DISCUSSION PAPER SERIES

IZA DP No. 6919

Truth-Telling: A Representative Assessment

Johannes Abeler
Anke Becker
Armin Falk

October 2012

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Truth-Telling: A Representative Assessment

Johannes Abeler
University of Oxford
and IZA

Anke Becker
University of Bonn

Armin Falk
University of Bonn
and IZA

Discussion Paper No. 6919
October 2012

IZA

P.O. Box 7240
53072 Bonn
Germany

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

IZA Discussion Paper No. 6919
October 2012

ABSTRACT

Truth-Telling: A Representative Assessment*

A central assumption of the canonical cheap talk literature is that people misreport their private information if this is to their material benefit. Recent evidence from laboratory experiments with student subjects suggests, however, that while many people do report the payoff-maximizing outcome, some report their private information truthfully or at least do not lie maximally. We measure truth-telling outside the laboratory by calling a representative sample of the German population at home. In our setup, participants have a strong monetary incentive to misreport, misreporting cannot be detected, and reputational concerns are negligible. Yet, we find that aggregate reporting behavior closely follows the expected truthful distribution. Our results underline the importance of lying costs and raise questions regarding the influence of the decision-making environment and the elicitation mode on reporting behavior.

JEL Classification: C93, D01, D82, D83

Keywords: private information, cheap talk, honesty, lying costs, representative experiment

Corresponding author:

Johannes Abeler
University of Oxford
Department of Economics
Manor Road
Oxford OX1 3UQ
United Kingdom
E-mail: johannes.abeler@economics.ox.ac.uk

* Financial support from the German Science Foundation (DFG) through SFB/TR 15, the Bonn Graduate School of Economics, the European Research Council (Starting Grant), and CESifo is gratefully acknowledged. We thank Steffen Altmann, Vince Crawford, Eddie Dekel, Holger Gerhardt, Sebastian Kube, Theo Offerman, Paul Schempp, Matthias Wibral, and Florian Zimmermann for helpful discussions. Valuable comments were also received from numerous seminar and conference participants.

1 Introduction

A potential buyer of a car stating his willingness to pay, a self-employed shopkeeper reporting her income to the tax authorities, an employee telling his boss that he worked the entire weekend: situations of asymmetric information are ubiquitous. How people report their private information is thus of fundamental importance for many areas in economics.

Economic theory usually assumes that agents are rational and behave fully strategically when reporting their type (e.g., Crawford & Sobel 1982): agents therefore misreport their private information if this is to their immediate benefit. In contrast, telling the truth is one of the most important norms, postulated by all moral systems and world religions.[1] Deviating from this norm will evoke disutility or a psychological cost for people who value this norm and could therefore make them misreport less than predicted by standard theory. Lying could also cause disutility for other reasons, e.g., because utility is derived from identity (Akerlof & Kranton 2000, Bénabou & Tirole 2011) or from self-reputation (similar to Bénabou & Tirole 2006). A recent theoretical literature is built on the assumption that agents face such a (psychological) lying cost when misrepresenting their type (e.g., Kartik et al. 2007, Chen et al. 2008, Kartik 2009, Saran 2011). These models capture the intuition that people are sometimes reluctant to lie. Yet, it remains an open empirical question how people actually report their private information. Using observational field data to understand reporting behavior is notoriously difficult since an individual's true type is usually unknown and misreporting cannot be identified. Even the reports are often not observed or are the consequence of equilibrium behavior which makes it difficult to deduce individual motives.

The cleanest evidence so far on how people report their private information comes from experiments conducted in tightly controlled laboratory situations. A robust result is that many subjects misreport their private information to their own advantage but that a substantial share of subjects refrain from reporting the payoff-maximizing type and that some are fully honest (e.g., Gneezy 2005, Charness & Dufwenberg 2006, Fischbacher & Heusi 2008, de Haan et al. 2011, Houser et al. 2011, Shalvi et al. 2011, Wibral et al. forthcoming, Serra-Garcia et al. forthcoming). These studies are a strong first indicator that lying costs do influence behavior. However, lab experiments do not allow for inferences with respect to the prevalence

[1] In few exceptional circumstances, some moral systems allow lying, e.g., if it serves a greater cause. In this paper, however, we are concerned with the much more common situation in which only the person who lies would benefit from misreporting.

of lying costs in the overall population since they have been conducted with student samples only (DellaVigna 2009, Falk & Heckman 2009). Also, decision making took place in an austere laboratory environment that is not representative of an everyday decision-making context, rendering generalizations to the level of behavior elsewhere more difficult. It could thus be that there are systematic differences in behavior of students in the laboratory and of non-student subjects outside the lab.

