Observability, Honesty, and the Social Image Costs of Lying *

Ciril Bosch-Rosa[1], Levent Neyse[2], and Daniele Nosenzo[3]

[1]Chair of Macroeconomics, Technische Universität Berlin, Germany
[2]WZB, Berlin and DIW, Berlin, Germany
[3]Aarhus University, Denmark

January 2, 2025

Abstract

We study the role of social image in influencing lying behavior through a pre-registered within-subject experiment embedded in the 2020 wave of the German Socio-Economic Panel Innovation Sample (SOEP-IS). By exogenously manipulating the observability of lying across two tasks, we explore how individuals respond to increased image costs of lying. By exploiting the rich comprehensive socio-demographic and psychological data from the SOEP-IS, we study how this response varies across substrata of the population. Our findings indicate that men and former citizens of East Germany (GDR) display a stronger aversion to lying under observable conditions. These results highlight the variability in image costs across demographic groups and underscore the importance of historical and cultural contexts in shaping ethical behaviors.

Keywords Lying, Social Image, Experiment, Survey

JEL Classification C91 · D84 · G11 · G41

*Corresponding author: cirilbosch@gmail.com. We thank for their helpful comments Urs Fischbacher and audiences at Aarhus University, BGSE Summer Forum, ESA World Meeting 2023, FAIR-NHH, IMT Lucca, University of Glasgow, University of Helsinki, WU Vienna. We also thank Martin Gerike for his help in preparing and managing the data. Bosch-Rosa and Neyse acknowledge the financial support from the Deutsche Forschungsgemeinschaft (DFG) through the CRC TRR 190 Rationality and Competition. Neyse acknowledges the financial support of the Leibniz Association (SAW10868 & K523/2023-*Lab*^{2}). Nosenzo acknowledges the financial support from the Aarhus University Research Foundation (AUFF Starting Grant 36835). The DIW Research Ethics Board approves the contents surveyed in the Socio-Economic Panel.

1 Introduction

How individuals report private information is a matter of substantial economic importance that has attracted considerable attention from theoretical and empirical scholars. Contrary to standard economic theory, a recent body of literature has shown that individuals are willing to forego significant material gains to honestly disclose private information (e.g., Fischbacher and Föllmi-Heusi, 2013; Jacobsen et al., 2018; Abeler et al., 2019; Gerlach et al., 2019). For example, in a meta-analysis of 90 experimental papers, Abeler et al. (2019) estimate that individuals forgo, on average, three-quarters of the potential gains from lying. Their analysis reveals that aversion to lying is mainly driven by two psychological costs: an intrinsic aversion to reporting false information ('intrinsic lying cost') and an aversion to being perceived as a liar ('image cost'). Other studies have also reached similar conclusions (e.g., Gneezy et al., 2018; Dufwenberg and Dufwenberg, 2018; Khalmetski and Sliwka, 2019).

Despite these advances in understanding the costs driving lying aversion, we still know little about the forces that shape such costs and their prevalence across different segments of society. Is the concern of being perceived as a liar common across people of varying socio-demographic backgrounds? Or is it more pronounced among individuals with specific personal or socio-economic characteristics, such as gender, income, or political affiliation? Addressing these questions is important for a number of reasons. First, by assessing the degree of heterogeneity in lying aversion, we can improve our understanding of how individuals report private information. Second, uncovering heterogeneity in the susceptibility to intrinsic or image-related lying costs may have important practical implications, as it could imply different approaches to tackle dishonesty across diverse population groups (e.g., stronger use of reputation mechanisms in groups more concerned with image costs). Finally – and perhaps more speculatively –, revealing the individual traits and attritbutes that correlate with lying costs may provide insights into the forces that determine these costs (e.g., if image costs vary systematically with socio-economic characteristics, this may suggest the social environment plays an important role in shaping image concerns).[1]

[1] See Abeler et al. (2024) for recent evidence on the causal influence of the social environment on lying behavior.

In this paper, we begin to address these questions by including a pre-registered lying experiment in the 2020 wave of the German Socio-Economic Panel Innovation Sample (SOEP-IS). Our experiment uses a within-subject manipulation to exogenously increase the image costs of lying. We study how this manipulation affects the lying behavior of individuals with diverse socio-demographic and personality profiles, as measured with survey instruments administered to our participants in concurrent and previous waves of the SOEP-IS panel.

Specifically, in our experiment participants take part in two information-reporting tasks. The first task is an adaptation of the classic die-rolling paradigm introduced by Fischbacher and Föllmi-Heusi (2013). In this task, participants privately observe the realization of a uniformly distributed 10-state random variable and then report it to the researcher. Participants are paid based on the state they report and maximize their payoff by reporting the highest payoff state, regardless of the true realization. Because the researcher cannot verify the true realization observed by the participants, we refer to this task as the 'unobservable' task. The second task follows the paradigm developed by Gneezy et al. (2018). In it, participants observe the realization of another 10-state uniformly distributed random variable and report it to the researcher. Again, participants are paid according to the state they report. However, in contrast to the unobservable task, the researcher can now observe *both* the state drawn and the report. In other words, the researcher can observe whether participants lie. We refer to this as the 'observable' task.

The key difference between the two tasks is that the observable task exogenously increases the likelihood that the researcher identifies dishonest reports, thereby raising participants' image cost of lying vis-á-vis the researcher.[2] Previous research has shown that increasing observability and image costs reduces dishonesty (e.g., Gneezy et al., 2018; Abeler et al., 2019; Crede and von Bieberstein, 2020; Fries et al., 2021; Huber et al., 2023; Parra, 2024).[3] We use our experimental design to measure how the exogenous variation in image costs influences the honesty of individuals who differ across a wide range of socio-

[2]It is important to stress that the material costs of dishonesty are *identical* across both tasks, as participants are paid solely based on their reports, irrespective of their honesty. In addition, participants are guaranteed anonimity and confidentiality of their responses, ruling out concerns about possible material repercussions of lying in the experiment.

[3]In fact, previous studies have also shown that merely varying the *perception* of observability, without actually making behavior more observable, is sufficient to reduce dishonesty, see e.g., Abeler et al. (2014); Gneezy and Kajackaite (2017); Lilleholt et al. (2020); Mol et al. (2020); Basic and Quercia (2022).

demographic and psychological characteristics through a difference-in-differences analysis.

To illustrate our approach, take the case of gender. Prior research has shown that women are less likely to lie in the unobservable task (e.g., Dreber and Johannesson, 2008; Childs, 2012; Capraro, 2018; Gneezy et al., 2018; Abeler et al., 2019). This suggests a difference in lying costs between men and women. However, it is unclear whether these differences arise from intrinsic or image costs. Our approach allows us to assess the extent to which image costs are responsible for these differences: if men reduce their lying relatively more than women in the observable task, this would indicate that image concerns play a relatively larger role for men than for women in lying tasks. Conversely, if the gender gap increases, we would infer that image costs are relatively more important for women than for men. If the gap remains unchanged, we would conclude that image costs carry the same weight for men and women.

We report two main results. First, we find significant evidence of lying across both tasks as participants systematically underreport lower-paying states and overreport higher-paying states. Relative to the expected report of 4.50 under truth-telling, the average report is 5.26 in the unobservable task and 4.95 in the observable task. These small deviations are consistent with Abeler et al. (2014), who report low levels of lying in a representative sample of the German population. Importantly, the differences between tasks are statistically significant, confirming participants' stronger aversion to lying in the observable tasks and highlighting the key role of image costs in truth-telling.

