

Khalmetski, Kiryl; Sliwka, Dirk

Working Paper

Disguising Lies - Image Concerns and Partial Lying in Cheating Games

CESifo Working Paper, No. 6347

Provided in Cooperation with:

Ifo Institute – Leibniz Institute for Economic Research at the University of Munich

Suggested Citation: Khalmetski, Kiryl; Sliwka, Dirk (2017) : Disguising Lies - Image Concerns and Partial Lying in Cheating Games, CESifo Working Paper, No. 6347, Center for Economic Studies and ifo Institute (CESifo), Munich

This Version is available at:

<https://hdl.handle.net/10419/155589>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



Disguising Lies – Image Concerns and Partial Lying in Cheating Games

Kiryl Khalmetski
Dirk Sliwka

CESIFO WORKING PAPER NO. 6347

CATEGORY 13: BEHAVIOURAL ECONOMICS

FEBRUARY 2017

An electronic version of the paper may be downloaded

- from the SSRN website: www.SSRN.com
- from the RePEc website: www.RePEc.org
- from the CESifo website: www.CESifo-group.org/wp

ISSN 2364-1428

CESifo

Center for Economic Studies & Ifo Institute

Disguising Lies – Image Concerns and Partial Lying in Cheating Games

Abstract

We study equilibrium reporting behavior in Fischbacher and Föllmi-Heusi (2013) type cheating games when agents have a fixed cost of lying and image concerns not to be perceived as a liar. We show that equilibria naturally arise in which agents with low costs of lying randomize among a set of the highest potential reports. Such equilibria induce a distribution of reports in line with observed experimental patterns. We also find that higher image concerns lead to an increase in the range of reported lies while the effect of the fixed cost of lying is the opposite.

JEL-Codes: D030, D820, D830, C720.

Keywords: cost of lying, image concerns, cheating game, truth-telling, deception.

Kiryl Khalmetski
University of Cologne
Department of Economics
Albertus-Magnus-Platz
Germany – 50923 Cologne
kiryl.khalmetski@uni-koeln.de

Dirk Sliwka
University of Cologne
Department of Economics
Albertus-Magnus-Platz
Germany – 50923 Cologne
sliwka@wiso.uni-koeln.de

February 3, 2017

We thank Johannes Abeler, Martin Dufwenberg and Joel Sobel for helpful comments. Financial support of the German Research Foundation (DFG) through the Research Unit “Design and Behavior” (FOR 1371) is gratefully acknowledged.

1 Introduction

Economists have only recently started to experimentally investigate lying behavior. Gneezy (2005) has shown that a significant share of subjects have a strong aversion to lying besides purely distributional preferences. Subsequent studies explored various determinants of individual lying aversion such as the payoff consequences for the sender and the receiver of the message (Erat and Gneezy, 2012), the form of communication (Lundquist et al., 2009), anticipated trust of the receiver (Charness and Dufwenberg, 2006; Sutter, 2009), or the size of the lie (Kajackaite and Gneezy, 2015).

An important experimental paradigm for studying motivations for honesty has been developed by Fischbacher and Föllmi-Heusi (2013) which already in the rather short time frame since its publication has been used in a large number of experimental studies. A recent meta-study by Abeler et al. (2016) already includes more than 70 studies using this paradigm. In this setting, subjects privately have to roll a die and then report the number they have observed. The payoff is simply equivalent to the reported number (except for the number 6 which leads to a payoff of zero). An important advantage of the method is that it avoids strategic interactions between subjects and thus facilitates the identification of intrinsic motives for lying behavior. Moreover, in the baseline design the experimenter cannot verify the reported messages on the individual level, hence, subjects can be certain that lies are not discovered, but lying behavior can still be inferred from the distribution of reported numbers. If all subjects told the truth each number should be reported with a likelihood of $1/6$. However, in the experiment the fractions of numbers 4 and 5 among all reported numbers were significantly above the expected $1/6$ (i.e., 27.2% for the 4 and 35.0% for the 5). Hence, a substantial number of subjects did not tell the truth - but also apparently many did not maximize their own material payoff (i.e., did not report the number 5). As a possible explanation, Fischbacher and Föllmi-Heusi (2013) suggest that some subjects might have a reputational concern motivating them to disguise their lying to an outside observer (which can be an experimenter or a future self) by reporting not the payoff-maximizing number. This conjecture was further substantiated by the experimental evidence that partial lying in cheating games turns to be significantly less frequent if lying can be individually verified by the experimenter (Gneezy et al., 2016; Kajackaite and Gneezy, 2015; Abeler et al., 2016).

We now formalize this idea by developing a model which assumes both a fixed

cost of lying *and* image concerns not to be viewed as a liar and study whether equilibrium behavior can indeed predict patterns such as those described in the experiment. In our model agents consider three motives: (i) they are better off when earning more money, (ii) they have an intrinsic cost of lying (which varies between individuals) and (iii) an agent’s utility is decreasing in the likelihood with which an outside observer believes her to have lied after having observed the reported number.¹ We show that such context naturally gives rise to equilibria in which agents who lie randomize their reporting behavior among a subset of the most favorable reports.

We furthermore show that the lower bound of the set of numbers liars report in equilibrium is decreasing in the strength of the agents’ image concerns. If image concerns are weak, all liars report the highest feasible number. In the opposite case when image concerns become very strong, liars randomize over the whole set of possible reports excluding the smallest number. Thus, too strong reputational concerns of the sender may actually backfire for the receiver since then she can never be sure to have obtained the correct information (except for the case of the lowest message). Notably, we show that this effect goes in the opposite direction of the effect of an increase in the fixed costs of lying, which leads to a smaller range of reported lies. Finally, we consider the effect of the stake size. While the total *rate* of lying unambiguously increases with the stake size, the *range* of reported lies can both increase and decrease depending on the parameters. The latter fact is due to the opposing effects from a relative decrease in the reputational and fixed costs of lying.

With very few calibrated parameters, the model produces a good fit of the predicted behavior to the actual data from Fischbacher and Föllmi-Heusi (2013) and related experiments. While our model is primarily motivated to explain the empirical evidence from experimental cheating games, it is worth noting that such games also capture structural properties of many real-life communication settings where: 1) transmitted information is not easily verifiable (or verifiable with a significant delay), 2) there are material incentives to misreport, and 3) the sender of the message might care about not being perceived as a liar at the time

¹Models of fixed (i.e., belief-independent) costs of lying were developed by Kartik et al. (2007) and Kartik (2009). Ottaviani and Sørensen (2006a) and Ottaviani and Sørensen (2006b) considered settings where the sender cares about the inferred precision of his private signal. The concern for the inferred social preferences was modeled in Bernheim (1994), Bénabou and Tirole (2006) and Ellingsen and Johannesson (2008) where an agent prefers to be perceived as having a certain *type* of preferences. In contrast to these models, we assume that the sender cares about the receiver’s inference regarding the *action* of lying. See Abeler et al. (2016) for further classification of the related models.

of communication. Potential examples include medical diagnoses, promises in political election campaigns, the sale of credence goods, and academic reference letters.

Simultaneously to our work, a few recent papers aim to explain the experimental data from cheating games. In Gneezy et al. (2016) and Abeler et al. (2016), the main model specifications also combine both fixed costs of lying and image concerns depending on the probability others assign that the agent lies. Unlike us, Dufwenberg and Dufwenberg (2016) incorporate only (belief-dependent) costs that are proportional to the size of a lie perceived by the receiver. In terms of qualitative analytical results, the closest study to ours is of Gneezy et al. (2016) who showed some of the important equilibrium properties also outlined in our analysis. While our model has somewhat stricter (though, empirically justified) restrictions on preferences (such as the invariance of the fixed lying cost to the observed number and the size of the lie), at the same time this allows us to provide a complete mathematical characterization of the (unique) equilibrium distribution of reports. Moreover, we analyze the comparative statics of the equilibrium with respect to an increase in the image concerns, fixed lying costs and monetary stakes, obtaining qualitatively different and testable predictions for each case that allow to explain several experimental findings.²

The remainder of the paper is organized as follows. Section 2 presents the model setting, Section 3 provides the equilibrium analysis, Section 4 shows the results of the calibration of the model to the actual data from Fischbacher and Föllmi-Heusi (2013) and Kajackaite and Gneezy (2015), and Section 5 concludes.

2 The Model

2.1 Benchmark Preferences: No Reputational Concerns

We consider a population of agents who play a variant of the Fischbacher and Föllmi-Heusi (2013) cheating game. First, an agent privately observes a uniformly distributed random (integer) number $y \in Y = \{0, 1, \dots, K\}$ such that $\Pr(y) = \frac{1}{K+1}$.³ She then has to report this number to a receiver making a report $x \in Y$ which leads to a payment equal to x to the agent. A perfectly selfish agent would then always report the highest possible number K .