In this paper, we measure how people report their private information outside the laboratory by calling participants on the phone at their home. Participants were drawn randomly from the German population, yielding a representative sample. An incentivized experiment was embedded in the interview. The experimental setup is related to the design of Fischbacher & Heusi (2008) and Bucciol & Piovesan (2011) and is extremely simple: participants were asked to toss a coin and report their type, i.e., either "heads" or "tails". Reporting tails yielded a payoff of 15 euros, which participants could choose to receive in cash or as an Amazon gift certificate, while reporting heads yielded a payoff of zero. Participants thus had a clear monetary incentive to report tails regardless of their true type. It was obvious that the true outcome was only known to the participants, as they tossed the coin privately at home. In this setup, we cannot draw reliable conclusions about the truthfulness of any individual report. But we can learn about aggregate behavior by comparing the distribution of reports to the true distribution of a fair coin (50 percent tails) and to the payoff-maximizing distribution (100 percent tails). This indirect observation therefore allows us to study the behavior of subjects in a situation in which private information is kept truly private and in which subjects do not face any risk of detection.[2] Moreover, the decision is non-strategic and reputational concerns are minimized as the interviewer is a stranger with whom no future interaction can be expected.[3]

[2]In other studies concerning how people report their private information (e.g., Gneezy 2005, Charness & Dufwenberg 2006) the experimenter knows or will later know the subject's true type (and the subject is aware of this) and can thus judge whether an individual was honest or not. In our experiment, only the participant knows his or her private information. Our setup is thus closer to situations in which information is truly private and only known by the individual, while Gneezy's and Charness & Dufwenberg's setup is more representative of situations in which the private information is known by more than one person, e.g., when filing a joint tax declaration. These papers are also interested in the interaction between sender and receiver, from which we abstract.

[3]Fischbacher & Heusi (2008) have conducted direct tests of the influence of reputation in this experimental paradigm. Subjects in their baseline condition rolled a die, reported their number, and were paid their report; in a double-blind control treatment, subjects were provided with enough money for the maximal report, took whatever they decided to take, put the remaining money in an envelope, and dropped the envelope into a letter box when leaving. The letter box was only opened after all subjects had left. A report could thus not be traced to any individual subject,

If all our participants were rational money maximizers, we would expect that all of them reported tails. If behavior on the phone was similar to previous, comparable laboratory experiments (e.g., Houser et al. 2011), we would expect about 75 percent of subjects reporting tails.

In contrast to these predictions, we find that almost all participants report their type honestly. If anything, participants report the payoff maximizing outcome less often than expected under truthful reporting. This latter effect, however, is small and disappears in a second treatment in which participants were asked to report the total number of tails in four consecutive coin tosses and received 5 euros times the number of reported tails. The resulting distribution of reports in the 4-coin treatment is indistinguishable from the distribution under complete truth-telling. Moreover, while previous studies (e.g., Dreber & Johannesson 2008) have found correlations between individual characteristics, like gender, and truth-telling, we do not find any robust correlations between individual characteristics and reporting behavior. This is not surprising if almost all participants report truthfully. Reports are thus solely determined by chance, namely the coin toss, which cannot be related to any individual characteristic. At the same time, our results show that lying costs are widespread and are influencing behavior regardless of gender, religious beliefs, education, or age: participants forego considerable amounts of money to avoid lying.

Our results add to the doubts that previous laboratory experiments have cast on the assumption of zero lying costs. If anything, lab experiments have understated the importance and extent of lying costs. This suggests that studying the theoretical implications of such costs (as in, e.g., Kartik et al. 2007) is a promising research avenue. The difference in behavior in our study and in previous lab studies further shows how malleable reporting behavior can be and opens many new questions about how exactly reporting private information depends on the decision-making context.[4] Intuitively, different norms might apply when making such a decision at home, representing a private and familiar environment. Similarly, people could be more attentive to their own moral rules, e.g., abstaining from lying when at home. In independent research, Waubert De Puiseau & Glöckner (2012) also find considerable

---

further reducing reputational concerns. Aggregate behavior, however, did not differ between the two treatments suggesting that reputation did not play a role in the baseline treatment.

[4]Previous research comparing behavior of student samples vs. non-students samples and behavior in the lab vs. outside the lab has mostly shown little differences, with some notable exceptions (e.g., Stoop et al. 2010, Falk et al. forthcoming). For an overview and critical discussion, see Falk & Heckman (2009) or Camerer (2011). Our field results differ quite strongly from behavior in comparable lab experiments (e.g., Fischbacher & Heusi 2008), suggesting that truth-telling is more context dependent than other behaviors.

truth-telling at home, though not as extreme as in our data, using an online panel in which participants answered questions at home on a computer. Irrespective of these differences between lab and field, our study establishes that lying costs are more important than previously assumed and are strongly influencing behavior.

In the next two sections, we present the design of the study and our hypotheses. Section 4 contains the results. We discuss policy implications in Section 5.