Second, we find that image costs are generally invariant across most individual characteristics we consider (age, education, income, employment status, religiosity, political orientation, risk attitude, patience, cognitive ability, interpersonal trust, and Big Five personality traits), with two exceptions. The first exception is gender, where we find that observability affects men's willingness to report high-paying numbers more than it affects women's. The second exception is citizenship in the former German Democratic Republic (GDR or East Germany). We find that image costs loom larger for subjects who lived in the GDR before German reunification. While former GDR citizens report significantly higher states in both tasks, the response to observability is stronger among former GDR citizens.

Our paper makes several contributions to the literature. First, we contribute to the broad literature on lying in two distinct ways. From a theoretical perspective, our study speaks to the recent spur of models that formalize the preferences for truth-telling (e.g., Gneezy et al., 2018; Dufwenberg and Dufwenberg, 2018; Abeler et al., 2019; Khalmetski and Sliwka, 2019). While these models typically allow for individual heterogeneity in the image and intrinsic costs of lying, they do not formalize how these costs might systematically depend on characteristics of the decision-maker. Our analysis suggests that disregarding individual-level heterogeneity may be a reasonable assumption for many, but not all, individual characteristics and traits, as we find evidence of variation in image-based lying costs related to gender and exposure to the socio-political environment.

Second, our results contribute to a better understanding of the factors that encourage or discourage lying behavior. In particular, the influence of citizenship in the GDR highlights the critical role of social factors in shaping the perceived costs of lying.[4] This chimes in with recent work by Abeler et al. (2024), who show the role of educational interventions in shaping children's preferences for honesty and aligns with previously documented socio-cultural differences among individuals raised in varying political systems (e.g., Algan and Cahuc, 2010; Alesina and Giuliano, 2015), especially the stark contrasts between West and East Germany (e.g., Ockenfels and Weimann, 1999; Alesina and Fuchs-Schündeln, 2007), characterized by mass surveillance conducted by the *Staatssicherheit* of the Ministry of State Security (Dennis and Laporte, 2014; Fulbrook, 2014; Lichter et al., 2021). In this line, Schudy et al. (2024) use data from a more recent SOEP-IS module and show that exposure to schooling in East Germany correlates with image-related lying concerns, emphasizing the importance of early life experiences in shaping the costs of dishonesty.

Third, our paper contributes to the behavioral and experimental literature studying the preferences of the general population (e.g., Falk et al., 2018; Cappelen et al., 2022; Holmen et al., 2023). We confirm that the prevalence of dishonest behavior appears to be relatively low in general population samples, especially when compared to student samples (see, Abeler et al. (2014) or Abeler et al. (2019)).

Finally, our results also contribute to more applied research and policy. Specifically, the observed heterogeneity in image costs indicates that there may be room for tailoring interventions to reduce dishonesty within specific demographic groups. For instance,

[4]This conclusion may also apply to gender, to the extent that it reflects nurture rather than nature (see, e.g., Booth and Nolen, 2012)

interventions that increase transparency and (real or perceived) observability may be particularly beneficial in contexts where men are the target of the intervention. More broadly, these findings highlight the importance that institutions and cultural contexts have over ethical standards, social norms and individual behavior (Hofstede, 2001; Becker et al., 2020; Cappelen et al., 2022).

2 Experimental Protocol and Design

Our experiment was run in collaboration with the German SOEP-IS, which each year allows researchers to integrate new modules (surveys and experiments) to its standard socio-economics survey (Richter and Schupp, 2015). Our experiment was integrated into the 2020 wave and conducted as Computer-Assisted Personalized Interviews (CAPI) between September and December 2020.[5]

The experiment consists of two reporting tasks widely used in the lying experimental literature. One task is a variant of the classic die-rolling paradigm introduced by Fischbacher and Föllmi-Heusi (2013), in which participants are asked to roll a fair 10-sided die in private under a cup.[6] Participants could roll the die as many times as they wished, but they were asked to report the outcome of the first roll. The possible outcomes were integers between 0 and 9. Participants knew that there was a one-to-one correspondence between the reported outcome of the roll and the monetary payment from the task (e.g., reporting a '4' corresponded to a 4 Euro payment). To ensure maximal privacy, the interviewer left the room before the participant rolled the die. Participants reported the outcome of the first roll directly on the survey tablet while the interviewer was in another room. We refer to this as the 'unobservable' task since no one, except the participants themselves, could observe the true outcome of the die roll. This setup precludes analysis of lying at the individual level as the researcher does not know whether any individual report is truthful, but allows for inferences about lying at the aggregate level.

The other task is a variant of the paradigm developed by Gneezy et al. (2018). Participants saw ten black boxes on their survey tablet and were told that each box was

[5] Despite the COVID-19 pandemic in 2020, SOEP-IS continued to do some interviews in person, adhering to strict hygiene protocols. For details, see https://www.diw.de/documents/publikationen/73/diw_01.c.818889.de/diw_ssp0986.pdf.

[6] In Fischbacher and Föllmi-Heusi (2013), subjects roll a 6-sided die. We opted for a 10-sided die to provide subjects with a wider rage of options and have less granular data.

associated with an integer between 0 and 9, both included. They were informed that clicking on a box would reveal its associated number and that the numbers had been randomly assigned to the boxes. Participants were also told they could click on as many boxes as they wished, but they were asked to report the number associated with the first box they clicked on. As in the unobservable task, there was a one-to-one correspondence between the reported outcome and the monetary payment. Also, as in the unobservable task, the interviewer was in another room while participants reported their values directly on the tablet without showing them to the interviewer. However, because both the draw of the state (the first box a participant clicked on) and the report were computerized, the researcher can now detect lying at the individual level. Participants were explicitly told in the instructions that this was the case. Hence, we call this task the 'observable' task.[7]

Each participant took part in both tasks in a randomized order. Participants were made aware at the outset of the experiment that they would be participating in two tasks, but they were given detailed instructions about the second task only after having completed the first. Participants knew they would receive a monetary payment for only one of the two tasks, which was randomly determined at the end of the experiment by having the participant flip a coin. The interviewer then observed the number reported by the participant (without learning the actual number observed) and paid out the corresponding amount in cash on the spot. The full set of instructions received by participants is reproduced in Appendix A.

3 Theoretical Background and Empirical Strategy

3.1 Theoretical Background

To illustrate the logic of our empirical strategy, consider an agent who privately observes a state of the world $t \in T$, where T is a subset of equally spaced natural numbers from 0 to 9, as in our experiment. The state is drawn i.i.d. across agents from a uniform

[7]Note that in the observable task, participants are still paid what they report, regardless of whether the report is truthful or not. Moreover, while participants are told that the numbers they clicked are recorded, at the beginning of the experiment they are reminded that the data is completely anonymous. The sentence reads: "All information from the game, like all other information from this interview, will be evaluated anonymously only and will not be associated with your name.". Thus, despite the fact that lying is detectable in the observable task, it should be clear to participants that the expected material costs and benefits of lying are the same as in the unobservable tasks.

distribution. After observing the state, the agent reports $r \in R$ to an audience (in our case, the researcher), where each element in R naturally maps to a corresponding element in T. The agent receives a monetary payoff equal to the report r.