²The focus of the comparative statics analysis in the other studies is somewhat different. In particular, Gneezy et al. (2016) primarily consider the effect of the underlying distribution of the observed numbers.

³Setting the lowest number to 0 is without loss of generality for the qualitative results.

Before we proceed to our full model that includes reputational concerns we start by investigating a benchmark case where agents have only a fixed cost of lying $l > 0$ such that their utility is reduced by l whenever they report $x \neq y$.⁴ Hereby, l is distributed according to a strictly increasing continuous cdf $F(l)$ with $F(l) = 0$ for $l \leq 0$ and $F(l) > 0$ for $l > 0$.⁵ Note that our model allows for a large probability mass on arbitrarily small lying costs: The assumption that $F(0) = 0$ just implies that agents who are indifferent between lying and telling the truth, will tell the truth (similarly to a lexicographic preference for truth-telling). We also assume that $F(K) < 1$, i.e., there is a strictly positive probability that there are agents who never lie even if this would yield the highest possible payoff (so that all numbers are reported in equilibrium). The utility function is then given by

$$u(l, x, y) = x - l \cdot I_{\mathbf{L}}(x, y), \quad (1)$$

where \mathbf{L} denotes the set of lies, i.e., $\mathbf{L} = \{(x, y) \in Y^2 : x \neq y\}$ and $I_{\mathbf{L}}$ is the indicator function. Such preferences imply that if an agent prefers lying by reporting some $x < K$ over truth-telling (i.e., $x - l > y$ for some $x \in Y$), then she prefers to lie to the highest extent, since $K - l > x - l$ for any $x < K$. Hence, in equilibrium no partial lying to $x < K$ can occur. This directly leads to the following result.

Proposition 1 *Under preferences given by (1), all liars report $x = K$ and the likelihood of reporting $x < K$ does not exceed $\frac{1}{K+1}$.*

Thus, a (belief-independent) lying cost cannot by itself explain overreporting of numbers strictly lower than K - the main empirical puzzle in the experiment of Fischbacher and Föllmi-Heusi (2013) - which motivates the incorporation of reputational costs into the agent's utility function in the next section.⁶

⁴The experiments of Gneezy et al. (2013), Kajackaite and Gneezy (2015) and Gneezy et al. (2016) show that the strength of (inference-independent) lying aversion does not depend on the size of a lie.

⁵Gneezy et al. (2013) provide empirical evidence for heterogeneous fixed lying costs with a large share of subjects being characterized by intermediate costs.

⁶Note that the second statement of Proposition 1 holds also for more complex specifications of the lying cost function. For instance, if the lying cost is concave in the size of the lie, then the net benefit from lying increases in the reported number. Consequently, no partial lies occur as well. In case of a convex (size-dependent) cost of lying, partial lying is possible. However, under such preferences, once an agent's type prefers to lie to some $x < K$ while observing $y < x$, the same type would also prefer to lie at least to $x + 1$ while observing x (as the marginal cost of lying in the latter case is weakly less steep). Hence, under uniform distribution of the observed states, the likelihood of lying to any given $x < K$ cannot exceed the likelihood of lying *after* observing x . As a result, still no $x < K$ would be overreported.

2.2 Reputational Concerns

In our full model, agents have also concern for reputation of being honest besides the fixed intrinsic lying costs. In particular, we assume that the agent's utility is decreasing in the likelihood that the receiver or some outside observer thinks that she has told lie, i.e., in $\Pr[y \neq x|x]$.⁷ Thus, her utility function is

$$u(l, x, y) = x - l \cdot I_L(x, y) - \eta \cdot \Pr[y \neq x|x] \quad (2)$$

such that parameter $\eta > 0$ captures the extent to which agents dislike being viewed as liars by the receiver. Since the utility function depends on the receiver's beliefs, we obtain a psychological game (Geanakoplos et al., 1989; Battigalli and Dufwenberg, 2009). The term $-\eta \cdot \Pr[y \neq x|x]$ is referred below as the agent's reputational payoff.

3 Equilibrium Analysis

3.1 General Characterization

We now consider Bayesian Nash equilibria of the game which are characterized by (i) the reporting strategy of an agent as a function of her cost of lying l and the observed number y and (ii) the receiver's beliefs about the likelihood that an agent lied which is a function of the reported number x . Denote the (mixed) strategy of an agent as

$$\sigma(l, y) = (p_{ly}^0, p_{ly}^1, \dots, p_{ly}^K),$$

which is a probability distribution over the $K+1$ pure reporting strategies, i.e., an agent with cost of lying l who has observed y reports number x with probability p_{ly}^x . The receiver's belief that a report x was truthful is denoted as

$$r(x) = \Pr[y = x|x].$$

Then, an equilibrium is a set of mixed strategies and beliefs satisfying the following conditions:

$$\forall(l, x, y) : p_{ly}^x > 0 \text{ only if } x \in \arg \max_{x'} u(l, x', y), \quad (3)$$

⁷See also Bénabou and Tirole (2006), Fischbacher and Föllmi-Heusi (2013) and Dufwenberg and Dufwenberg (2016) for discussions regarding the agent herself being a plausible audience.

$$\forall(l, y) : \sum_{x=0}^K p_{ly}^x = 1, \quad (4)$$

$$\forall(x) : r(x) = \frac{\Pr[y = x \wedge x \text{ reported}]}{\Pr[x \text{ reported}]} = \frac{\int_0^\infty p_{lx}^x dF(l)}{\sum_{y=0}^K (\int_0^\infty p_{ly}^x dF(l))}. \quad (5)$$

The last condition ensures that beliefs are formed by Bayes rule.^{8,9}

Denote by X_L the set of reports chosen by lying agents in equilibrium. We first show that in any equilibrium X_L is non-empty.

Lemma 1 *There exists no equilibrium with complete truth-telling.*

Proof: Assume by contradiction that all agents tell the truth. Then, $r(x) = 1$ for all $x \in Y$. But then agents with sufficiently small costs of lying l observing $y < K$ would deviate and report $x = K$. ■

When X_L is a singleton we have a pure strategy equilibrium, and it is straightforward to show that the only feasible candidate with this property is one in which $X_L = \{K\}$.¹⁰ To see that, suppose that all agents who lie choose $x' < K$. In that case someone reporting $x = K$ must be a truth-teller and thus $r(x') < r(K) = 1$. But then the lying agents could earn a higher material payoff and reputation when reporting the highest possible report K .

Note, however, that such an equilibrium where $X_L = \{K\}$ may not exist. The reason is the following: suppose that sufficiently many liars report K . Then such a report can lead to a strong loss in reputation. In turn, when η is large enough liars will have an incentive to deviate and report a lower number. We can indeed show:

Lemma 2 *In any equilibrium in which the set of reports chosen by liars X_L is a singleton, all liars report $x = K$. If η is sufficiently large, X_L can never be a singleton.*

Proof:

The first part of the lemma follows from the above considerations.

⁸Recall that there are always people who never lie (since $F(K) < 1$). Hence any value of x is reported with strictly positive probability by a truthful agent on the equilibrium path.

⁹For our game, the defined equilibrium is equivalent to sequential equilibrium for psychological games developed by Battigalli and Dufwenberg (2009).

¹⁰The assumption that all agents are either pure liars or pure truth-tellers is without loss of generality since the mass of agents whose lying cost l makes them indifferent between lying and truth-telling has zero measure.

Let us show the second part. Assume that in equilibrium all liars report $x = K$. Then, the total likelihood of lying $\Pr[x = K|y \neq K]$ is bounded from below for any η . Indeed, consider agents observing $y = 0$ with $l < K - 1$. These types will never tell the truth in equilibrium, since they have a strict incentive to deviate to $x = K - 1$ earning $K - 1 - l$, which is then strictly higher than the payoff of 0 they earn from truth-telling (while there is no reputational loss in both cases). Thus, the probability of lying is bounded by the fraction of these types, i.e.,

$$\Pr[x = K|y \neq K] \geq \frac{1}{K+1}F(K-1). \quad (6)$$

Furthermore, note that agents observing K never report some $x' \neq K$ (earning $x' - l$) since otherwise it should hold $x' > K - \eta(1 - r(K))$ so that all liars to K would have a strict incentive to also deviate to x' (being indifferent in terms of lying costs). Then, by Bayes rule

$$\begin{aligned} r(K) &= \frac{\Pr[y = K \wedge K \text{ reported}]}{\Pr[K \text{ reported}]} = \frac{\frac{1}{K+1}}{\frac{1}{K+1} + \Pr[x = K|y \neq K]} \\ &\leq \frac{\frac{1}{K+1}}{\frac{1}{K+1} + \frac{1}{K+1}F(K-1)} = \frac{1}{1 + F(K-1)}, \end{aligned} \quad (7)$$

i.e., $r(K)$ is bounded from above by $\frac{1}{1+F(K-1)} < 1$.