2 Design

The computer-assisted telephone interviews were operated by the Institute for Applied Social Sciences (infas), a private and well-known German research institute. They were conducted between November 2010 and February 2011.[5] The average interview lasted 20 minutes (standard deviation: 5.5 minutes). Telephone numbers were selected using a random digit dialing technique: numbers were generated randomly based on a data set of all potential telephone numbers in Germany. Only landline numbers were used in this study, as 92 percent of German households have a landline number. The selection of the participant within each household was also random: only the member of the household whose birthday was the most recent among all household members was eligible to participate. We restricted participation to those aged between 18 and 70 years at the time of the interview.

The survey part of the study consisted of questions relating to the participants' socio-demographic background, their economic and political preferences, their current living and financial situation, their religious beliefs, their attitudes towards opportunistic behavior and everyday crime, as well as their belief about other people's behavior in the experiment.

The experimental part of the study was embedded in the survey. Before the experiment started the participant was reminded that the resulting data would be anonymized, and that infas and the University of Bonn guaranteed the correct payment. The interviewer then asked the participant to take a coin and explained the rules of the experiment: the task was to toss the coin and report whether heads or tails came up. If the participant reported heads, they received no payment. If the participant reported tails, they would receive 15 euros, either in cash via regular mail or as an Amazon gift certificate code. The alphanumeric 14-digit gift certificate

[5]The interviews were conducted in the infas telephone studio. Infas ensures a high quality of interviews by supervising interviews randomly. Supervisors are present in the telephone studio at all times and interviews can be monitored without the interviewer noticing this.

code could be received via email or directly on the phone at the end of the interview. Then, the participant was asked to toss the coin and report the outcome. We will call this treatment "1-coin treatment." 658 people participated in this version of our experiment. A translation of the exact experimental instructions can be found in Appendix A.

In a second treatment, 94 people were interviewed and participated in the following variation of the experiment. Participants were asked to take a coin, toss it four times, and report the number of times that tails came up. For each time participants reported tails they received 5 euros. Thus, they could earn 0, 5, 10, 15, or 20 euros. We will call this treatment "4-coin treatment."

3 Hypotheses

The standard economic prediction in our setup is straightforward: depending on the treatment, people will report tails one or four times, respectively. This is the payoff maximizing outcome as there are no exogenous costs linked to misreporting, no possibility of detection and no fines. The setup is extremely simple and subjects should have no trouble identifying the payoff maximizing choice. Moreover, the setup is highly anonymous, discouraging any reputational concerns because of repeated interaction. Since the coin was tossed at the home of the participant, it was obvious that the interviewer could not secretely observe the true outcome of the coin toss. This might be a concern in otherwise similar laboratory experiments but not in our phone study.

If, however, some participants incur a psychological cost or derive direct disutility from falsely reporting their private information per se we should expect both heads and tails to be reported in the experiment. There are a few recent theoretical papers that assume such a cost. For example, Kartik (2009) and Kartik et al. (2007) build on Crawford & Sobel's (1982) cheap-talk model and derive predictions for the case that some agents incur costs when misreporting their private information (see also, e.g., Saran 2011). Assuming some degree of heterogeneity in the incurred costs when misreporting, it is then a question of the trade-off between psychological costs and monetary benefits of misreporting how many participants will report heads and how many report tails.

The two treatments differ in crucial aspects and allow different types of lying costs to influence behavior. On the one hand, marginal incentives to lie are higher in the 1-coin treatment. If lying costs are compared to these marginal incentives,

the 1-coin treatment will favor lying more strongly than the 4-coin treatment. On the other hand, lying costs might be related to self-reputation or identity arguments (e.g., Bénabou & Tirole 2006, Akerlof & Kranton 2000). Participants in the 1-coin treatment have to make a clear, binary choice whether to lie or not; lying in such a setting will impact self-reputation or identity and thus make lying more costly. Participants in the 4-coin treatment can make a finer choice between being fully honest, exaggerating a little bit, or lying maximally; this could render small lies compatible with a positive self-reputation and thus enhance lying (Mazar et al. 2008). Such non-maximal lying has already been shown to be important by Fischbacher & Heusi (2008).

Regarding potential differences in reporting behavior according to individual characteristics, we would expect that women are more honest than men (as already shown by Dreber & Johannesson 2008, Houser et al. 2011). More religious participants would be expected to be more honest, since religious priming leads to less lying and more pro-social behavior (Mazar et al. 2008, Shariff & Norenzayan 2007). Income could be positively correlated with honesty because of the lower marginal utility of the monetary rewards or negatively correlated because of reverse causality. A similarly ambiguous hypothesis can be derived for education or the social environment, e.g., the size of the community or family status.

4 Results

658 people were interviewed and participated in the 1-coin treatment of the experiment. Here, participants tossed the coin once and reported the outcome. A report of tails yielded 15 euros, a report of heads 0 euros. In the 4-coin treatment, they reported the outcome of four consecutive coin tosses. For each reported tails they received 5 euros. 94 people participated in this treatment.