We model the agent's utility following the 'Reputation for Honesty + LC' model in Abeler et al. (2019), which, based on their analysis, is the model that more accurately describes the existing regularities in lying behavior documented in the literature. Utility depends on the monetary payoff as well as on two types of lying costs — intrinsic lying costs and image costs. Individuals are heterogeneous in their concerns for lying costs, with θ^{INT} and θ^{IMG} denoting the weight agents place on intrinsic and image costs, respectively. These concerns are jointly distributed across individuals according to a distribution Θ. The agent's utility is given by:

$$U(r, c(r,t), Pr(r \neq t|r); \theta^{INT}, \theta^{IMG}) = r - \theta^{INT} c(r,t) - \theta^{IMG}(Pr(r \neq t|r)), \quad (1)$$

where the first term is the agent's monetary payoff r and the second term captures the intrinsic lying cost through the function $c(r,t)$. This function describes the psychological cost associated with differences between the true state and the report by taking value 0 when the agent reports the true state $r = t$ and $c(r,t) > 0$ otherwise. The weight θ^{INT} captures the extent to which the agent is concerned with incurring the cost $c(r,t)$. The third term describes the agent's social image cost, as the individual weight θ^{IMG} is multiplied by $Pr(r \neq t|r)$, which is the probability that a given report r is untruthful. Utility increases in the first term and (weakly) decreases in the second and third terms.

Consider the two tasks performed by participants in our experiment. The first and second terms of (1) do not vary across the observable and unobservable tasks since the mapping between reports and states is the same in both tasks. However, the third term differs. In the unobservable task, the probability $Pr(r \neq t|r)$ is based on the Bayesian inferences made by the audience about whether an agent who makes a report r is lying, which corresponds to the fraction of liars at r in equilibrium. In contrast, in the observable task, the audience can perfectly observe both t and r for each participant in the experiment. Hence, for any report, $Pr(r \neq t|r)$ takes either value 0 if $r = t$ or value 1 otherwise. Thus, for any given value of θ^{IMG}, the observable task weakly increases the image cost of making an untruthful report compared to the unobservable task.

3.2 Empirical Strategy

Our empirical strategy relies on the exogenous increase in image costs induced by the observable task to gauge the weight θ^{IMG} that different types of individuals place on image considerations. Consider two groups of individuals differing by some characteristic, e.g., their gender. Suppose these two groups face the same $Pr(r \neq t|r)$ increase between the unobservable and observable tasks.[8] If the concern for image θ^{IMG} are similarly distributed across the two groups, we expect observability to affect their behavior in a similar way. However, suppose we observe that men reduce their reports relatively more than women when comparing their behavior across tasks: this indicates that θ^{IMG} is not distributed equally across groups and is, on average, larger for men than for women.

We operationalize this intuition by conducting a (pre-registered) difference-in-differences regression analysis with the participants' reports in the two tasks as the dependent variable. The independent variables consist of a treatment dummy, which takes value one if the task is observable, and a series of pre-registered individual characteristics listed in Table 1 and described in detail in the following subsection. Our interest lies in the interaction between the treatment dummy and the individual characteristics. We interpret any statistically significant interaction as evidence of differences in the weight placed on image considerations by individuals varying along that specific characteristic. The sign of the interaction effect relative to the observability dummy informs us which group of individuals places a higher weight on image concerns.

Take again gender as an example. Our inferences would be based on the following regression model:

$$r = \alpha + \beta_1 Observable + \beta_2 Female + \beta_3 Observable * Female + X + \epsilon \tag{2}$$

where $Observable$ takes value one for the participant's report in the observable task and value zero for his/her report in the unobservable task, $Female$ is a dummy taking value

[8]This implies that in the unobservable task the audience's beliefs are not conditional on any group-level characteristic. This may or may not be a mild assumption depending on the extent to which the individual characteristics are immediately observable to the audience, which, in turn, hinges on the nature of the audience itself (e.g., do subjects care about the inferences made by the interviewer or the researcher who will have access to all data?). There is some evidence that audiences may condition their inferences of honesty on visible characteristics. For example, Lohse and Qari (2021) show that men are perceived to be more dishonest than women, and anticipate this being the case. This would point to a weaker effect of our observability manipulation among men than women.

one if the participant is a female, and X is a vector of controls that include the other individual characteristics and their interaction with the observability dummy.

Based on previous literature, we expect β_1 to be negative as participants make, on average, lower reports when the researcher can observe both the true state and the report. Additionally, based on previous findings (Abeler et al., 2019), we expect β_2 to be negative as women lie on average less than men in unobservable tasks. The coefficient of interest is β_3. If β_3 is significantly different from zero and negative, this indicates that observability has a stronger effect on women than men, i.e., women place a relatively higher weight on image concerns than men. If β_3 is positive and significant, we can infer the opposite, that is, that men are significantly more concerned with image than women. If β_3 is not significantly different from zero, this suggests that women and men place a similar weight on image concerns.

3.3 Individual-level Covariates

The empirical strategy sketched above will be performed for the individual-level characteristics reported in Table 1. These variables were measured in the 2020 or earlier waves of the German Socio-Economic Panel Innovation Sample (SOEP-IS).[9]

The socio-demographic variables were measured through standard survey questions asking participants to report their gender (male/female), age at the time of the 2020 survey, number of years spent in education or training, net monthly income in the participant's household, and whether they were resident in East or West Germany before German reunification in 1989. We also use a variable measuring a participant's employment status in 2020. This was constructed by combining two survey questions, one asking people about their current employment category (fully employed, unemployed, or in some other form of employment, e.g., apprenticeship, voluntary service, part-time employment, etc.), and the other identifying people who were retired at the time of the interview.

Additionally, we were interested in exploring the correlation between dishonesty and participant's political and religious identity. The former was measured by asking participants whether they leaned towards a particular political party in Germany. If they answered positively, they were further asked to indicate the party they leaned towards.

[9]Whenever possible, we used responses to questions asked in the 2020 wave. Otherwise, we used a respondent's most recent answer in a wave prior to 2020.

Table 1: Summary Statistics of the Pre-Registered covariates

Pre-Registered Variable	**Observations**	**Sample Average**	**Year Collected**
	Socio-Demographic		
Female (binary)	1,318	52.7%	2020*
Age (in years)	1,318	54.6 (18.5)	2020*
Education (in years)	1,279	12.5 (2.8)	2020
Employment status	1,318		2020
Retired (binary)		41.9%	
Fully Employed (binary)		29.1%	
Unemployed (binary)		12.3%	
Other type of employment (binary)		16.7%	
Household net income (in 1000's Euro / month)	1,262	3.2 (2.1)	2020
Living in East Germany 1989 (binary)	1,316	23.2%	2020*
Political Interest	1,271		2020
No Interest (binary)		50.6%	
Moderate (binary)		41.8%	
Extreme Left (binary)		3.7%	
Extreme Right (binary)		3.8%	
Religious (binary)	850	61.3%	2020*
	Personality Traits		
Willingness to take risks (scale 0-10)	1,318	5.0 (2.2)	2020
Correct CRT Answers (scale 1-3)	933	0.9 (1.0)	2020
Trust in People (scale 1-4)	1,219	2.6 (0.5)	2020
Patience (scale 0-10)	1,228	5.9 (2.5)	2018
Big Five Personality Traits	1,246		2019
Openness		4.8 (1.1)	
Conscientiousness		5.7 (0.9)	
Extraversion		5.0 (1.1)	
Agreeableness		5.5 (0.9)	
Neuroticism		3.8 (1.3)	

Note: The second column reports the number of observations with non-missing values for each covariate. The third column reports summary statistics for each covariate, with relative frequencies for binary variables and mean and standard deviation (in parentheses) for non-binary variables. For the personality traits variables, higher values in the response scale indicate a stronger degree of the trait (more willingness to take risks, higher cognitive ability, more trust, more patience, more openness/conscientiousness/extraversion/agreeableness/neuroticism). Variables with an asterisk in the year of collection are collected when participants join the SOEP panel and some (e.g., *Female*) can be updated in subsequent waves of the panel: for all these variables, we use the entries recorded in the 2020 wave.