Moreover, if $X_L = \{K\}$ any agent who lies must prefer to report $x = K$ to reporting $x = K - 1$ or

$$\begin{aligned} K - 1 &\leq K - \eta \cdot (1 - r(K)) \\ \Leftrightarrow 1 - \frac{1}{\eta} &\leq r(K). \end{aligned} \quad (8)$$

Conditions (7) and (8) thus yield $1 - \frac{1}{\eta} \leq \frac{1}{1+F(K-1)}$ which leads to a contradiction when η is sufficiently large. ■

Hence, when η is sufficiently large liars will not all report the same number in equilibrium as otherwise someone reporting this number is perceived as a liar with a too high probability which makes a deviation attractive. However, there may be equilibria where liars report numbers from a larger set $X_L \subseteq Y$. Intuitively, randomizing reports over a larger set makes it easier to disguise a lie. It is important to note that this can only be the case when liars are indifferent between

all messages in X_L or

$$x - \eta \cdot (1 - r(x)) = \rho \text{ for all } x \in X_L \quad (9)$$

where ρ is some constant. To see this, note that if a person lies, the “optimal lie” (i.e., the value of x maximizing $x - \eta \cdot (1 - r(x)) - l$) does not depend on the fixed cost of lying l . We use this to further characterize necessary conditions for all feasible equilibria:

Proposition 2 *In any equilibrium the following properties hold:*

- (i) *Agents who lie make a report which is larger or equal than a cut-off value x_L where $1 \leq x_L \leq K$.*
- (ii) *The probability that someone who reports x is a truth-teller is*

$$r(x) = \begin{cases} 1 - \frac{x-\rho}{\eta} & \text{if } x \geq x_L \\ 1 & \text{if } x < x_L, \end{cases} \quad (10)$$

where ρ is some constant in the interval $(\max\{0, K - \eta\}, K)$.

- (iii) *Agents who have observed a value $y \geq x_L$ report their observed values truthfully.*

Proof:

First, note that in any equilibrium X_L is non-empty by Lemma 1. Let us show that agents who observed $y \in X_L$ report their true value. Assume by contradiction that there exist $y', y'' \in X_L$ such that some agents observing y' lie by reporting $x = y''$. For these agents, the following must hold

$$\begin{aligned} y' - \eta(1 - r(y')) &\leq y'' - \eta(1 - r(y'')) - l \\ \Leftrightarrow y' - \eta(1 - r(y')) &< y'' - \eta(1 - r(y'')). \end{aligned} \quad (11)$$

This implies that no agent observing $y \neq y'$ would like to lie to y' , strictly preferring to report y'' . Hence, $y' \notin X_L$ which is a contradiction. It follows that all untruthful reports are done by the agents observing $y \notin X_L$.

Next, we show that X_L consists of the largest elements of Y by contradiction: suppose that x_L is the smallest element of X_L and there is a specific value $x' > x_L$ which is not an element of X_L so that $r(x') = 1 > r(x_L)$. Then $u(x', y) > u(x_L, y)$ for all $y \neq x_L$ (as both material and reputational payoffs are higher by reporting x') and liars to x_L would deviate and report x' instead. Thus, $x_L \notin X_L$ which is a contradiction.

The existence of the cut-off value, together with condition (9), leads to the characterization of the probability of truth-telling (10). Moreover, $x_L \neq 0$ as the opposite implies that $X_L = Y$ and hence nobody would lie since, as we have shown, agents observing $y \in X_L$ report their number truthfully. This is a contradiction by Lemma 1.

Since agents observing $y \in X_L$ tell the truth while liars must use all messages in X_L by definition, we have

$$1 > r(x) > 0 \text{ for all } x \in X_L, \quad (12)$$

which together with $K \in X_L$ and condition (9) leads to $\rho \in (K - \eta, K)$. Furthermore, ρ must be strictly positive. Otherwise, we would have $\rho - l < y$ for all (l, y) which would imply that liars observing $y \notin X_L$ have a strict incentive to deviate to truth-telling. Thus, $\rho \in (\max\{0, K - \eta\}, K)$. ■

In equilibrium we must thus observe a specific pattern in the association between a report and the reputation for being a truth-teller given this report. This reputation is linearly decreasing within the set of reports chosen by liars: the loss in money when choosing a lower report must be exactly offset by an equivalent gain in reputation so that the sum of monetary and reputational payoffs remains constant at some level ρ (see (9)).

As Proposition 2 also shows, all agents who have observed a $y \geq x_L$ tell the truth. The intuition for this is simple: reporting a different value within X_L would lead to the identical sum of monetary and reputational payoffs but comes along with a utility loss due to the cost of lying l .

We can now consider the behavior of agents who have observed a true value y strictly below the cut-off value x_L . When reporting the truth such an agent's utility is y and when lying she obtains $\rho - l$. Hence, she will report truthfully if and only if

$$\rho - y \leq l. \quad (13)$$

This implies:

Lemma 3 *If an agent observes a value $y < x_L$ she reports truthfully if and only if her cost of lying l exceeds a cut-off value $l^*(y) = \rho - y$.*

We can use Proposition 2 and Lemma 3 to characterize the equilibrium distribution of reported numbers:

Proposition 3 *In equilibrium the probability that x is reported is*

$$\Pr[x \text{ reported}] = \begin{cases} \frac{\eta}{(K+1)(\eta-x+\rho)} & \text{if } x \geq x_L \\ \frac{1}{K+1} (1 - F(\rho - x)) & \text{if } x < x_L \end{cases} \quad (14)$$

which is strictly increasing in x . Moreover, $\Pr[x \text{ reported}] > \frac{1}{K+1}$ for $x \geq x_L$ and $\Pr[x \text{ reported}] \leq \frac{1}{K+1}$ for $x < x_L$.

Proof:

By Bayes rule,

$$r(x) = \frac{\Pr[y = x \wedge x \text{ reported}]}{\Pr[x \text{ reported}]}. \quad (15)$$

At the same time, by Proposition 2 for any $x \geq x_L$

$$\Pr[y = x \wedge x \text{ reported}] = \frac{1}{K+1} \quad (16)$$

as x is observed with probability $\frac{1}{K+1}$ and all agents who observe it tell the truth. We then obtain from (15) and (16) that

$$\Pr[x \text{ reported}] = \frac{1}{K+1} \frac{1}{r(x)} \text{ for } x \geq x_L, \quad (17)$$

which is strictly larger than $\frac{1}{K+1}$. By substituting $r(x) = 1 - \frac{x-\rho}{\eta}$ from Proposition 2 we obtain that

$$\Pr[x \text{ reported}] = \frac{\eta}{(K+1)(\eta-x+\rho)} \quad (18)$$

in this case.

For any $x < x_L$ the likelihood that x is reported is (given Lemma 3)

$$\begin{aligned} \Pr[x \text{ reported}] &= \Pr[x \text{ observed}] \cdot \Pr[\rho - x \leq l] \\ &= \frac{1}{K+1} (1 - F(\rho - x)) \leq \frac{1}{K+1}. \end{aligned} \quad (19)$$

The claim that $\Pr[x \text{ reported}]$ is strictly increasing in x for $x \geq x_L$ follows directly from (18). It is left to show that $\Pr[x \text{ reported}]$ strictly increases in x also for $x < x_L$. Given (19), a sufficient condition for this is that

$$\rho - (x_L - 1) \geq 0, \quad (20)$$

in which case $F(\rho - x)$ strictly decreases in x for $x \leq x_L - 1$. Assume the

opposite, i.e., that $\rho - (x_L - 1) < 0 \Leftrightarrow x_L - 1 > \rho$. Then, any agent who lies would be better off reporting $x_L - 1$ (getting at least $x_L - 1 - l$) instead of an $x \geq x_L$ (getting $\rho - l$) which leads to a contradiction by Proposition 2. ■

Proposition 3 thus implies a specific pattern in the observed distribution of reports. First of all, all numbers above the threshold x_L are reported more frequently than they are actually observed, and all numbers below the threshold are reported (weakly) less frequently than observed.¹¹ The reason is straightforward: only agents who observed $y < x_L$ lie and they only lie to $x \geq x_L$.

Furthermore, the higher x the higher is the likelihood that x is reported. This is due to two effects: For $x < x_L$ all agents who report such an x are truth-tellers. But within this set the incentives to lie are stronger for lower observed numbers as here the material gain from lying is larger. For larger values of x ($x \geq x_L$) the fact that the likelihood of a report x must increase in x is due to the reputational payoff obtained in equilibrium. As laid out in the above, all numbers reported by liars in equilibrium must lead to the same total payoff. In turn, any material gain must be offset by an equivalent reputational loss. As higher values of x come along with a higher material gain, more liars must thus choose to report them in equilibrium.