> **Result 1:** *In the 1-coin treatment, the distribution of actual reports is very close to the truthful distribution; participants report the payoff-maximizing outcome slightly less often than expected if everyone reported truthfully. In the 4-coin treatment, the distribution of reports is indistinguishable from the truthful distribution.*

Figure 1 illustrates aggregate behavior (the dashed line corresponds to the expected distribution if every participant reported the true outcome of the coin toss). 55.6 percent of participants report heads as the outcome of the coin toss, yielding a payoff

of zero, the remaining participants report tails yielding a payoff of 15 euros. The payoff-maximizing outcome is reported slightly less often than in 50 percent of the cases and although the difference is small in terms of effect size, it is significant (Binomial test, $p < 0.01$).[6] Figure 2 shows aggregate behavior in the 4-coin treatment. Again, reporting behavior follows the expected distribution under complete honesty extremely closely (the dashed line corresponds to the truthful distribution). In fact, the distribution of reported outcomes is statistically indistinguishable from the truthful distribution (Kolmogorov-Smirnov test, $p = 0.61$; binomial tests of the expected against the observed frequency, all five $p > 0.13$) and in particular, there is no significant over-reporting of zero. Looking at behavior in both treatments we can therefore summarize that the payoff-maximizing outcome is reported by much fewer participants than expected if no one incurred lying costs. It is also reported less often than suggested by previous lab experimental studies, which find some truth-telling but also many instances of the payoff-maximizing report. Instead, it is close to the distribution that would arise if every participant reported his or her type truthfully. Overall, we therefore conclude that most participants in our experiment reported truthfully in the aggregate.

Previous studies have shown that truth-telling does correlate with observable characteristics, e.g. gender or religiosity (Dreber & Johannesson 2008, Houser et al. 2011, Mazar et al. 2008, Shariff & Norenzayan 2007). In contrast, if our conjecture that almost all participants report truthfully is correct, an individual's reported outcome will only be driven by their random coin toss; if this is the case, reporting cannot be correlated with any individual characteristic, as these are orthogonal to the chance move. Therefore, if we do not find such a correlation, our finding of (almost) complete honesty is corroborated.[7] More specifically, we conduct regression

[6]We can only speculate about why some people obviously falsely claimed to be of the payoff minimizing type. The design of the experiment allows to rule out reputational concerns towards the interviewer as an important factor. Privacy concerns could drive this effect: reporting the type that gives zero payoff makes it unnecessary to hand over one's address. Note that we do not find any such effect in the 4-coin treatment in which the option of reporting zero is less salient than in the 1-coin treatment. However, we ensured that privacy concerns were minimized in both treatments by giving participants the opportunity to receive the payment as a gift certificate code directly via the phone. Choosing this payment mode made it unnecessary to hand over any additional contact details. Another possibility would be self-image concerns: refraining from easily and safely earning 15 euros could be a strong signal to oneself that one is not greedy and thereby flattering for one's self-image.

[7]It could still be that a subgroup of people, which we cannot identify with our background information, reports tails more often than actually true while another subgroup reports tails less often. This could result in the two effects offsetting each other, which would result in a similar picture of aggregate behavior. However, we consider this to be extremely unlikely as our analysis shows that this is not the case for any of the numerous subgroups that we can identify with our data. Such an effect would further need to recreate the distinct distributions of Figures 1 and 2,

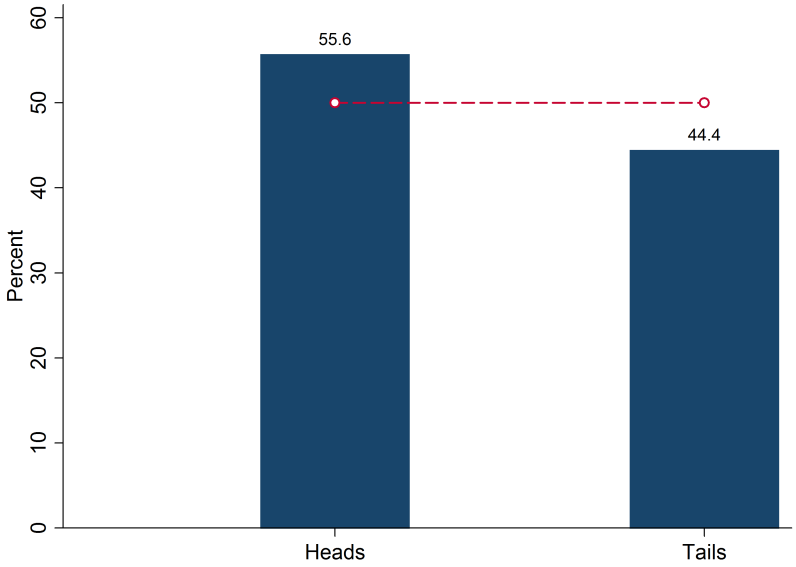

Figure 1: Aggregate Behavior in the 1-coin Treatment. Reporting heads yielded no payoff; reporting tails yielded a payoff of 15 euros. The dashed line corresponds to the expected distribution if every participant reported the true outcome of their coin toss.