We construct our 'political interest' variable by combining the answers to those two questions ('No interest' in politics if they indicated not to lean towards any particular party and "Moderate"/"Extreme Left"/"Extreme Right" if they indicated inclination towards a party that is conventionally viewed as moderate/far left/far right).[10] Religiousness was measured by asking people whether they belonged to a church or religious community,

[10]We coded *SPD*, *CDU*, *CSU*, *FDP*, and *Grünen* as moderate. *Linke* as far left, and *AfD* and *NPD* as far right.

coding their response on a binary variable indicator.[11]

The personality traits variables were constructed using responses to the following survey questions. Willingness to take risks was measured using the standard SOEP question '*Are you generally a person who is willing to take risks or do you try to avoid taking risks?*', coded on an 11-point scale from '*very risk averse*' to '*very willing to take risks*'. Cognitive ability was measured by counting the number of correct answers respondents gave in the three-item Cognitive Reflection Test (CRT) by Frederick (2005). Trust was measured by averaging a respondent's degree of agreement (on a 4-point scale, from 1 = *'agree completely'* to 4 = *'totally disagree'*) with three questions about whether people can be trusted and relied upon (*In general, people can be trusted*; *These days, you cannot rely on anyone anymore*; and *When dealing with strangers, it is better to be careful before trusting them*). Patience was measured using the question '*Would you describe yourself as an impatient or a patient person in general?*', coded on an 11-point scale from 0 '*very impatient*' to 10 '*very patient*' (Cobb-Clark et al., 2022, 2024). Finally, the Big 5 Personality traits were derived from the short, 15-item version of the Big Five Inventory (Hahn et al., 2012).

3.4 Pre-registration and Sample Selection

In total, 1,603 respondents were offered to participate in the experiment. Of these, 1,330 accepted to participate, and 273 declined. Of the 1,330 who accepted to participate, 12 did not report values in at least one of the two tasks and are excluded from the dataset, leaving us with 1,318 observations.[12]

Our study and analysis plan were pre-registered at the Open Science Framework before we had access to the data: https://osf.io/wg4d7. The pre-registration specified the individual-level characteristics that are the focus of the analysis, listed in Table 1. It also specified our estimates would be based on OLS regressions with a Type-I error rate $\alpha = 0.05$ as the threshold for significance. Importantly, the pre-registration discussed a

[11]We do not distinguish between subjects who belong to different religions, mainly because we have only a few people identifying as religious (329 out of the 850 who were asked the question.)

[12]In Table B.1 of Appendix B, we study whether those respondents who accepted to participate in the experiment are different from those who opted out. Overall, there are only a few traits that explain the decision to participate robustly across regression specifications. Participation is negatively correlated with age and positively correlated with Extreme Right political support and willingness to take risks (although the latter effect is significant only in two out of three specifications).

concern about potential missing observations. Because the SOEP-IS is a panel, it has attrition, and participants are regularly refreshed. Moreover, in any given year, not all panel participants participate in all modules, and some might opt out if offered. For example, as mentioned above, our module was only offered to a subset of the panel (1,603 participants), of which 273 declined to participate. This means that we had anticipated that a significant part of the sample likely had missing values for the pre-registered variables (see Appendix C for a copy of the pre-registration).

We accounted for this problem in the pre-registration by indicating that we would omit from the analysis variables that contained excessive missing observations. However, the problem was more severe than we had anticipated. While most variables have missing data for no more than 5% of the subjects, only five of our pre-registered variables do not have any missing values. This results in a compounding problem once we require participants to have data for multiple variables *simultaneously* (e.g., when running regressions): if we restrict the analysis to subjects without any missing values, we lose about 73% of observations and are left only with 350 subjects. Moreover, upon inspecting the data, we found some evidence that the missing data was not randomly distributed and potentially correlated with our treatment manipulation. We report details in Appendix D where we show indications that those subjects who answer all questions, and therefore *do not* have missing observations, respond more to the observability treatment.

Taken together, these considerations convinced us to pursue a more comprehensive strategy for handling missing data than what we had initially pre-registered. Specifically, our analysis below will be based on three different models, reflecting different approaches to handling missing data with associated advantages and disadvantages. The first and third approaches described below are not pre-registered, while the second approach follows the pre-registration.

Our first approach will restrict the analysis to only the subset of individuals who do not have missing information in any of the variables. This allows us to use the full set of covariates we had pre-registered but at the cost of a significant loss of statistical power. We call this approach the "Unrestricted Covariates Set" approach.

Our second approach follows the pre-registration and omits those variables which have excessive missing observations. The cutoff is set to those variables with more than 5% of missing data, which results in the omission of religiousness and all personality traits variables except risk attitudes.[13] With this approach, we reduce the loss of observations due to missing data and drop approximately 10% of the full sample (135 subjects) for the analysis. We call this approach the "Restricted Covariate Set" approach.

As a third alternative, we use the "Missing Indicator" approach to handle missing data in our analysis (Rubin, 1976; Groenwold et al., 2012). This method allows us to retain all observations by accounting for missingness directly in our regression model. Specifically, for each observation i and covariate j, we define a dummy variable $M_{i,j}$ that equals one if $X_{i,j}$ is missing and zero if it is observed, and then replace missing values in $X_{i,j}$ with a placeholder (in our case, zero), resulting in $\tilde{X}_{i,j}$. By including *both* $\tilde{X}_{i,j}$ and $M_{i,j}$ as separate predictors in our regression model, this approach allows us to estimate the effect of $X_{i,j}$ among those with observed data while controlling for systematic differences associated with missing data.[14] While the missing indicator method was originally criticized (e.g., Greenland and Finkle, 1995; Jones, 1996), it is commonly used across social (A. Hoffman and Strezhnev, 2023) and medical sciences (Cho et al., 2021), as the recent literature has shown it is a valid approach under most data conditions (e.g., Blake et al., 2020; Song et al., 2021)

4 Results

4.1 Aggregate behavior

We start our analysis by providing an overview of aggregate behavior in the two tasks. In Figure 1 we plot the relative frequency of reports made in the unobservable (left panel) and observable tasks (right panel). The dashed horizontal line represents the expected relative frequency of each report under truth-telling. It is clear from the figure that there is underreporting of the low-paying numbers (0 to 4) and overreporting of the high-paying numbers (5 to 9).[15] A series of binomial tests that compare the theoretical frequency of

[13]The dropped variables are trust, patience, Big Five traits, and number of correct CRT answers.

[14]For example, if income data is missing more often for individuals with lower education, simply omitting these observations or ignoring the missingness could bias our results. By including a missingness indicator for income, we control for the possibility that missing income data is related to other variables like education or employment status, allowing us to disentangle the influence of missing income from its relationship with other variables.

[15]In the observable task we can observe, for each subject, whether they reported a number different from the one they were asked to report. A total of 131 subjects (10.21%) did so. A further 35 subjects made a report without clicking on any box (which is another form of misreporting).

each report under truth-telling to their observed frequency, reveal statistically significant differences between the theoretical and empirical distributions, both in the unobservable and observable task. The significant cases are indicated by stars plotted at the top of the bar graph in Figure 1.[16]

Importantly, the differences between empirical and theoretical distributions appear to differ across the two tasks, with smaller deviations from truth-telling in the observable task. The average report in the unobservable task is 5.26 (st. dev. 2.75), while it is 4.94 (st. dev. 2.86) in the observable task. A matched-pairs Wilcoxon signed-rank test confirms the distribution of reports differ between unobservable and observable tasks ($p < 0.01$).[17] An OLS regression (Table B.3 in Appendix B) controlling for the order in which the tasks were presented to subjects as well as their individual characteristics, confirms the result: reports are on average lower in the observable than unobservable task.