3.2 Existence and Uniqueness

We have shown so far that an equilibrium of the game can be characterized by the cut-off value x_L and the gross utility from lying (not including the fixed lying cost) ρ . The following result shows that the equilibrium values of these variables are directly linked.

Lemma 4 *In any equilibrium x_L is equal to the smallest integer strictly larger than ρ , that is*

$$x_L(\rho) = \min \{x \in \{1, \dots, K\} \mid x > \rho\}. \quad (21)$$

Proof:

By Proposition 2, the equilibrium likelihood that a report x is truthful is

$$r(x) = 1 - \frac{x - \rho}{\eta} < 1 \text{ for all } x \geq x_L, \quad (22)$$

¹¹The number $x_L - 1$ can still be reported with probability $\frac{1}{K+1}$ if $\rho = x_L - 1$, which is not precluded by the incentive constraint (20).

where the inequality follows from the fact that there are always liars reporting an $x \geq x_L$. This implies that $x > \rho$ for all $x \geq x_L$ and in turn we must have $x_L > \rho$. At the same time, $x_L \leq \rho + 1$ is given by (20). ■

In equilibrium, the total likelihood of sending a false message by an agent observing $y < x_L$ must be equal to the total likelihood of receiving a false message. We now use this “accounting property” to show the existence and investigate the uniqueness of equilibrium.

Theorem 1 *All equilibria of the game induce the same distribution of reported numbers characterized by $\rho^* \in (\max\{0, K - \eta\}, K)$, which is implicitly defined by the equation*

$$\sum_{y=0}^{x_L(\rho^*)-1} F(\rho^* - y) = \sum_{x=x_L(\rho)}^K \left(\frac{x - \rho^*}{\eta - x + \rho^*} \right) \quad (23)$$

that always has a unique solution in this interval.

Proof:

By Proposition 2 and Lemma 3 the total fraction of liars is

$$\Pr[\text{Lie}] = \sum_{y=0}^{x_L(\rho)-1} \frac{1}{K+1} F(\rho - y) \quad (24)$$

as a function of ρ . At the same time we can determine the total probability that a reported number is a lie as

$$\Pr[\text{Lie}] = \sum_{x=x_L(\rho)}^K \Pr[x \text{ reported}] \cdot \Pr[x \text{ is a lie} | x \text{ reported}].$$

Using Proposition 3 and the fact that $\Pr[x \text{ is a lie} | x \text{ reported}] = 1 - r(x) = \frac{x-\rho}{\eta}$, this is equal to

$$\sum_{x=x_L(\rho)}^K \frac{\eta}{(K+1)(\eta - x + \rho)} \frac{x - \rho}{\eta} = \sum_{x=x_L(\rho)}^K \frac{x - \rho}{(K+1)(\eta - x + \rho)}. \quad (25)$$

In equilibrium the fraction of liars (24) must be equal to the fraction of reported lies (25) which then leads to condition (23).

In a next step, we show that (23) always has a unique solution on $(\max\{0, K - \eta\}, K)$. In order to do so, consider the function

$$\theta(\rho) = \sum_{y=0}^{x_L(\rho)-1} F(\rho - y) - \sum_{x=x_L(\rho)}^K \frac{x - \rho}{\rho + \eta - x} \quad (26)$$

which is strictly increasing in ρ on $(\max\{0, K - \eta\}, K)$. Let us show that $\theta(\rho) = 0$ always has a unique solution $\rho^* \in (\max\{0, K - \eta\}, K)$.

First, we can show that

$$\lim_{\rho \rightarrow \max\{0, K - \eta\}^+} \theta(\rho) < 0, \quad (27)$$

$$\lim_{\rho \rightarrow K^-} \theta(\rho) > 0. \quad (28)$$

Indeed, when $\rho \rightarrow \max\{0, K - \eta\}$ (from above) then either $\theta(\rho) \rightarrow -\infty$ (if $\max\{0, K - \eta\} = K - \eta$) or $\theta(\rho) \rightarrow -\sum_{x=1}^K \frac{x}{\eta - x} < 0$ (if $\max\{0, K - \eta\} = 0$). When $\rho \rightarrow K$ (from below) then $x_L = K$ by Lemma 4 so that $\theta(\rho) \rightarrow \sum_{y=0}^{K-1} F(K - y) > 0$.

We now show that $\theta(\rho)$ is continuous on $(\max\{0, K - \eta\}, K)$. First note that $\theta(\rho)$ is continuous when ρ is not an integer as in this case $x_L(\rho)$ does not vary by Lemma 4. We now demonstrate that the function is continuous also at integer values of ρ . Suppose we have an integer value of $\rho' \in (\max\{0, K - \eta\}, K)$. Then $x_L(\rho') = \rho' + 1$ and

$$\theta(\rho') = \sum_{y=0}^{\rho'} F(\rho' - y) - \sum_{x=\rho'+1}^K \frac{x - \rho'}{\rho' + \eta - x}.$$

It is straightforward to see that $\lim_{\varepsilon \rightarrow 0} \theta(\rho' + \varepsilon) = \theta(\rho')$, since x_L does not vary. Now consider $\lim_{\varepsilon \rightarrow 0} \theta(\rho' - \varepsilon)$. First note that $x_L(\rho' - \varepsilon) = \rho'$ for $\varepsilon < 1$. Hence,

$$\begin{aligned} \theta(\rho' - \varepsilon) &= \sum_{y=0}^{\rho'-1} F(\rho' - \varepsilon - y) - \sum_{x=\rho'}^K \left(\frac{x - \rho' + \varepsilon}{\rho' - \varepsilon + \eta - x} \right) \\ &= \sum_{y=0}^{\rho'} F(\rho' - \varepsilon - y) - F(\rho' - \varepsilon - \rho') \\ &\quad - \sum_{x=\rho'+1}^K \left(\frac{x - \rho' + \varepsilon}{\rho' - \varepsilon + \eta - x} \right) - \left(\frac{\rho' - \rho' + \varepsilon}{\rho' - \varepsilon + \eta - \rho'} \right) \end{aligned}$$

$$\begin{aligned}
&= \sum_{y=0}^{\rho'} F(\rho' - \varepsilon - y) - F(-\varepsilon) \\
&\quad - \sum_{x=\rho'+1}^K \left(\frac{x - \rho' + \varepsilon}{\rho' - \varepsilon + \eta - x} \right) - \left(\frac{\varepsilon}{\eta - \varepsilon} \right).
\end{aligned}$$

But then, as $\lim_{\varepsilon \rightarrow 0} F(-\varepsilon) = 0 = F(0)$,

$$\lim_{\varepsilon \rightarrow 0} \theta(\rho' - \varepsilon) = \sum_{y=0}^{\rho'} F(\rho' - y) - \sum_{x=\rho'+1}^K \left(\frac{x - \rho'}{\rho' + \eta - x} \right) = \theta(\rho'), \quad (29)$$

and thus the function is continuous. This together with (27), (28), the fact that $\theta(\rho)$ is increasing and the intermediate value theorem implies that $\theta(\rho) = 0$ has a unique solution $\rho^* \in (\max\{0, K - \eta\}, K)$. Given that ρ uniquely pins down the equilibrium distribution of reported numbers by Proposition 3 and Lemma 4, the claim follows. \blacksquare

The intuition behind the condition (23) is the following. Recall that only agents who observed an $y < x_L(\rho)$ lie. And only reports $x \geq x_L(\rho)$ can be lies. In equilibrium the likelihood that an agent who observed $y < x_L(\rho)$ decided to lie (the left-hand side of (23)) must be equal to the likelihood that an agent who reported $x \geq x_L(\rho)$ actually lied (the right-hand side of (23)). The left-hand side of (23) is strictly increasing in ρ . Intuitively, the higher the gross payoff of lying ρ the more agents lie. At the same time, the right-hand side of (23) – the total probability that a reported number is a lie – is strictly decreasing in ρ . Intuitively, the larger the fraction of lies among the reported numbers the smaller is the reputation of somebody reporting an $x \geq x_L$ and the smaller is thus ρ .

Therefore there is a unique value of ρ that solves this equation and, by Proposition 3 and Lemma 4, we thus can infer that any equilibrium must induce the same distribution of messages. We can now combine these insights to characterize the equilibrium strategies p_{ly}^x chosen by liars. Note that there are multiple equilibria but – as the following result shows – all are payoff equivalent to the unique mixed strategy equilibrium where all lying agents observing $y < x_L$ pursue symmetric randomization strategies, i.e., $p_{ly}^x = p^x$ for any $y < x_L$, $x \geq x_L$ and $l < \rho - y$.