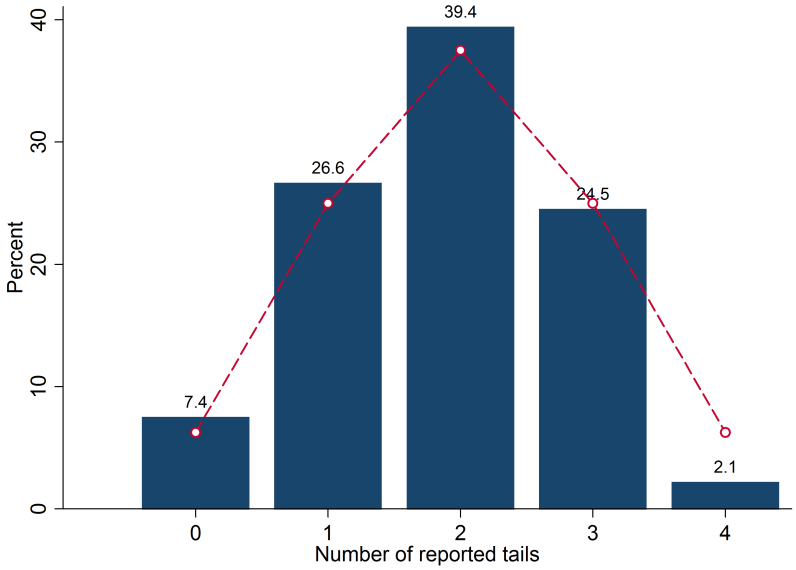

Figure 2: Aggregate Behavior in the 4-coin Treatment. The payoff was 5 euros times the number of tails reported. The dashed line corresponds to the expected distribution if every participant reported the true outcomes of their coin tosses.

analyses for the two experiments in order to examine whether there are systematic effects of individual characteristics on reporting behavior. First, we regress the report only on clearly exogenous variables such as age and gender, in a second step adding religious denomination. We then include income, the size of the city the individual lives in, and education dummies. Finally, we look at the effect of an individual's religiousness (interacted with denomination), their risk and trust preferences, and their belief about the reporting behavior of other participants.

> **Result 2:** *There is no significant correlation between reporting behavior and any individual characteristic.*

First, we look for potential group differences in terms of reporting behavior in the 1-coin treatment by conducting Probit regressions of the reported outcome on the respective characteristics (see Table 1). No characteristic except for one's belief about others' behavior is significantly associated with reporting in the experiment: participants who think many other participants report tails dishonestly, are *less* likely to report tails themselves. This belief is, however, not significant if we include it as the only explanatory variable ($p = 0.15$). Note in particular that neither gender nor any religion-related variable is significantly correlated with reporting. Conducting the same regressions as in Table 1 using OLS leaves the results unchanged. Next, we check whether these results also hold in the 4-coin treatment. We run Ordered Logit regressions of the reported number of tails on the same explanatory variables as before. Table 2 illustrates the results from this estimation. Only the coefficient for trust is significant. This effect is, however, not robust to the inclusion or exclusion of other explanatory variables. The effect is also not present in the 1-coin treatment. In contrast to the 1-coin treatment, the belief coefficient shows no significant association with reporting behavior in this treatment and the point estimate has the opposite sign.

When we elicited beliefs about the behavior of other participants, most people stated that they believed many people to make a false report in the experiment. Participants in the 1-coin treatment believed that on average 46.7 percent (std. dev.: 28.5) of all participants would report tails when they actually had tossed heads (this would translate into a 73.4 percent share of tails if all tails are reported as tails). In the 4-coin treatment participants believed that, on average, 43.7 percent (std. dev.: 28.4) of all other participants reported higher types than actually true and that 28.1 percent (std. dev.: 26.3) reported four times tails, the payoff-maximizing report. Thus, participants' beliefs strongly deviate from observed behavior. It is thus

which is again very unlikely.

Table 1: Covariates in 1-Coin Experiment (Probit Regressions)

Dependent Variable: 1 if Reported Tails

	(1)	(2)	(3)	(4)	(5)
Age	0.002	0.002	0.002	0.002	0.001
	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
1 if Female	0.041	0.040	0.079*	0.065	0.066
	(0.039)	(0.039)	(0.047)	(0.048)	(0.049)
1 if Protestant		0.010	0.021	-0.038	-0.084
		(0.046)	(0.056)	(0.140)	(0.141)
1 if Catholic		0.006	-0.004	-0.004	-0.064
		(0.048)	(0.057)	(0.150)	(0.152)
Income			0.002	0.003	0.004
			(0.005)	(0.005)	(0.005)
1 if Professional Education			0.041	0.053	0.056
			(0.062)	(0.062)	(0.063)
1 if Academic Education			-0.016	-0.009	-0.015
			(0.055)	(0.056)	(0.057)
City Size			0.000	0.000	0.000
			(0.000)	(0.000)	(0.000)
Religiousness				0.029	0.023
				(0.020)	(0.020)
Religiousness*Catholic				-0.010	0.004
				(0.035)	(0.036)
Religiousness*Protestant				0.005	0.013
				(0.034)	(0.035)
Risk Tolerance					-0.003
					(0.018)
Trust					-0.028
					(0.018)
Belief about other					-0.218**
Participants					(0.092)
N.Obs.	658	658	465	464	454