Overall, these results show that the behavior in our experiment aligns with the previous literature. Specifically, we find that: 1) there are deviations from truth-telling in both the unobservable and observable tasks, 2) there is overreporting of both maximal and non-maximal values, and 3) observability reduces misreports.

[16]See Table B.2 in Appendix B for p-values from these tests. The table also reports multiple-hypotheses-corrected p-values based on the Bonferroni procedure.

[17]In the pre-registration we mistakenly registered a Kolmogorov-Smirnov test to detect differences in reports across treatments. This would be incorrect since the reported values across tasks are not independent.

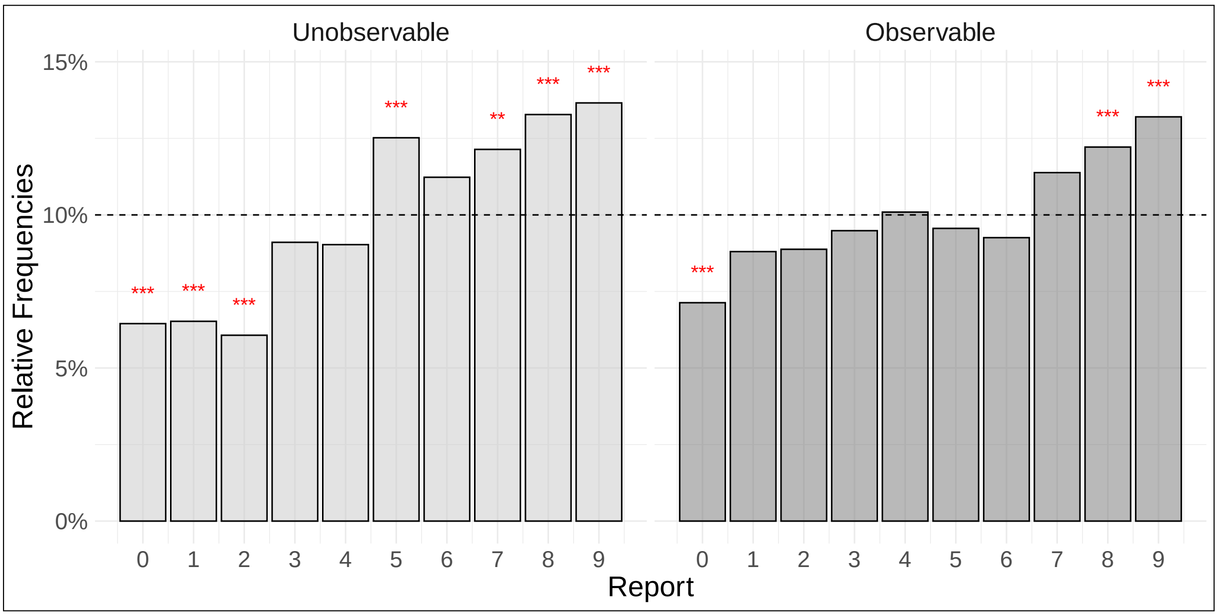

Figure 1: Empirical distribution of reports in the unobservable task (left panel) and observable task (right panel). The dotted horizontal bar marks the expected relative frequency of each report under the assumption of truthful reporting. The stars indicate significance levels based on binomial tests comparing the observed and expected frequency of each report (*** = 1% level; ** = 5% level). Table B.2 in Appendix B for p-values from these tests and multiple-hypotheses-corrected p-values based on the Bonferroni procedure. Additionally, Figure E.1 in Appendix E shows the disaggregated data based on the order of the tasks.

4.2 Correlates of the Social Image Cost of Lying

In this section, we address the paper's main question: assessing the incidence of image lying cost across different substrata of the population. To do so, we exploit the exogenous manipulation of observability in our within-subject design and run a series of regressions of subject i's reported value in task t ($report_{i,t}$) on the interaction between the individual characteristics and a treatment dummy *observable*, which takes value one in the observable task and zero otherwise. As discussed in Section 3, our focus lies on the coefficient of these interactions, as they reveal whether individual characteristics explain the differences in reports between the unobservable and observable tasks.

Table 2 reports the results. As discussed in Section 3.2, we run three different models encompassing the different approaches to handle missing values. In column 1, we use the "Unrestricted Covariates Set" method: we include in the regression all covariates, which restricts the sample to the 350 subjects (26.5% of the full sample) who do not have any missing data. In column 2, we use the "Restricted Covariates Set" approach (our pre-registered model). We restrict our analysis to those covariates for which the missing data represents less than 5% of the subjects. This results in a larger sample (1,183 subjects, which is 89.7% of the full sample) at the cost of dropping religiousness and all personality trait variables except risk preferences. Finally, in column 3 we report results based on the “Missing Indicator” method. This allows us to include all covariates and all 1,318 subjects, although the method has been criticized as discussed earlier. In all cases, we also include fixed effects for subject’s state of residence and a dummy (*Unobservable first*) which controls for order effect and takes value one if the unobservable task came first and zero otherwise.

Focusing on the coefficients of the interaction terms, for most covariates we observe little variation in the effects of observability. Across all models, we do not see heterogeneous effects across age, employment status, level of education, political attitudes, or any personality trait variables. The only covariates that display statistically significant interactions with observability are gender, being a citizen of East Germany (GDR) before reunification, and religiousness. However, the latter effect is only detected in column 1, which relies on a considerably restricted sample, while it disappears in column 3, where we use the full sample. The effects of gender and GDR appear instead more robust.

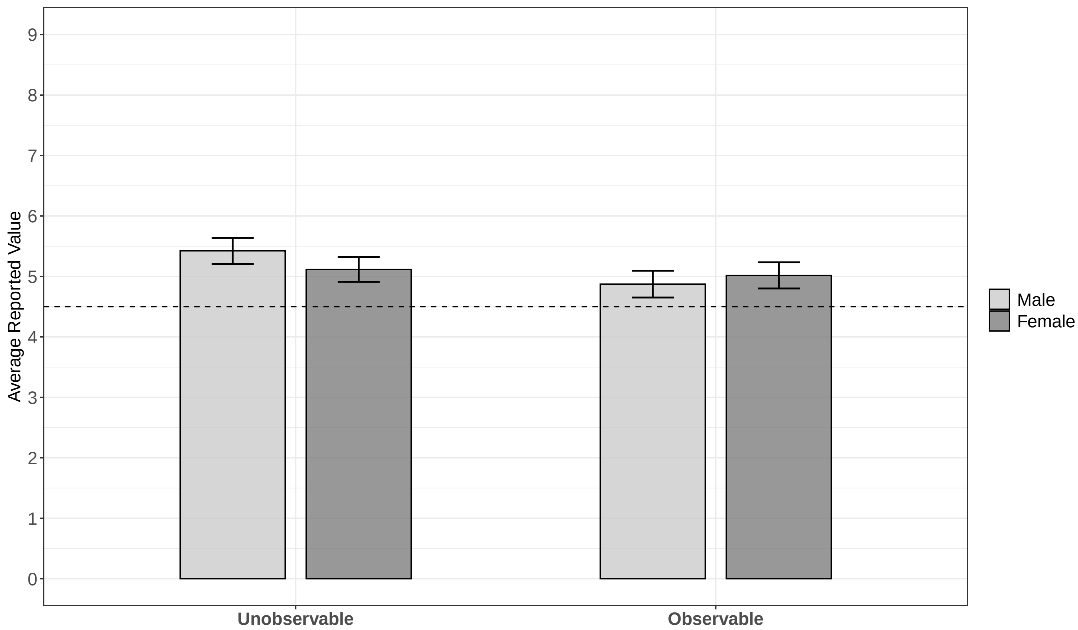

Figure 2: Average reported number in the unobservable and observable tasks by men and women. Bars represent 95% confidence intervals, the dashed line represents the expected average report. Based on the full sample of 1,318 subjects.