Proposition 4 *There is a unique mixed strategy equilibrium with symmetric lying strategies. In this equilibrium an agent lies if and only if $x < x_L(\rho^*)$ and $l < \rho^* - y$ and otherwise tells the truth. All liars report $x \geq x_L(\rho^*)$ with proba-*

bility

$$p^x = \frac{1}{\sum_{y=0}^{x_L(\rho^*)-1} F(\rho^* - y)} \frac{x - \rho^*}{\eta - x + \rho^*}. \quad (30)$$

All other equilibria yield the same distribution of messages and are payoff-equivalent.

Proof:

To derive p^x for given ρ recall that we know from Proposition 3 that in equilibrium

$$\Pr[x \text{ reported}] = \frac{\eta}{(K+1)(\eta - x + \rho)} \text{ for all } x \geq x_L.$$

But by the law of total probability this must be equal to

$$\begin{aligned} \Pr[x \text{ reported}] &= \frac{\eta}{(K+1)(\eta - x + \rho)} \\ &= \Pr[y = x \wedge x \text{ reported}] + \Pr[y \neq x \wedge x \text{ reported}] \\ &= \frac{1}{K+1} + \sum_{y=0}^{x_L-1} \Pr[x \text{ reported}|y] \frac{1}{K+1} \\ &= \frac{1}{K+1} + \sum_{y=0}^{x_L-1} F(\rho - y) p^x \frac{1}{K+1} \end{aligned} \quad (31)$$

$$\Leftrightarrow \frac{\eta}{\eta - x + \rho} = 1 + \sum_{y=0}^{x_L-1} F(\rho - y) p^x. \quad (32)$$

Solving for p^x yields

$$p^x = \frac{1}{\sum_{y=0}^{x_L-1} F(\rho - y)} \frac{x - \rho}{\eta - x + \rho}.$$

Note that the reporting strategies characterized by p^x , x_L , $l^*(y)$ and beliefs $r(x)$ that are all uniquely pinned down by the equilibrium value of ρ^* , indeed constitute an equilibrium. In particular, the consistency of strategies with incentive constraint (3) and beliefs with Bayesian rule (5) has been shown previously. Besides, it is easy to verify that the remaining equilibrium condition (4) holds for the derived p^x since the condition (23) is satisfied. Hence, the existence of a unique (symmetric) equilibrium is guaranteed.

Finally, note that there can be multiple other equilibria yielding the same distribution of messages (uniquely characterized by ρ^* by Theorem 1). Such equilibria entail asymmetric lying strategies, i.e., strategies where p_{ly}^x is not identical for all $y < x_L(\rho^*)$ and $l < l^*(y)$. However, since ρ^* uniquely defines $r(x)$, x_L and $l^*(y)$, the total payoffs of all agents are the same in all these equilibria. ■

Proposition 4 implies that in the symmetric equilibrium the likelihood assigned by liars to report x is increasing and convex in x . Intuitively, once the reputation from reporting a number becomes low it gets less sensitive to a marginal increase in the probability mass of agents lying to this number (by Bayes' formula). Consequently, one requires a larger increase in p^x to generate the same amount of reputational loss.

At the same time, there are multiple payoff-equivalent equilibria yielding the same distribution of messages. This result is due to the fact that for a given cost of lying l and observed value y an agent is indifferent between all reports in $x \geq x_L$. Hence, all equilibria in which p_{ly}^x differs across l and y induce exactly the same payoffs as the equilibrium where $p_{ly}^x = p^x$ given that the total distribution of reported numbers remains the same. In any asymmetric equilibrium the role of the random variable l (or y) is thus purely that of a coordination device. In particular, there are also equilibria with partial lying where a given type of agent does not randomize between reports but instead plays a pure strategy, and the equilibrium distribution of reports arises due to different groups of lying agents playing different strategies.

We have thus shown that in any equilibrium agents with sufficiently small lying costs choose to lie when the realized number they observed is smaller than a cut-off value x_L . These agents then (generally) randomize among reports $x \geq x_L$. In the next section, we investigate the determinants of this cut-off value x_L in order to further characterize equilibrium behavior.

3.3 Comparative Statics

3.3.1 Effect of a Change in the Image Concerns

We now consider how a change in the agents' image concerns affects the equilibrium distribution of the reported numbers.

Proposition 5 *If the agents' image concerns η increase, then*

- (i) x_L weakly decreases and
- (ii) the likelihood that an agent lies strictly decreases.

Proof:

- (i) We start with the accounting condition (23)

$$\theta(\rho, \eta) = \sum_{y=0}^{x_L(\rho)-1} F(\rho - y) - \sum_{x=x_L(\rho)}^K \frac{x - \rho}{\rho + \eta - x} = 0. \quad (33)$$

By Theorem 1, this implicitly defines a function $\rho^*(\eta) \in (\max\{0, K - \eta\}, K)$ such that $\theta(\rho^*(\eta), \eta) = 0$. Consider some $0 < \eta' < \eta''$. As $\theta(\rho, \eta)$ is strictly increasing in η for a given ρ we must have that

$$\theta(\rho^*(\eta''), \eta') < \theta(\rho^*(\eta''), \eta'') = 0 = \theta(\rho^*(\eta'), \eta') \quad (34)$$

such that

$$\theta(\rho^*(\eta''), \eta') < \theta(\rho^*(\eta'), \eta'). \quad (35)$$

At the same time, for given η , $\theta(\rho, \eta)$ is strictly increasing in ρ on $(\max\{0, K - \eta\}, K)$. This together with (35) implies that

$$\rho^*(\eta'') < \rho^*(\eta')$$

for any $\eta'' > \eta' > 0$. It follows that $\rho^*(\eta)$ is strictly decreasing in η and, by Lemma 4, $x_L(\rho^*(\eta))$ must then be (weakly) decreasing in η .

(ii) Lemma 3 implies

$$\Pr[\text{Lie}] = \frac{1}{K+1} \sum_{y=0}^{x_L(\rho^*)-1} F(\rho^* - y). \quad (36)$$

From the proof of claim (i) we know that ρ^* is strictly decreasing in η , which together with $\rho^* > 0$ by Theorem 1 leads to the claim. \blacksquare

Hence, stronger image concerns reduce the fraction of liars but enlarge the set of numbers they report. An intuition for the latter effect is that for high values of η liars suffer more from a loss in reputation. Hence, the incentives for liars to deviate to a lower number (which would not be conceived as a lie) become stronger. In equilibrium this pushes x_L downwards.

Next, we can show that in the limits of η the highest and the lowest feasible values of x_L are reached.

Proposition 6 (i) *There is a cut-off value $\bar{\eta}$ such that $x_L(\eta) = 1$ for all $\eta \geq \bar{\eta}$.*
(ii) *There is a cut-off value $\underline{\eta}$ such that $x_L(\eta) = K$ for all $\eta \leq \underline{\eta}$.*

Proof:

(i) Consider again condition (23):

$$\sum_{y=0}^{x_L(\rho^*(\eta))-1} F(\rho^*(\eta) - y) = \sum_{x=x_L(\rho^*(\eta))}^K \left(\frac{x - \rho^*(\eta)}{\rho^*(\eta) + \eta - x} \right).$$

Since $\rho^*(\eta) > 0$ by Theorem 1, the right-hand side, which is continuously decreasing in ρ^* on $(\max\{0, K - \eta\}, K)$, is bounded from above for a given η by

$$\sum_{x=x_L(\rho^*(\eta))}^K \left(\frac{x}{\eta - x} \right)$$

which tends to 0 if $\eta \rightarrow \infty$. Consequently, the left-hand side must go to 0 as well. Since

$$\sum_{y=0}^{x_L(\rho^*(\eta))-1} F(\rho^*(\eta) - y) \geq F(\rho^*(\eta) - 0) = F(\rho^*(\eta)) \quad (37)$$

we then obtain that $F(\rho^*(\eta)) \rightarrow 0$. As F is a continuously increasing function and $F(0) = 0$, this implies that $\lim_{\eta \rightarrow \infty} \rho^*(\eta) = 0$. This together with Lemma 4 leads to the claim.

(ii) In equilibrium, $\rho^* > K - \eta$ by Theorem 1. Consequently, at least for any $\eta < 1$ we must have $\rho^* > K - 1$ so that $x_L(\eta) = K$ by Lemma 4. ■

Note that the result of Proposition 6 (i) has potentially interesting welfare implications. As we have seen already in the above, larger image concerns have a dual effect on reporting strategies: on the one hand, the fraction of lies is reduced, but on the other hand, it becomes harder to infer whether a person is telling the truth from observing a lower report as liars also report smaller numbers. When $\eta \geq \bar{\eta}$ so that $x_L = 1$ an observer can never be sure that a reported message is not a lie (except for the lowest message). If such an observer would have to take a decision based on this information and mistakes are extremely costly, this might preclude her from taking a (potentially welfare-improving) action given the unavoidable risk of a mistake. Thus, image concerns of the agent may backfire for the receiver under certain circumstances.

