Bitte geben Sie die ungefähre Zahl der Monate an.

Im Folgenden beantworten Sie bitte noch ein paar Fragen zu Lotterien, bei denen Sie noch einmal zusätzlich zu den €12 Geld verdienen oder auch verlieren können, falls Sie sich entscheiden, die Lotterien zu akzeptieren.

Unten sind 6 verschiedene Lotterien aufgelistet. Sie können für jede der 6 Lotterien wählen, ob Sie die Lotterie akzeptieren oder ablehnen möchten. Falls Sie eine Lotterie ablehnen, bleibt Ihre Auszahlung unverändert. Falls Sie eine Lotterie akzeptieren, werden Sie ausgehend von den €12 entweder einen zusätzlichen Gewinn oder einen zusätzlichen Verlust realisieren.

Am Ende des Experiments wird zufällig eine der 6 Lotterien ausgewählt. Sie sollten also jede Lotterieentscheidung so fallen, als wäre es Ihre einzige Entscheidung. Die ausgewählte Lotterie wird anschließend ausgelost, damit feststeht, ob sich der zusätzliche Gewinn oder Verlust realisiert.

Lotterie 1: Mit 50% Wahrscheinlichkeit verlieren Sie €2 und mit 50% Wahrscheinlichkeit gewinnen Sie €6. (akzeptieren / ablehnen)
Lotterie 2: Mit 50% Wahrscheinlichkeit verlieren Sie €3 und mit 50% Wahrscheinlichkeit gewinnen Sie €6. (akzeptieren / ablehnen)
Lotterie 3: Mit 50% Wahrscheinlichkeit verlieren Sie €4 und mit 50% Wahrscheinlichkeit gewinnen Sie €6. (akzeptieren / ablehnen)
Lotterie 4: Mit 50% Wahrscheinlichkeit verlieren Sie €5 und mit 50% Wahrscheinlichkeit gewinnen Sie €6. (akzeptieren / ablehnen)
Lotterie 5: Mit 50% Wahrscheinlichkeit verlieren Sie €6 und mit 50% Wahrscheinlichkeit gewinnen Sie €6. (akzeptieren / ablehnen)
Lotterie 6: Mit 50% Wahrscheinlichkeit verlieren Sie €7 und mit 50% Wahrscheinlichkeit gewinnen Sie €6. (akzeptieren / ablehnen)

# PREVIOUS DISCUSSION PAPERS

356 Petrishcheva,Vasilisa, Riener, Gerhard and Schildberg-Hörisch, Hannah, Loss Aversion in Social Image Concerns, November 2020.

355 Garcia-Vega, Maria, Kneller, Richard and Stiebale, Joel, Labor Market Reform and Innovation: Evidence from Spain, November 2020.

354 Steffen, Nico, Economic Preferences, Trade and Institutions, November 2020.

353 Pennerstorfer, Dieter, Schindler, Nora, Weiss, Christoph and Yontcheva, Biliana, Income Inequality and Product Variety: Empirical Evidence, October 2020.

352 Gupta, Apoorva, R&D and Firm Resilience During Bad Times, October 2020.

351 Shekhar, Shiva and Thomes, Tim Paul, Passive Backward Acquisitions and Downstream Collusion, October 2020.
Forthcoming in: Economics Letters.

350 Martin, Simon, Market Transparency and Consumer Search – Evidence from the German Retail Gasoline Market, September 2020.

349 Fischer, Kai and Haucap, Justus, Betting Market Efficiency in the Presence of Unfamiliar Shocks: The Case of Ghost Games during the COVID-19 Pandemic, August 2020.

348 Bernhardt, Lea, Dewenter, Ralf and Thomas, Tobias, Watchdog or Loyal Servant? Political Media Bias in US Newscasts, August 2020.

347 Stiebale, Joel, Suedekum, Jens and Woessner, Nicole, Robots and the Rise of European Superstar Firms, July 2020.

346 Horst, Maximilian, Neyer, Ulrike and Stempel, Daniel, Asymmetric Macroeconomic Effects of QE-Induced Increases in Excess Reserves in a Monetary Union, July 2020.

345 Riener, Gerhard, Schneider, Sebastian O. and Wagner, Valentin, Addressing Validity and Generalizability Concerns in Field Experiments, July 2020.