Figure 2 illustrates the heterogeneous effect for gender. The figure plots the average reports of men and women in the unobservable and observable tasks (based on the full

Table 2: OLS regressions

	Unrestricted $report_{i,t}$	Restricted $report_{i,t}$	Missing Indicator $report_{i,t}$
Unobservable First	0.004	0.092	0.110
	(0.235)	(0.120)	(0.113)
Observable	-1.143	-0.968	-1.062
	(2.869)	(0.822)	(0.785)
Age	-0.008	-0.002	-0.003
	(0.015)	(0.007)	(0.007)
Observable × Age	0.011	0.007	0.006
	(0.020)	(0.010)	(0.010)
Female	-0.968***	-0.469***	-0.444***
	(0.352)	(0.173)	(0.168)
Observable × Female	1.157**	0.535**	0.435
	(0.492)	(0.238)	(0.231)
GDR	-0.640	0.712***	0.558**
	(0.524)	(0.273)	(0.256)
Observable × GDR	-1.305**	-0.530**	-0.508
	(0.518)	(0.263)	(0.262)
Unemployed	-0.109	-0.080	-0.047
	(0.507)	(0.287)	(0.269)
Other Employment	0.973	0.266	0.241
	(0.520)	(0.249)	(0.238)
Retired	-0.052	-0.353	-0.293
	(0.582)	(0.285)	(0.273)
Observable × Unemployed	-1.029	-0.346	-0.235
	(0.706)	(0.416)	(0.395)
Observable × Other Employment	-0.997	-0.243	-0.216
	(0.676)	(0.362)	(0.350)
Observable × Retired	0.105	0.389	0.469
	(0.784)	(0.386)	(0.375)
Education	-0.082	-0.055	-0.064**
	(0.073)	(0.033)	(0.031)
Observable × Education	0.010	0.005	-0.020
	(0.090)	(0.044)	(0.036)
Net income	-0.090	-0.051	-0.035
	(0.100)	(0.046)	(0.042)
Observable × Net income	0.035	-0.007	0.003
	(0.155)	(0.062)	(0.055)
Moderate	-0.740**	0.099	0.030
	(0.362)	(0.175)	(0.168)
Extreme Left	0.021	-0.224	-0.298
	(0.836)	(0.485)	(0.491)
Extreme Right	-0.913	-0.461	-0.270
	(0.690)	(0.459)	(0.438)
Observable × Moderate	0.507	-0.122	0.016
	(0.500)	(0.245)	(0.231)
Observable × Extreme Left	-0.672	0.550	0.842
	(1.231)	(0.677)	(0.678)
Observable × Extreme Right	1.069	0.860	0.698
	(1.133)	(0.676)	(0.649)
Risk willing	-0.038	0.000	-0.003
	(0.072)	(0.036)	(0.035)
Observable × Risk willing	0.081	0.012	-0.002
	(0.101)	(0.051)	(0.050)
Religious	-0.165		0.139
	(0.359)		(0.205)
Observable × Religious	1.053**		-0.160
	(0.468)		(0.260)
CRT	0.256		0.111
	(0.202)		(0.092)
Observable × CRT	-0.249		0.041
	(0.263)		(0.119)
Trust	0.167		-0.013
	(0.246)		(0.117)
Observable × Trust	-0.319		-0.066
	(0.315)		(0.134)
Patience	-0.050		0.000
	(0.066)		(0.031)
Observable × Patience	0.058		-0.012
	(0.083)		(0.040)
Extraversion	-0.034		0.029
	(0.167)		(0.074)
Observable × Extraversion	0.012		0.073
	(0.227)		(0.101)
Conscientiousness	0.193		0.202**
	(0.198)		(0.087)
Observable × Conscientiousness	0.013		-0.032
	(0.262)		(0.113)
Openness	0.044		-0.053
	(0.142)		(0.071)
Observable × Openness	-0.133		0.052
	(0.209)		(0.096)
Neuroticism	0.007		0.025
	(0.145)		(0.064)
Observable × Neuroticism	0.231		0.119
	(0.208)		(0.086)
Agreeableness	0.366**		0.068
	(0.171)		(0.083)
Observable × Agreeableness	-0.195		-0.043
	(0.227)		(0.108)
Constant	4.084	6.894***	5.192***
	(2.292)	(0.677)	(0.979)
Observations	700	2366	2636
R^2	0.124	0.032	0.047
Land Fixed Effects	Yes	Yes	Yes

** $p < 0.05$, *** $p < 0.01$

Note: OLS regressions where the dependent variable is $report_{i,t}$, the value reported by individual i in task t. Robust standard errors clustered at the individual level. The three models differ in their approach to handling missing data. Column 1 uses the “Unrestricted Covariates Set” approach, column 2 the “Restricted Covariates Set” approach, and column 3 the “Missing Indicator” method. See section 3.4 for details. In column 3, the regression model also includes dummy variables capturing, for each covariate, whether an observation is missing. These variables are excluded from the table output to ease readability. See Table B.4 in Appendix B for the full output.

sample of 1,318 observations). In the unobservable task, the average report is 5.42 (st. dev. 2.74) for men and 5.12 (st. dev. 2.75) for women. This difference is statistically significant in all three models of Table 2, which is in line with previous evidence on gender differences in truth-telling in unobservable tasks. In the observable task, the difference in reports becomes smaller and reverts in sign: the average report is 4.87 (st. dev. 2.82) for men and 5.02 (st. dev. 2.91) for women. Linear restriction tests confirm this difference is insignificantly different from zero in all three models of Table 2 (all $p > 0.584$). Moreover, the coefficient of the interaction term is positive and significant in Columns 1 and 2 ($p - vals = 0.019$ and 0.025, respectively), but not in Column 3 ($p - val = 0.060$).

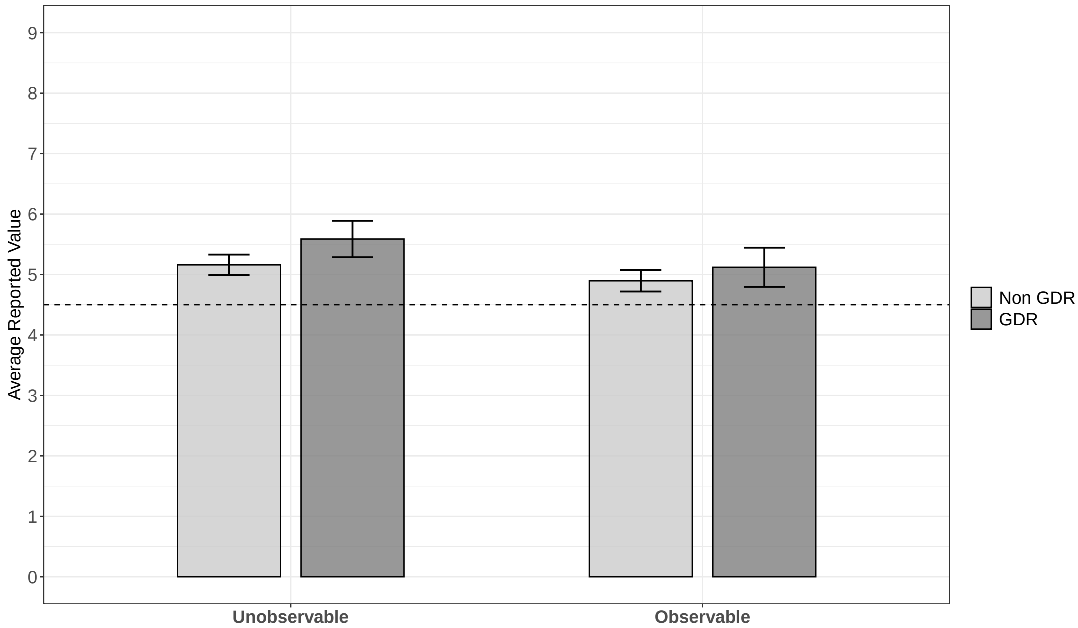

Figure 3: Average reported number in the unobservable and observable tasks by subjects who were and were not citizens of GDR in 1989. Bars represent 95% confidence intervals, the dashed line represents the expected average report. Based on the full sample of 1,318 subjects.