Finally, we can show that if image concerns become very large, lying disappears even for types observing the lowest number:

Corollary 1 *If $\eta \rightarrow \infty$, then the rate of lying tends to 0 for all $y \in Y$. The distribution of the observed messages converges to a uniform distribution.*

Proof:

By Proposition 6 (i), if $\eta \rightarrow \infty$ then at some point x_L becomes 1 so that all agents observing $y \geq 1$ tell the truth (by Proposition 2). At the same time, by Lemma 3 the rate of lying for $y = 0$ is equal to $F(\rho)$, which tends to 0 by the proof of Proposition 6 (i). ■

It is important to note that the latter is of course a limit result. In equilibrium there will always be liars as we have shown in Lemma 1. But when η becomes very large the loss in reputation when reporting a high number is also sizeable even if there is only a small fraction of liars, and this, in turn, prevents other agents from lying.

3.3.2 Effect of a Change in the Fixed Costs of Lying

Let us now consider how the equilibrium is affected by a change in the distribution of the fixed lying costs F . We consider an increase in F in the sense of first-order stochastic dominance (FOSD). That is, assume that the family of lying costs distributions can be parametrized by λ , such that for any $\lambda' < \lambda''$ and $z > 0$

$$F_{\lambda'}(z) > F_{\lambda''}(z). \quad (38)$$

Proposition 7 *If the distribution of lying costs F increases in the sense of FOSD, then*

- (i) x_L weakly increases and
- (ii) the likelihood that an agent lies strictly decreases.

Proof:

(i) Consider the accounting condition (23) now denoted as a function of ρ and λ :

$$\theta(\rho, \lambda) = \sum_{y=0}^{x_L(\rho)-1} F_{\lambda}(\rho - y) - \sum_{x=x_L(\rho)}^K \frac{x - \rho}{\rho + \eta - x} = 0. \quad (39)$$

By Theorem 1, this implicitly defines a function $\rho^*(\lambda) \in (\max\{0, K - \eta\}, K)$ such that $\theta(\rho^*(\lambda), \lambda) = 0$. By (38), for any given $\rho > 0$ and any $\lambda' < \lambda''$

$$\theta(\rho, \lambda') > \theta(\rho, \lambda'').$$

Hence, we must have that

$$\theta(\rho^*(\lambda''), \lambda') > \theta(\rho^*(\lambda''), \lambda'') = 0 = \theta(\rho^*(\lambda'), \lambda') \quad (40)$$

such that

$$\theta(\rho^*(\lambda''), \lambda') > \theta(\rho^*(\lambda'), \lambda'). \quad (41)$$

At the same time, for given λ , $\theta(\rho, \eta)$ is strictly increasing in ρ on $(\max\{0, K - \eta\}, K)$. Consequently, (41) implies $\rho^*(\lambda'') > \rho^*(\lambda')$ for any $\lambda' < \lambda''$ and thus $\rho^*(\lambda)$ must be strictly increasing in λ . By Lemma 4, $x_L(\rho^*(\lambda))$ must then be (weakly) increasing in λ .

(ii) Lemma 3 and Theorem 1 imply

$$\Pr[\text{Lie}] = \frac{1}{K+1} \sum_{y=0}^{x_L(\rho^*)-1} F_\lambda(\rho^* - y) = \frac{1}{K+1} \sum_{x=x_L(\rho^*)}^K \frac{x - \rho^*}{\rho^* + \eta - x}. \quad (42)$$

From the proof of claim (i) we know that ρ^* is strictly increasing in λ , which together with the fact that the right-hand side of (42) is decreasing in ρ^* leads to the claim. \blacksquare

Thus, the effect of an increase in the fixed lying costs goes in the opposite direction of the effect of an increase in the image concerns not to be perceived as a liar: the range of reported lies *shrinks* as the fixed lying costs increase. The intuition for this effect is that at higher lying costs less agents lie and thus the reputational loss from reporting a high number decreases. In turn, the relative incentive to disguise lies by reporting smaller values is reduced.

Next, we consider again limit results. Let us take a convention that if $\lambda \rightarrow \infty$ then $F_\lambda(z) \rightarrow 0$, and if $\lambda \rightarrow 0$ then $F_\lambda(z) \rightarrow 1$ for any $z > 0$. First, we can show that if lying costs get very large, agents lie only to the highest possible number.

Proposition 8 *There is a cut-off value $\bar{\lambda}$ such that $x_L(\lambda) = K$ for all $\lambda \geq \bar{\lambda}$.*

Proof:

Consider the equilibrium condition (23) which implicitly defines ρ^* as a function of λ :

$$\sum_{y=0}^{x_L(\rho^*(\lambda))-1} F_\lambda(\rho^*(\lambda) - y) = \sum_{x=x_L(\rho^*(\lambda))}^K \left(\frac{x - \rho^*(\lambda)}{\rho^*(\lambda) + \eta - x} \right). \quad (43)$$

If $\lambda \rightarrow \infty$, then $F_\lambda(\rho - y) \rightarrow 0$ for any given ρ and y . Since $\rho^*(\lambda)$ is bounded from above by K (by Theorem 1), this implies that the left-hand side of (43) and hence the right-hand side of (43) converge to 0 for $\lambda \rightarrow \infty$. Since the denominator on the

right-hand side of (43) is bounded from above and all terms in the corresponding sum are positive (given that $x_L > \rho^*$ by Lemma 4 and $\rho^* > K - \eta$ by Theorem 1), this is only possible if $x_L = K$ and $\rho^*(\lambda) \rightarrow K$. ■

Moreover, analogously to the previous case, one can show that lying disappears once the fixed lying costs get very large.

Corollary 2 *If $\lambda \rightarrow \infty$, then the rate of lying tends to 0 for all $y \in Y$. The distribution of the observed messages converges to a uniform distribution.*

Proof:

As shown in the proof of Proposition 8, if $\lambda \rightarrow \infty$ then $\sum_{y=0}^{x_L(\rho^*)-1} F_\lambda(\rho^* - y) \rightarrow 0$. By Lemma 3, this implies that the total likelihood of lying also converges to 0. ■

At the same time, it is not generally true that x_L converges to the minimal possible value of 1 if λ approaches 0 (given that it is increasing in λ by Proposition 7). As the following result shows, $\lim_{\lambda \rightarrow 0} x_L$ can be as high as K depending on the size of the agents' image concerns η :¹²

Proposition 9 *i) If η is sufficiently large, then $\lim_{\lambda \rightarrow 0} x_L = 1$.*

ii) If η is sufficiently small, then $\lim_{\lambda \rightarrow 0} x_L = K$.

Proof:

(i) Proposition 6 implies that for any given λ and sufficiently large η we must have $x_L(\lambda, \eta) = 1$. Since $x_L(\lambda, \eta)$ is non-decreasing in λ by Proposition 7, the claim holds.

(ii) By Theorem 1, $\rho^* > K - \eta$. Consequently, at least for any $\eta < 1$ we must have $\rho^* > K - 1$ so that $x_L(\rho^*) = K$ for any distribution F_λ . ■

If the lying costs become very small in the population of agents, the intrinsic incentives to truthfully report numbers lower than K disappear. Yet, the image concerns not to be perceived as a liar remain. If η is large this still provides an incentive to report low numbers to disguise the lie. If, however, η is small, this incentive becomes weaker and agents focus on reporting the highest number for any lying costs.

¹²More generally, one can show that $\lim_{\lambda \rightarrow 0} x_L$ can take any value between 1 and K depending on η . The proof is available upon request.

Proposition 9 shows that partial lying to reports below K – the main feature of the characterized equilibrium – can emerge also if the fixed lying costs are negligible.¹³ At the same time, the next result demonstrates that intrinsic lying costs are still required to explain the typical pattern observed in experiments using the paradigm that at the same time i) at least several lowest numbers are not overreported (i.e., $x_L \geq 2$; that is there are at least two lowest reports which are reported with a probability between 0 and $\frac{1}{K+1}$), and (ii) the reported numbers have full support on Y (i.e., even $x = 0$ is reported with a positive probability). In particular, if the fixed lying costs converge to 0 while $\lim_{\lambda \rightarrow 0} x_L \geq 2$, the lowest number is almost never reported in contrast to this evidence.