344 Fischer, Kai and Haucap, Justus, Does Crowd Support Drive the Home Advantage in Professional Soccer? Evidence from German Ghost Games during the COVID-19 Pandemic, July 2020.

343 Gösser, Niklas and Moshgbar, Nima, Smoothing Time Fixed Effects, July 2020.

342 Breitkopf, Laura, Chowdhury, Shyamal, Priyam, Shambhavi, Schildberg-Hörisch, Hannah and Sutter, Matthias, Do Economic Preferences of Children Predict Behavior?, June 2020.

341 Westphal, Matthias, Kamhöfer, Daniel A. and Schmitz, Hendrik, Marginal College Wage Premiums under Selection into Employment, June 2020.

340 Gibbon, Alexandra J. and Schain, Jan Philip, Rising Markups, Common Ownership, and Technological Capacities, June 2020.

339 Falk, Armin, Kosse, Fabian, Schildberg-Hörisch, Hannah and Zimmermann, Florian, Self-Assessment: The Role of the Social Environment, May 2020.

338 Schildberg-Hörisch, Hannah, Trieu, Chi and Willrodt, Jana, Perceived Fairness and Consequences of Affirmative Action Policies, April 2020.

337 Avdic, Daniel, de New, Sonja C. and Kamhöfer, Daniel A., Economic Downturns and Mental Wellbeing, April 2020.

336 Dertwinkel-Kalt, Markus and Wey, Christian, Third-Degree Price Discrimination in Oligopoly When Markets Are Covered, April 2020.

335 Dertwinkel-Kalt, Markus and Köster, Mats, Attention to Online Sales: The Role of Brand Image Concerns, April 2020.

334 Fourberg, Niklas and Korff, Alex, Fiber vs. Vectoring: Limiting Technology Choices in Broadband Expansion, April 2020.
Published in: Telecommunications Policy, 44 (2020), 102002.

333 Dertwinkel-Kalt, Markus, Köster, Mats and Sutter, Matthias, To Buy or Not to Buy? Price Salience in an Online Shopping Field Experiment, April 2020.
Revised version published in: European Economic Review, 130 (2020), 103593.

332 Fischer, Christian, Optimal Payment Contracts in Trade Relationships, February 2020.

331 Becker, Raphael N. and Henkel, Marcel, The Role of Key Regions in Spatial Development, February 2020.

330 Rösner, Anja, Haucap, Justus and Heimeshoff, Ulrich, The Impact of Consumer Protection in the Digital Age: Evidence from the European Union, January 2020.
Forthcoming in: International Journal of Industrial Organization.

329 Dertwinkel-Kalt, Markus and Wey, Christian, Multi-Product Bargaining, Bundling, and Buyer Power, December 2019.
Published in: Economics Letters, 188 (2020), 108936.

328 Aghelmaleki, Hedieh, Bachmann, Ronald and Stiebale, Joel, The China Shock, Employment Protection, and European Jobs, December 2019.

327 Link, Thomas, Optimal Timing of Calling In Large-Denomination Banknotes under Natural Rate Uncertainty, November 2019.

326 Heiss, Florian, Hetzenecker, Stephan and Osterhaus, Maximilian, Nonparametric Estimation of the Random Coefficients Model: An Elastic Net Approach, September 2019.

325 Horst, Maximilian and Neyer, Ulrike, The Impact of Quantitative Easing on Bank Loan Supply and Monetary Policy Implementation in the Euro Area, September 2019.
Published in: Review of Economics, 70 (2019), pp. 229-265.

324 Neyer, Ulrike and Stempel, Daniel, Macroeconomic Effects of Gender Discrimination, September 2019.

323 Stiebale, Joel and Szücs, Florian, Mergers and Market Power: Evidence from Rivals' Responses in European Markets, September 2019.

322 Henkel, Marcel, Seidel, Tobias and Suedekum, Jens, Fiscal Transfers in the Spatial Economy, September 2019.

321 Korff, Alex and Steffen, Nico, Economic Preferences and Trade Outcomes, August 2019.

320 Kohler, Wilhelm and Wrona, Jens, Trade in Tasks: Revisiting the Wage and Employment Effects of Offshoring, July 2019.