Figure 3 plots the average report in the unobservable and observable tasks disaggregated depending on whether the subject was a citizen of East Germany before reunification. The average report is larger for former citizens of East Germany both in the unobservable task (GDR: 5.68, st. dev. 2.70 vs. Not GDR: 5.13, st. dev. 2.76, difference statistically significant in columns 2 and 3) and in the observable task (GDR: 5.10, st. dev. 2.92 vs. Not GDR: 4.91, st. dev. 2.86, difference statistically significant in column 1 only). Comparing the difference-in-differences across tasks, the data indicate that former citizens of East Germany respond more strongly to observability. The coefficient of the interaction term between the GDR dummy and the task dummy is significantly different from zero in the first two models ($p = 0.012$ in column 1 and $p = 0.044$ in column 2), while the level of statistical significance just exceeds the 5% threshold in column 3 ($p = 0.053$).

Overall, the evidence across the three regression specifications suggests that the effect of observability varies across gender as well as between individuals who were citizens in East vs West Germany before reunification and were thus exposed to radically different economic and political systems. It is important to emphasize that these effects are observed in regression specifications that control for the influence that additional covariates may have on reporting behavior, including the interactions between these covariates and our observability manipulation. This means, for example, that the differential effect of former East German citizens is observed while holding constant the influence of factors (such as income) that may differ across former West and East German citizens – both in their direct influence and through their interaction with observability. Thus, our results suggest that social image costs may loom larger among men and former citizens of East Germany for reasons not fully captured by the control variables in our model but that instead arise from factors specific to these groups themselves.

5 Discussion and Conclusions

One of the most robust findings in the behavioral economics literature over the past decades is that people are averse to lying. This aversion has been found across cultures (Gächter and Schulz, 2016; Aycinena et al., 2022) and under different experimental conditions (Jacobsen et al., 2018; Abeler et al., 2019; Gerlach et al., 2019). Importantly, there is consensus in the literature about the main drivers for lying aversion — 'intrinsic lying costs' and 'image lying costs' (Dufwenberg and Dufwenberg, 2018; Gneezy et al., 2018; Abeler et al., 2019; Khalmetski and Sliwka, 2019). However, despite such wide consensus, little is known about the relative importance of these costs across different segments of the population. Are the intrinsic costs of lying stronger for women than for men? Do younger people place more weight on image lying costs? And how are these costs shaped by the socio-economic circumstances individuals face throughout their lives? In this paper, we begin to address these questions using a pre-registered survey experiment conducted in the 2020 wave of the German Socio-Economic Panel Innovation Sample (SOEP-IS) to reach a large sample of the German population.

Our experiment exogenously manipulates the importance of image costs by varying the observability of participants' behavior in two standard information-reporting tasks (Fischbacher and Föllmi-Heusi, 2013; Gneezy et al., 2018). We then study the extent to which this increase in image costs leads to behavioral differences in dishonesty among individuals who belong to different strata of the population. We find that, although image costs do not seem to vary systematically across most of the individual-level characteristics we consider, they differ across gender and former citizens of West and East Germany. Specifically, men and former citizens of East Germany appear more influenced by image concerns than women and former West Germany citizens.

Remarkably, we observe little variation across many of the socio-demographic and personality factors we consider. Some of these factors – such as trust, conscientiousness, or agreeableness – have previously been linked to prosocial and moral behavior. To the extent that our results from the lying paradigm generalize to other decision-making environments, our study suggests these links may not operate through a social image channel but rather through an intrinsic preference for prosociality. More generally, the lack of correlation between socio-demographic outcomes and lying behavior confirms the results of Abeler et al. (2014), who did not detect any effects of socio-economic measures on the lying behavior of a representative German sample, using a much larger sample.

In concurrent work, Schudy et al. (2024) included a two-question survey instrument in the 2023 SOEP-IS wave to separately capture the influence of intrinsic and image costs on dishonesty. In line with our results, they find limited evidence of an association between lying costs and political affiliation, age, or selection into particular employment industries. They report a positive relation between religiosity (measured through exposure to religious parents) and image costs, albeit based on a small sample, which is also the limitation we faced in our analysis.

Schudy et al. (2024) also find that lying costs are associated with both gender and exposure to East Germany. In particular, they report an association between image costs and being schooled in East Germany, confirming our finding that exposure to a socio-political environment characterized by extensive state surveillance shapes image concerns. However, their gender results diverge from ours, as Schudy et al. (2024) observe gender differences in lying behavior mediated through intrinsic costs, while we detect a differential effect of image costs. Related work by Basic (2018) and Lohse and Qari (2021) further reflect the complexity of this picture. In the context of a dictator game, Basic (2018) report that men respond more strongly than women to self-image manipulations but not to social image manipulations. Lohse and Qari (2021) find that women, but not men, appear to be more honest in a face-to-face reporting task relative to a task where reporting does not involve face-to-face interaction. However, men respond more strongly when the audience's inferred probability of dishonesty is linked to the possibility of an audit, suggesting a stronger strategic response for men than women. Taken together, these results suggest a complex relationship between gender, lying costs, and dishonest behavior that warrants further investigation.

Overall, the results from our paper and other related work represent initial steps toward understanding the deeper mechanisms that shape dishonesty and moral behavior more broadly. The emerging picture highlights an intriguing picture the impact of social environments and historical contexts have in shaping individual preferences (see also Gerber and Jackson, 1993; Hofstede, 2001; Becker et al., 2020; Cappelen et al., 2022, for the importance of culture and institutions for individual preferences). However, more work is needed to probe the robustness of our observed links and to develop a more comprehensive understanding of the underlying mechanisms that determine *how* and *when* historical and cultural contexts can influence moral behavior and dishonesty.

References

A. Hoffman, D. and A. Strezhnev (2023): "Longer trips to court cause evictions," *Proceedings of the National Academy of Sciences*, 120, e2210467120. Cited on page 14.

Abeler, J., A. Becker, and A. Falk (2014): "Representative evidence on lying costs," *Journal of Public Economics*, 113, 96–104. Cited on pages 3, 4, 5, and 21.

Abeler, J., A. Falk, and F. Kosse (2024): "Malleability of preferences for honesty," *The Economic Journal*, forthcoming. Cited on pages 2 and 5.

Abeler, J., D. Nosenzo, and C. Raymond (2019): "Preferences for truth-telling," *Econometrica*, 87, 1115–1153. Cited on pages 2, 3, 4, 5, 8, 10, and 20.

Alesina, A. and N. Fuchs-Schündeln (2007): "Good-bye Lenin (or not?): The effect of communism on people's preferences," *American Economic Review*, 97, 1507–1528. Cited on page 5.