Notes: Probit estimates. Marginal effects are shown, robust standard errors are in parentheses. "Risk tolerance" is the answer to the general risk question of the GSOEP ("How do you consider yourself? Are you in general a rather risk-loving person, or do you try to avoid risks? Use a scale from 1, meaning that you are not at all willing to take risks, to 7, meaning that you are absolutely willing to take risks.") and "trust" is the answer to the WVS trust question ("Generally speaking: Do you think one can trust other people, or that one should rather be careful when dealing with other people? Please indicate your answer on a scale from 1 to 7, with 1 meaning that one should be careful when dealing with other people, and 7 meaning that one can trust other people."). "Belief about other participants" is the belief of this participant about the share of other participants who reported "tails" while actually "heads" came up. Significance at the 1, 5, and 10 percent level is denoted by ***, **, and *, respectively.

Table 2: Covariates in 4-Coin Experiment (Ordered Logit Regressions)

Dependent Variable: Number of Reported Tails (0–4)

	(1)	(2)	(3)	(4)	(5)
Age	-0.015	-0.017	0.006	0.008	-0.013
	(0.012)	(0.012)	(0.020)	(0.023)	(0.027)
1 if Female	0.120	0.159	0.177	0.072	0.075
	(0.389)	(0.399)	(0.496)	(0.526)	(0.573)
1 if Protestant		-0.099	0.073	-1.052	-1.588
		(0.473)	(0.622)	(1.258)	(1.269)
1 if Catholic		0.292	0.703	-0.407	-1.425
		(0.452)	(0.565)	(1.373)	(1.520)
Income			0.039	0.042	0.025
			(0.176)	(0.184)	(0.211)
1 if Professional Education			-0.591	-0.721	-0.759
			(0.587)	(0.620)	(0.693)
1 if Academic Education			-0.640	-0.757	-0.764
			(0.697)	(0.693)	(0.729)
City Size			-0.002	-0.002	-0.002
			(0.002)	(0.002)	(0.002)
Religiousness				-0.271	-0.289
				(0.238)	(0.250)
Religiousness*Catholic				0.328	0.475
				(0.318)	(0.345)
Religiousness*Protestant				0.336	0.395
				(0.283)	(0.276)
Risk Tolerance					-0.324
					(0.251)
Trust					0.491**
					(0.215)
Belief about other					0.715
Participants					(1.311)
N.Obs.	94	94	62	62	60

Notes: Ordered Logit estimates. Robust standard errors are in parentheses. "Risk tolerance" is the answer to the general risk question of the GSOEP ("How do you consider yourself? Are you in general a rather risk-loving person, or do you try to avoid risks? Use a scale from 1, meaning that you are not at all willing to take risks, to 7, meaning that you are absolutely willing to take risks.") and "trust" is the answer to the WVS trust question ("Generally speaking: Do you think one can trust other people, or that one should rather be careful when dealing with other people? Please indicate your answer on a scale from 1 to 7, with 1 meaning that one should be careful when dealing with other people, and 7 meaning that one can trust other people."). "Belief about other participants" is the belief of this participant about the share of other participants who reported four times "tails" while actually "heads" came up at least once. Significance at the 1, 5, and 10 percent level is denoted by ***, **, and *, respectively.

even more surprising that people decided to refrain from exploiting the opportunity to receive a considerable amount of money when they believed that many others would do so (see, e.g., López-Pérez 2010, López-Pérez 2012).

Two further aspects of our analysis are worth noting. First, when running OLS regressions using the same predictor variables as above, we find that only one of the 10 specifications has an adjusted R^2 above 0 (at 0.0146), all other adjusted R^2 values are negative. Moreover, the resulting adjusted R^2 tend to decrease in the number of included variables. This again underlines our conclusion: the tested predictor variables do not increase explained variance in the dependent variable compared to pure chance. Second, we also tested the correlation between reported number and answers to the survey questions that we did not include in the main specifications of Tables 1 and 2. These include a person's citizenship and country of birth, various personal characteristics, a person's current job situation and their current or recent position in the professional hierarchy, a person's willingness to tell white lies in different situations, a person's family status and living situation (whether one lives with a partner and the number of people belonging to the household), the frequency of church attendance, a person's political preference, and the individual's tendency to behave in an opportunistic way as well as the belief about others' willingness to behave like that. Testing these variables as predictors in Probit and Ordered Logit regressions in the two different data sets, akin to Tables 1 and 2, we find no robust association between any of them and reporting behavior. Summing up, the overall picture is confirmed: no individual characteristic, whether exogenous or endogenous, is systematically associated with reporting behavior suggesting that almost all participants in our study tell the truth.