319 Cobb-Clark, Deborah A., Dahmann, Sarah C., Kamhöfer, Daniel A. and Schildberg-Hörisch, Hannah, Self-Control: Determinants, Life Outcomes and Intergenerational Implications, July 2019.

318 Jeitschko, Thomas D., Withers, John A., Dynamic Regulation Revisited: Signal Dampening, Experimentation and the Ratchet Effect, July 2019.

317 Jeitschko, Thomas D., Kim, Soo Jin and Yankelevich, Aleksandr, Zero-Rating and Vertical Content Foreclosure, July 2019.

316 Kamhöfer, Daniel A. und Westphal, Matthias, Fertility Effects of College Education: Evidence from the German Educational Expansion, July 2019.

315 Bodnar, Olivia, Fremerey, Melinda, Normann, Hans-Theo and Schad, Jannika, The Effects of Private Damage Claims on Cartel Stability: Experimental Evidence, June 2019.

314 Baumann, Florian and Rasch, Alexander, Injunctions Against False Advertising, October 2019 (First Version June 2019).
Published in: Canadian Journal of Economics, 53 (2020), pp. 1211-1245.

313 Hunold, Matthias and Muthers, Johannes, Spatial Competition and Price Discrimination with Capacity Constraints, May 2019 (First Version June 2017 under the title “Capacity Constraints, Price Discrimination, Inefficient Competition and Subcontracting”).
Published in: International Journal of Industrial Organization, 67 (2019), 102524.

312 Creane, Anthony, Jeitschko, Thomas D. and Sim, Kyoungbo, Welfare Effects of Certification under Latent Adverse Selection, March 2019.

311 Bataille, Marc, Bodnar, Olivia, Alexander Steinmetz and Thorwarth, Susanne, Screening Instruments for Monitoring Market Power – The Return on Withholding Capacity Index (RWC), March 2019.
Published in: Energy Economics, 81 (2019), pp. 227-237.

310 Dertwinkel-Kalt, Markus and Köster, Mats, Salience and Skewness Preferences, March 2019.
Published in: Journal of the European Economic Association, 18 (2020), pp. 2057–2107.

309 Hunold, Matthias and Schlütter, Frank, Vertical Financial Interest and Corporate Influence, February 2019.

308 Sabatino, Lorien and Sapi, Geza, Online Privacy and Market Structure: Theory and Evidence, February 2019.

307 Izhak, Olena, Extra Costs of Integrity: Pharmacy Markups and Generic Substitution in Finland, January 2019.

306 Herr, Annika and Normann, Hans-Theo, How Much Priority Bonus Should be Given to Registered Organ Donors? An Experimental Analysis, December 2018.
Published in: Journal of Economic Behavior and Organization, 158 (2019), pp.367-378.

305 Egger, Hartmut and Fischer, Christian, Increasing Resistance to Globalization: The Role of Trade in Tasks, December 2018.
Published in: European Economic Review, 126 (2020), 103446.

304 Dertwinkel-Kalt, Markus, Köster, Mats and Peiseler, Florian, Attention-Driven Demand for Bonus Contracts, October 2018.
Published in: European Economic Review, 115 (2019), pp.1-24.

303 Bachmann, Ronald and Bechara, Peggy, The Importance of Two-Sided Heterogeneity for the Cyclicality of Labour Market Dynamics, October 2018.
Forthcoming in: The Manchester School

302 Hunold, Matthias, Hüschelrath, Kai, Laitenberger, Ulrich and Muthers, Johannes, Competition, Collusion and Spatial Sales Patterns – Theory and Evidence, September 2018.
Forthcoming in: Journal of Industrial Economics.

301 Neyer, Ulrike and Sterzel, André, Preferential Treatment of Government Bonds in Liquidity Regulation – Implications for Bank Behaviour and Financial Stability, September 2018.

300 Hunold, Matthias, Kesler, Reinhold and Laitenberger, Ulrich, Hotel Rankings of Online Travel Agents, Channel Pricing and Consumer Protection, September 2018 (First Version February 2017).
Forthcoming in: Marketing Science.