Proposition 10 *If η is such that $\lim_{\lambda \rightarrow 0} x_L \geq 2$, then $\lim_{\lambda \rightarrow 0} \Pr[x = 0] = 0$.*

Proof:

Note that for any λ

$$\rho^*(\lambda) \geq x_L(\lambda) - 1 \geq 1, \quad (44)$$

where the first inequality is by Lemma 4, and the second inequality is due to Proposition 7 and $\lim_{\lambda \rightarrow 0} x_L \geq 2$. Hence, $\rho^*(\lambda)$ is bounded from below by 1. By Lemma 3, the lying rate conditional on observing 0 is $F_\lambda(\rho^*)$ which converges to 1 when $\lambda \rightarrow 0$ (and ρ^* is bounded from below) such that $\lim_{\lambda \rightarrow 0} \Pr[x = 0] = 0$. ■

Hence, having non-negligible fixed costs of lying in the model (besides image concerns - recall Proposition 1) is necessary to explain the whole pattern of experimental data.¹⁴

3.3.3 Effect of a Change in the Monetary Stakes

As a final step in the theoretical analysis, we consider the effect of the size of monetary stakes on the structure of the equilibrium reporting behavior. Thus assume now that the agent's utility is

$$u(l, x, y) = s \cdot x - l \cdot I_{\mathbf{L}}(x, y) - \eta \cdot \Pr[y \neq x | x],$$

where s denotes the stake size. Note that the agent's choice problem is equivalent to maximizing

$$x - \frac{l}{s} \cdot I_{\mathbf{L}}(x, y) - \frac{\eta}{s} \cdot \Pr[y \neq x | x]$$

¹³Note that we still need $F(0) = 0$, i.e., the assumption that people avoid the lie when they are otherwise indifferent between lying and telling the truth, to ensure the existence of the considered equilibrium (see the proof of Theorem 1).

¹⁴This is consistent with the results of Abeler et al. (2016) who came to a similar conclusion within their modeling framework.

such that an increase in stake size can be analogously expressed as a joint decrease in image concerns and lying costs. Then, we can apply all of our earlier results substituting the image concerns parameter and costs of lying appropriately. In particular, an increase in the stake size (i.e., $s > 1$) corresponds to a shift in the distribution of the costs of lying in the sense of FOSD since for any $z > 0$

$$F(z) = \Pr[l \leq z] < \Pr[l/s \leq z] = F_s(z), \quad (45)$$

where $F_s(z)$ is the distribution of the normalized lying costs.

Proposition 11 *An increase in monetary stakes raises the likelihood that an agent lies.*

Proof:

For any $\eta' < \eta''$ and $\lambda' < \lambda''$ we must have that

$$\Pr[\text{Lie}|\eta', \lambda'] > \Pr[\text{Lie}|\eta', \lambda''] > \Pr[\text{Lie}|\eta'', \lambda''], \quad (46)$$

where the first inequality follows from Proposition 5 (ii) and the second from Proposition 7 (ii). Hence, any joint reduction in η and λ (which is outcome equivalent to an increase in s as shown above) causes a (strict) increase in the lying rate. ■

Proposition 11 is in line with the experimental evidence of Kajackaite and Gneezy (2015) who show that the total rate of lying significantly increases with the stake size once controlling for the perceived negative consequences from being eventually detected as a liar (which in turn might be correlated with monetary stakes).¹⁵

At the same time, our model does not provide clear directional predictions regarding the effect of the stake size on the *range* of reported lies, i.e., on x_L . As outlined above, an increase in the stake size is equivalent to a joint decrease in both the reputational lying costs (η) and the fixed lying costs (λ). Yet, according to Propositions 5 and 7, these two effects push x_L in the opposite directions.

¹⁵In some other studies not controlling for this (e.g., Mazar et al., 2008) the evidence for the stake size effect on lying is ambiguous. In terms of our model, an increase in the perceived negative consequences of lying with the monetary stakes would correspond to an exogenous increase in λ , thus indeed countervailing the positive effect of the stake size on lying.

The meta-study by Abeler et al. (2016) doesn't find stake size effects when comparing lying behavior across different experiments in different labs. However, if they analyze only studies in which the stake size is varied within the the same experiment, they find positive stake size effects.

The net effect depends on the parameter values, in particular on the form of distribution F . If the resulting decrease in the relative fixed lying costs is more prominent, the range of reported lies can expand with s (see Appendix for a numerical example). Notably, this is in line with the effect of the stake size in Fischbacher and Föllmi-Heusi (2013) who observed a shift in reporting the second-highest number of 4 from 17.72% to 27.50% (with only the second value being significantly different from 1/6) as the stake size was tripled.¹⁶

4 Empirical Calibration

The model predicts the main qualitative patterns observed in Fischbacher and Föllmi-Heusi (2013)-type cheating games. In particular, it rationalizes the existence of a threshold x_L with numbers below (above) x_L being under(over)reported (see Proposition 3). And, moreover, it shows why there are “partial lies”, i.e., why not all liars report the payoff maximizing number.

Of course, there may be other behavioral mechanisms at work in addition to those we capture in our model. But it is yet instructive to verify how well the model can be calibrated to fit the actual experimental data in Fischbacher and Föllmi-Heusi (2013). For the distribution of lying costs F , we assume that it is drawn from a normal distribution with mean 0 and standard deviation σ , while being left-truncated at 0.¹⁷ The calibration was obtained by finding the values of η and σ which minimize the sum of squared deviations of the predicted and the observed values.

The parameter values yielding the best fit of the message frequencies predicted by the model to the data in the Baseline treatment of Fischbacher and Föllmi-Heusi (2013) are $\eta = 3.65$ and $\sigma = 2.96$ (Fig. 1). One can see that the model predicts the main data patterns from the experiment, in particular the substantial amount of partial lying to $x = 4$, as well as a positive fraction of subjects reporting 0.¹⁸ Notably, setting $\eta = 0$, i.e., using the benchmark model without reputational concerns results in a much worse fit than that of the main model (see Fig. 2).¹⁹

¹⁶The rates of reporting smaller numbers were lower than 1/6 in both treatments. The control treatment was formed within the same subject sample, and hence was different from the baseline treatment reported in our Section 4.

¹⁷We assume that the distribution of lying costs is shifted towards 0 given the frequent experimental observation that the probability mass of individual social preference parameters is shifted towards payoff maximization (see, e.g., Engel, 2011).

¹⁸Similarly good fit is obtained by calibrating the model to the aggregated data from the meta-study of Abeler et al. (2016).

¹⁹The mean squared error is 6.79 in the main model with image concerns and 29.90 in the benchmark model (in terms of percentage points). According to the F-test, the first model fits

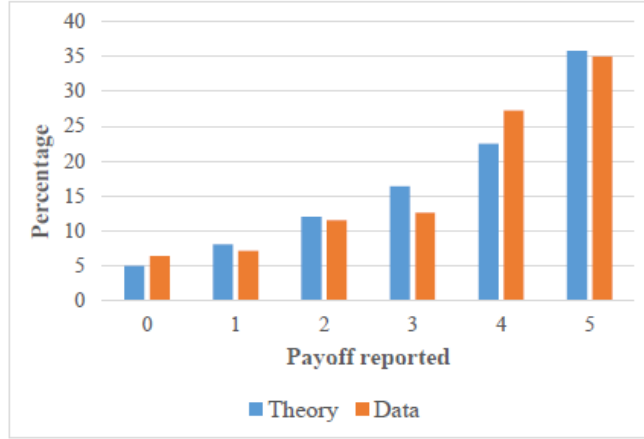


Figure 1: Empirical calibration of the main model to the Baseline Treatment in Fischbacher and Föllmi-Heusi (2013).

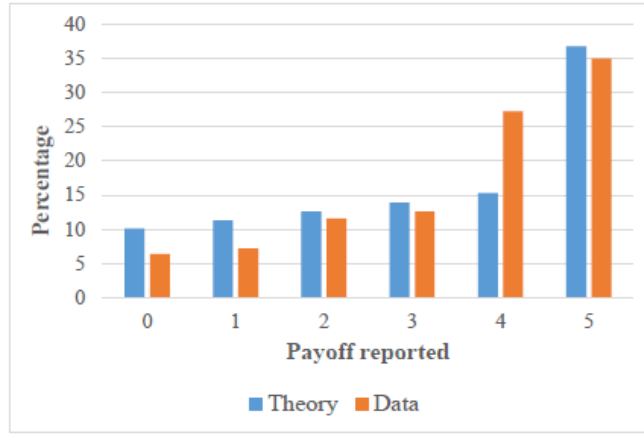


Figure 2: Empirical calibration of the benchmark model without reputational concerns ($\eta = 0$) to the Baseline Treatment in Fischbacher and Föllmi-Heusi (2013).

Finally, the model is also capable to explain some of the comparative statics results observed in the experimental cheating games. As noted above, Kajackaite and Gneezy (2015) compared two treatments which differed in whether the experimenter was able to verify (ex-post) the true number rolled by the participant (as in one case the number was protocolled by the computer software). Note that in the limit case when the agent is absolutely sure that her lying will be ex-post verified by the observer, x is not informative about the likelihood of lying anymore (which would imply $\eta = 0$ in our model), while the image loss from lying would be captured instead by the fixed cost of lying l . In line with this prediction, calibrating the empirical results from the two treatments in Kajackaite and Gneezy the data significantly better ($p = 0.021$).