5 Conclusion

Using a representative sample of the German population we conducted telephone interviews during which respondents participated in an incentivized experiment. Depending on the treatment, they could earn money by reporting tails as the outcome of one or four coin tosses. We find that almost all participants report their coin toss(es) honestly: the distributions of reports are extremely close to the true distribution of a fair coin toss or four coin tosses, respectively. Moreover, reports are not correlated with any individual characteristic, including gender which has been shown in previous lab studies to predict honesty.

Our results strongly underline doubts about the generalizability of economic

models assuming that people always lie maximally when it is financially beneficial and when they have the possibility to do so. Apparently, people do not only care for the trade-off between expected costs (caused by some form of external punishment when making a false claim and being detected) and expected benefits (namely the monetary gain from reporting the maximal type without being caught). Our results reinforce the results of previous lab experimental research that many people face an intrinsic aversion to misreporting their private information, i.e., a lying cost. If anything, these lying costs are stronger in our setting outside the lab. A fruitful avenue for future research would therefore be to explore the implications of lying cost models in different settings, e.g., how lying costs and concerns for reputation interact. It is also not clear what drives such a lying cost. It could be based on a concern for self-reputation (Bénabou & Tirole 2006) or come from social norms (e.g., Pruckner & Sausgruber forthcoming, Abeler 2012) or preferences over identity (Akerlof & Kranton 2000, Bénabou & Tirole 2011).

At the same time, this study does not imply that everybody always reports their private information truthfully. The level of lying costs seems to interact with the context in which people are asked to report their type (see also Mazar & Ariely 2006, Mazar et al. 2008). While we find that almost everyone reports the outcome of the coin toss truthfully in the telephone study, previous studies run with student subjects in the laboratory find that a substantial share of participants report higher types or lie maximally. But even the lab experiments find evidence for truth-telling or non-maximal lying. The difference in behavior on the phone and in the lab shows how malleable reporting behavior can be.

Our results therefore point to important policy implications: institutions, like tax authorities or insurance companies, could make use of the context dependence of reporting behavior when designing decision-making environments. As we find strong evidence for widespread lying costs, appropriate mechanisms might be much less complex than those resulting when assuming that agents have no qualms about lying. It might be possible to change reporting behavior in simple and low-cost ways in the spirit of libertarian paternalism (Thaler & Sunstein 2003). Further research is necessary to uncover what the crucial aspects of the decision-making environment are that induce truth-telling.

Methodologically, our findings suggest that survey responses are much more reliable than previously assumed: if people report the truth even when there is a strong monetary incentive to lie, there is less reason to assume that they lie when there are no, or only non-monetary, incentives to do so.

6 References

Abeler, J. (2012), 'Preferences for truth-telling', *mimeo*.

Akerlof, G. & Kranton, R. (2000), 'Economics and identity', *The Quarterly Journal of Economics* **115**(3), 715–753.

Bénabou, R. & Tirole, J. (2006), 'Incentives and prosocial behavior', *American Economic Review* **96**(5), 1652–1678.

Bénabou, R. & Tirole, J. (2011), 'Identity, morals, and taboos: Beliefs as assets', *Quarterly Journal of Economics* **126**, 805–855.

Bucciol, A. & Piovesan, M. (2011), 'Luck or cheating? A field experiment on honesty with children', *Journal of Economic Psychology* **32**(1), 73–78.

Camerer, C. (2011), 'The promise and success of lab-field generalizability in experimental economics', *SSRN Discussion Paper*.

Charness, G. & Dufwenberg, M. (2006), 'Promises and partnership', *Econometrica* **74**(6), 1579–1601.

Chen, Y., Kartik, N. & Sobel, J. (2008), 'Selecting cheap-talk equilibria', *Econometrica* **76**(1), 117–136.

Crawford, V. P. & Sobel, J. (1982), 'Strategic information transmission', *Econometrica* **50**(6), 1431–1451.

de Haan, T., Offerman, T. & Sloof, R. (2011), 'Money talks? An experimental investigation of cheap talk and burned money', *Discussion Paper*.

DellaVigna, S. (2009), 'Psychology and economics: Evidence from the field', *Journal of Economic Literature* **47**(2), 315–372.

Dreber, A. & Johannesson, M. (2008), 'Gender differences in deception', *Economics Letters* **99**(1), 197–199.

Falk, A. & Heckman, J. J. (2009), 'Lab experiments are a major source of knowledge in the social sciences', *Science* **326**(5952), 535–538.

Falk, A., Meier, S. & Zehnder, C. (forthcoming), 'Do lab experiments misrepresent social preferences? The case of self-selected student samples', *Journal of the European Economic Association*.