299 Odenkirchen, Johannes, Pricing Behavior in Partial Cartels, September 2018.

298 Mori, Tomoya and Wrona, Jens, Inter-city Trade, September 2018.

297 Rasch, Alexander, Thöne, Miriam and Wenzel, Tobias, Drip Pricing and its Regulation: Experimental Evidence, August 2018.
Published in: Journal of Economic Behavior and Organization, 176 (2020), pp. 353-370.

296 Fourberg, Niklas, Let’s Lock Them in: Collusion under Consumer Switching Costs, August 2018.

295 Peiseler, Florian, Rasch, Alexander and Shekhar, Shiva, Private Information, Price Discrimination, and Collusion, August 2018.

294 Altmann, Steffen, Falk, Armin, Heidhues, Paul, Jayaraman, Rajshri and Teirlinck, Marrit, Defaults and Donations: Evidence from a Field Experiment, July 2018.
Published in: Review of Economics and Statistics, 101 (2019), pp. 808-826.

293 Stiebale, Joel and Vencappa, Dev, Import Competition and Vertical Integration: Evidence from India, July 2018.

292 Bachmann, Ronald, Cim, Merve and Green, Colin, Long-run Patterns of Labour Market Polarisation: Evidence from German Micro Data, May 2018.
Published in: British Journal of Industrial Relations, 57 (2019), pp. 350-376.

291 Chen, Si and Schildberg-Hörisch, Hannah, Looking at the Bright Side: The Motivation Value of Overconfidence, May 2018.
Published in: European Economic Review, 120 (2019), 103302.

290 Knauth, Florian and Wrona, Jens, There and Back Again: A Simple Theory of Planned Return Migration, May 2018.

289 Fonseca, Miguel A., Li, Yan and Normann, Hans-Theo, Why Factors Facilitating Collusion May Not Predict Cartel Occurrence – Experimental Evidence, May 2018.
Published in: Southern Economic Journal, 85 (2018), pp. 255-275.

288 Benesch, Christine, Loretz, Simon, Stadelmann, David and Thomas, Tobias, Media Coverage and Immigration Worries: Econometric Evidence, April 2018.
Published in: Journal of Economic Behavior & Organization, 160 (2019), pp. 52-67.

287 Dewenter, Ralf, Linder, Melissa and Thomas, Tobias, Can Media Drive the Electorate? The Impact of Media Coverage on Party Affiliation and Voting Intentions, April 2018.
Published in: European Journal of Political Economy, 58 (2019), pp. 245-261.

286 Jeitschko, Thomas D., Kim, Soo Jin and Yankelevich, Aleksandr, A Cautionary Note on Using Hotelling Models in Platform Markets, April 2018.

285 Baye, Irina, Reiz, Tim and Sapi, Geza, Customer Recognition and Mobile Geo-Targeting, March 2018.

284 Schaefer, Maximilian, Sapi, Geza and Lorincz, Szabolcs, The Effect of Big Data on Recommendation Quality. The Example of Internet Search, March 2018.

283 Fischer, Christian and Normann, Hans-Theo, Collusion and Bargaining in Asymmetric Cournot Duopoly – An Experiment, October 2018 (First Version March 2018).
Published in: European Economic Review, 111 (2019), pp.360-379.

282 Friese, Maria, Heimeshoff, Ulrich and Klein, Gordon, Property Rights and Transaction Costs – The Role of Ownership and Organization in German Public Service Provision, February 2018.
Published in: International Journal of Industrial Organization, 72 (2020), 102637.

281 Hunold, Matthias and Shekhar, Shiva, Supply Chain Innovations and Partial Ownership, February 2018.

280 Rickert, Dennis, Schain, Jan Philip and Stiebale, Joel, Local Market Structure and Consumer Prices: Evidence from a Retail Merger, January 2018.
Forthcoming in: Journal of Industrial Economics under the title "The Effect of Mergers on Retail Prices: Evidence from Germany".

Older discussion papers can be found online at:
http://ideas.repec.org/s/zbw/dicedp.html

**Heinrich-Heine-Universität Düsseldorf**

**Düsseldorfer Institut für Wettbewerbsökonomie (DICE)**

Universitätsstraße 1, 40225 Düsseldorf

ISSN 2190-992X (online)
ISBN 978-3-86304-355-1