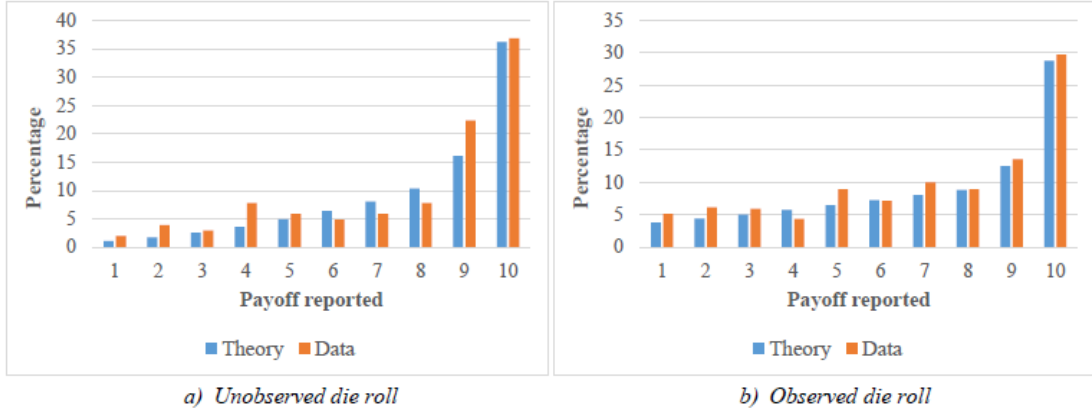


Figure 3: Empirical calibration of the main model to the experimental treatments in Kajackaite and Gneezy (2015).

(2015) with the model results in that the treatment with privately observed die roll corresponds to relatively larger sensitivity to the inference-driven image loss ($\eta = 3.18$ vs. $\eta = 2.44$) while yielding relatively lower estimate of the fixed costs of lying ($\sigma = 4.27$ vs. $\sigma = 9.06$), see Fig. 3. Note also that the estimation of η in the private treatment in Kajackaite and Gneezy (2015) is close to that in the Baseline treatment of Fischbacher and Föllmi-Heusi (2013) (3.18 vs. 3.65).

5 Conclusion

We have shown that incomplete lying behavior naturally arises in the Fischbacher and Föllmi-Heusi (2013) setting when agents have (i) a fixed cost of lying and (ii) image concerns not to be perceived as a liar. When image concerns are sufficiently strong agents can “disguise” a lie only when liars randomize over a set of reports. This allows to reduce larger reputational losses occurring when all liars report the highest feasible report. In particular, as we have shown an agent’s reputation to be honest must be strictly decreasing in the reported number so that any monetary gain is offset by an equivalent loss in reputation.

We furthermore have shown how the distribution of the reported numbers changes with the agents’ image concerns. When they are weak, all liars report the number yielding the maximum feasible payoff. As image concerns increase, the lower bound of the set of numbers reported by the liars shifts to the left. In the limit, liars randomize over all but the lowest feasible number. Hence, higher image concerns of the agent may actually backfire in terms of the quality of information transmission: in the limit case, the receiver can never be sure that the reported number is not a lie, except for the lowest message. In contrast,

higher fixed costs of lying always lead to a reduction in the range of reported lies.

The model can explain the stylized facts identified by Fischbacher and Föllmi-Heusi (2013): a) the fraction of people reporting a payoff of 0 is positive, b) the fraction of people reporting 5 is above $1/6$; c) the fraction of people reporting 4 is above $1/6$. Moreover, the model can replicate the observed difference in behavior once the experimenter is able to verify lying ex-post (as in Kajackaite and Gneezy, 2015), attributing this to a shift in both lying costs and inference-based image concerns.

Overall, our model provides a useful workhorse for the analysis of empirical data from cheating experiments. The model may of course be extended in several directions. For instance, one could allow that lying costs depend upon the distance between the true and reported number. Note, however, that even size-dependent (fixed) costs of lying cannot *alone* explain the overreporting of the second-highest number (see footnote 6). Thus, the key ingredient yielding partial lying in the considered setting is an image concern not to be perceived as a liar, which we aimed to demonstrate in the clearest way putting aside additional interactions with more complex preferences. Moreover, with only two components, i.e., fixed costs of lying and image concerns, our model attains already a reasonably good fit to the actually observed distribution of lies in cheating games. Future experiments may further verify the predictions of the model for instance by exogenously manipulating the strength of image concerns by varying the audience of outside observers (as in Ariely et al., 2009) or exogenously changing the costs of lying by imposing random monitoring and punishments for liars.

6 Appendix: Numerical Simulation of the Effect of the Monetary Stakes on x_L

When we substitute l and η according to a shift in the stake size to $s \cdot x$, equilibrium condition (23) becomes

$$\theta(\rho, s) = \sum_{y=0}^{x_L(\rho, s)-1} F_s(\rho - y) - \sum_{x=x_L(\rho, s)}^K \frac{x - \rho}{\rho + \frac{\eta}{s} - x} = 0, \quad (47)$$

where $F_s(z)$ is given by (45) (see Section 3.3.3). The equilibrium is characterized by a unique value $\rho^* \in (\max\{0, K - \frac{\eta}{s}\}, K)$ solving the equation.

Consider an example with $K = 5$, $\eta = 3.65$, and F being a normal distribution with mean μ and standard deviation $\sigma = 1$ (left-truncated at 0) so that $F_s(z) =$

$F(z|\mu/s, \sigma/s)$. Then, if $\mu = 0$ and s increases from 1 to 2, the value of $\rho^* \in (\max\{0, K - \frac{\eta}{s}\}, K)$ solving equation (47) changes from 2.681 to 3.584, and hence x_L increases from 3 to 4. At the same time, if $\mu = 5$, then the same increase in s leads to a decrease in ρ^* from 4.237 to 3.840 so that x_L drops from 5 to 4.

References

- Abeler, J., D. Nosenzo, and C. Raymond (2016). Preferences for truth-telling - a meta study. CESifo Working Paper Series 6087, CESifo Group Munich.
- Ariely, D., A. Bracha, and S. Meier (2009). Doing good or doing well? Image motivation and monetary incentives in behaving prosocially. *American Economic Review* 99(1), 544–555.
- Battigalli, P. and M. Dufwenberg (2009). Dynamic psychological games. *Journal of Economic Theory* 144(1), 1–35.
- Bénabou, R. and J. Tirole (2006). Incentives and prosocial behavior. *American Economic Review* 96(5), 1652–1678.
- Bernheim, B. D. (1994). A theory of conformity. *Journal of Political Economy* 100, 841–847.
- Charness, G. and M. Dufwenberg (2006). Promises and partnership. *Econometrica* 74(6), 1579–1601.
- Ellingsen, T. and M. Johannesson (2008). Pride and prejudice: The human side of incentive theory. *American Economic Review* 2011, 990–1008.
- Engel, C. (2011). Dictator games: A meta study. *Experimental Economics* 14(4), 583–610.
- Erat, S. and U. Gneezy (2012). White lies. *Management Science* 58(4), 723–733.
- Fischbacher, U. and F. Föllmi-Heusi (2013). Lies in disguise - an experimental study on cheating. *Journal of the European Economic Association* 11(3), 525–547.
- Geanakoplos, J., D. Pearce, and E. Stacchetti (1989). Psychological games and sequential rationality. *Games and Economic Behavior* 1, 60–79.
- Gneezy, U. (2005). Deception: The role of consequences. *American Economic Review* 95, 384–394.

- Gneezy, U., A. Kajackaite, and J. Sobel (2016). Lying aversion and the size of the lie. Technical report, UCSD.
- Gneezy, U., B. Rockenbach, and M. Serra-Garcia (2013). Measuring lying aversion. *Journal of Economic Behavior & Organization* 93, 293–300.
- Kajackaite, A. and U. Gneezy (2015). Lying costs and incentives. Technical report, UCSD.
- Kartik, N. (2009). Strategic communication with lying costs. *The Review of Economic Studies* 76(4), 1359–1395.
- Kartik, N., M. Ottaviani, and F. Squintani (2007). Credulity, lies, and costly talk. *Journal of Economic theory* 134(1), 93–116.
- Lundquist, T., T. Ellingsen, E. Gribbe, and M. Johannesson (2009). The aversion to lying. *Journal of Economic Behavior & Organization* 70(1-2), 81–92.
- Mazar, N., O. Amir, and D. Ariely (2008). The dishonesty of honest people: A theory of self-concept maintenance. *Journal of marketing research* 45(6), 633–644.
- Ottaviani, M. and P. N. Sørensen (2006a). Reputational cheap talk. *The Rand journal of economics* 37(1), 155–175.
- Ottaviani, M. and P. N. Sørensen (2006b). The strategy of professional forecasting. *Journal of Financial Economics* 81(2), 441–466.
- Sutter, M. (2009). Deception through telling the truth?! Experimental evidence from individuals and teams. *The Economic Journal* 119(534), 47–60.