Fischbacher, U. & Heusi, F. (2008), 'Lies in disguise – An experimental study on cheating', *Discussion Paper*.

Gneezy, U. (2005), 'Deception: The role of consequences', *American Economic Review* **95**(1), 384–394.

Houser, D., Vetter, S. & Winter, J. K. (2011), 'Fairness and cheating', *GMU Working Paper in Economics No. 11-02*.

Kartik, N. (2009), 'Strategic communication with lying costs', *Review of Economic Studies* **76**(4), 1359–1395.

Kartik, N., Ottaviani, M. & Squintani, F. (2007), 'Credulity, lies, and costly talk', *Journal of Economic Theory* **134**(1), 93–116.

López-Pérez, R. (2010), 'Guilt and shame: An axiomatic analysis', *Theory and Decision* **69**, 569–586.

López-Pérez, R. (2012), 'The power of words: A model of honesty and fairness', *Journal of Economic Psychology* **33**(3), 642–658.

Mazar, N., Amir, O. & Ariely, D. (2008), 'The dishonesty of honest people: A theory of self-concept maintenance', *Journal of Marketing Research* **45**(6), 633–644.

Mazar, N. & Ariely, D. (2006), 'Dishonesty in everyday life and its policy implications', *Journal of Public Policy & Marketing* **25**(1), 117–126.

Pruckner, G. & Sausgruber, R. (forthcoming), 'Honesty on the streets: A field study on newspaper purchasing', *Journal of the European Economic Association*.

Saran, R. (2011), 'Bilateral trading with naive traders', *Games and Economic Behavior* **72**(2), 544–557.

Serra-Garcia, M., van Damme, E. & Potters, J. (forthcoming), 'Lying about what you know or about what you do', *Journal of the European Economic Association*.

Shalvi, S., Dana, J., Handgraaf, M. J. & De Dreu, C. K. (2011), 'Justified ethicality: Observing desired counterfactuals modifies ethical perceptions and behavior', *Organizational Behavior and Human Decision Processes* **115**(2), 181–190.

Shariff, A. F. & Norenzayan, A. (2007), 'God is watching you: Priming god concepts increases prosocial behavior in an anonymous economic game', *Psychological Science* **18**(9), 803–809.

Stoop, J., Noussair, C. & van Soest, D. (2010), 'From the lab to the field: Cooperation among fishermen', *MPRA Discussion Paper* .

Thaler, R. H. & Sunstein, C. R. (2003), 'Libertarian paternalism', *American Economic Review* **93**(2), 175–179.

Waubert De Puiseau, B. & Glöckner, A. (2012), 'Investigating cheating behavior in students compared to the general public', *mimeo*.

Wibral, M., Dohmen, T., Falk, A., Klingmüller, D. & Weber, B. (forthcoming), 'Testosterone administration reduces lying in men', *PLOS One*.

A Experimental Instructions

These are the instructions for the 1-coin experiment translated into English. The instructions for the 4-coin treatment were adapted according to the rules of the 4-coin treatment but otherwise identical.

The experiment's purpose is without any exception scientific, and all legal regulations of data protection are strictly respected and the anonymity of the data analysis is fully warranted. Your participation in the experiment is, of course, on a voluntary basis. Infas and the University of Bonn guarantee the correct and proper payment of the achieved amount of money.

In case the participant had doubts about any part of the experiment he or she was given the phone number of the supervisor of the study at Infas, so that he or she could make sure that everything was legitimate.

You can receive the payment in cash via regular mail (we will then need your address), or we will give you a gift certificate code which equals the respective amount via email or via phone. The gift certificate is valid for a purchase at Amazon. The address will only be used for mailing the payment of the money you earned and will be deleted from our database afterwards. You will need a coin to take part in the experiment (any coin). Please get the coin first. I will then explain the details of the experiment.

Now, I will explain the experiment. Your coin has one side showing a number, and another one, often showing some image. After receiving the instructions, I will ask you to toss the coin, and you will have to tell me afterwards which side of the coin came up. You receive 15 euros if "tails" comes up.[8] You receive 0 euros if "heads" comes up.

The interviewers were instructed to react with the following sentence to any comments or questions that could possibly influence the participants decision: "I can only give you the instructions. What side of the coin you report to me, is completely up to you." They were also reminded, to react neutrally towards any comment by the participant concerning the number they reported, and to refrain from making any comments or jokes.

Once again: You toss the coin and tell me which side came up. If "tails" came up, you receive 15 euros. Please don't start just yet. Do you understand the rules?

If participant said that he or she understood the rules, the interviewer continued

[8]The German word for "tails" is "Zahl", literally "number".

with the experiment.

Please put aside your telephone receiver (but please do not hang up), and toss the coin. Afterwards please tell me which side of the coin came up. Please toss the coin now.

In case the participant hung up by accident he or she was called again immediately.

If necessary, the interviewer read the following question: Which side came up?