Alesina, A. and P. Giuliano (2015): "Culture and institutions," *Journal of Economic Literature*, 53, 898–944. Cited on page 5.

Algan, Y. and P. Cahuc (2010): "Inherited trust and growth," *American Economic Review*, 100, 2060–2092. Cited on page 5.

Aycinena, D., L. Rentschler, B. Beranek, and J. F. Schulz (2022): "Social norms and dishonesty across societies," *Proceedings of the National Academy of Sciences*, 119, e2120138119. Cited on page 20.

Basic, Z. (2018): "Essays on Image Concerns and Norm-Enforcing Behavior." *PhD diss., Dissertation, Bonn, Rheinische Friedrich-Wilhelms-Universit√§tBonn.Citedonpages* 21*and* 22.

Basic, Z. and S. Quercia *(2022): "The influence of self and social image concerns on lying,"* Games and Economic Behavior*, 133, 162–169. Cited on page 3.*

Becker, A., B. Enke, and A. Falk *(2020): "Ancient origins of the global variation in economic preferences," in* AEA Papers and Proceedings*, American Economic Asso-*

ciation 2014 Broadway, Suite 305, Nashville, TN 37203, vol. 110, 319–323. Cited on pages 6 and 22.

Blake, H. A., C. Leyrat, K. E. Mansfield, L. A. Tomlinson, J. Carpenter, and E. J. Williamson *(2020): "Estimating treatment effects with partially observed covariates using outcome regression with missing indicators,"* Biometrical Journal*, 62, 428–443. Cited on page 14.*

Booth, A. L. and P. Nolen *(2012): "Gender Differences in Risk Behaviour: Does Nurture Matter?"* The Economic Journal*, 122, F56–F78. Cited on page 5.*

Cappelen, A. W., B. Enke, and B. Tungodden *(2022): "Moral universalism: Global evidence," Tech. rep., National Bureau of Economic Research. Cited on pages 5, 6, and 22.*

Capraro, V. *(2018): "Gender differences in lying in sender-receiver games: A meta-analysis,"* Judgment and Decision Making*, 13, 345–355. Cited on page 4.*

Childs, J. *(2012): "Gender differences in lying,"* Economics Letters*, 114, 147–149. Cited on page 4.*

Cho, B., Y. Han, M. Lian, G. A. Colditz, J. D. Weber, C. Ma, and Y. Liu *(2021): "Evaluation of racial/ethnic differences in treatment and mortality among women with triple-negative breast cancer,"* JAMA Oncology*, 7, 1016–1023. Cited on page 14.*

Cobb-Clark, D. A., S. C. Dahmann, D. A. Kamhöfer, and H. Schildberg-Hörisch *(2022): "The predictive power of self-control for life outcomes,"* Journal of Economic Behavior & Organization*, 197, 725–744. Cited on page 12.*

——— *(2024): "Surveillance and Self-Control,"* The Economic Journal*, 134, 1666–1682. Cited on page 12.*

Crede, A.-K. and F. von Bieberstein *(2020): "Reputation and lying aversion in the die roll paradigm: Reducing ambiguity fosters honest behavior,"* Managerial and Decision Economics*, 41, 651–657. Cited on page 3.*

Dennis, M. and N. Laporte *(2014):* The Stasi: Myth and Reality*, Routledge. Cited on page 5.*

Dreber, A. and M. Johannesson *(2008): "Gender differences in deception,"* Economics Letters*, 99, 197–199. Cited on page 4.*

Dufwenberg, M. and M. A. Dufwenberg *(2018): "Lies in disguise–A theoretical analysis of cheating,"* Journal of Economic Theory*, 175, 248–264. Cited on pages 2, 5, and 20.*

Falk, A., A. Becker, T. Dohmen, B. Enke, D. Huffman, and U. Sunde *(2018): "Global Evidence on Economic Preferences*,"* The Quarterly Journal of Economics*, 133, 1645–1692. Cited on page 5.*

Fischbacher, U. and F. Föllmi-Heusi *(2013): "Lies in disguise —an experimental study on cheating,"* Journal of the European Economic Association*, 11, 525–547. Cited on pages 2, 3, 6, and 21.*

Frederick, S. *(2005): "Cognitive Reflection and Decision Making,"* Journal of Economic Perspectives*, 19, 25–42. Cited on page 12.*

Fries, T., U. Gneezy, A. Kajackaite, and D. Parra *(2021): "Observability and lying,"* Journal of Economic Behavior & Organization*, 189, 132–149. Cited on page 3.*

Fulbrook, M. *(2014):* A history of Germany 1918-2014: The divided nation*, John Wiley & Sons. Cited on page 5.*

Gächter, S. and J. F. Schulz *(2016): "Intrinsic honesty and the prevalence of rule violations across societies,"* Nature*, 531, 496–499. Cited on page 20.*

Gerber, E. R. and J. E. Jackson *(1993): "Endogenous preferences and the study of institutions,"* American Political Science Review*, 87, 639–656. Cited on page 22.*

Gerlach, P., K. Teodorescu, and R. Hertwig *(2019): "The truth about lies: A meta-analysis on dishonest behavior."* Psychological Bulletin*, 145, 1. Cited on pages 2 and 20.*

Gneezy, U. and A. Kajackaite *(2017): "Incentives and cheating,,"* Games and Economic Behavior*, 102, 433–44. Cited on page 3.*

Gneezy, U., A. Kajackaite, and J. Sobel *(2018): "Lying aversion and the size of the lie,"* American Economic Review*, 108, 419–53. Cited on pages 2, 3, 4, 5, 6, 20, and 21.*

Greenland, S. and W. D. Finkle *(1995): "A critical look at methods for handling missing covariates in epidemiologic regression analyses,"* American Journal of Epidemiology*, 142, 1255–1264. Cited on page 14.*

Groenwold, R. H., I. R. White, A. R. T. Donders, J. R. Carpenter, D. G. Altman, and K. G. Moons *(2012): "Missing covariate data in clinical research: when and when not to use the missing-indicator method for analysis,"* Cmaj*, 184, 1265–1269. Cited on page 14.*

Hahn, E., J. Gottschling, and F. M. Spinath *(2012): "Short measurements of personality - Validity and reliability of the GSOEP Big Five Inventory (BFI-S),"* Journal of Research in Personality*, 46, 355–359. Cited on page 12.*

Hofstede, G. *(2001):* Culture's consequences: Comparing values, behaviors, institutions and organizations across nations*, Sage publications. Cited on pages 6 and 22.*

Holmen, M., F. Holzmeister, M. Kirchler, M. Stefan, and E. Wengstrom *(2023): "Economic Preferences and Personality Traits Among Finance Professionals and the General Population,"* The Economic Journal*, 133, 2949–2977. Cited on page 5.*

Huber, C., C. Litsios, A. Nieper, and T. Promann *(2023): "On social norms and observability in (dis) honest behavior,"* Journal of Economic Behavior & Organization*, 212, 1086–1099. Cited on page 3.*

Jacobsen, C., T. R. Fosgaard, and D. Pascual-Ezama *(2018): "Why do we lie? A practical guide to the dishonesty literature,"* Journal of Economic Surveys*, 32, 357–387. Cited on pages 2 and 20.*

Jones, M. P. *(1996): "Indicator and stratification methods for missing explanatory variables in multiple linear regression,"* Journal of the American Statistical Association*, 91, 222–230. Cited on page 14.*

Khalmetski, K. and D. Sliwka (2019): *"Disguising lies ‚ÄîImage concerns and partial lying in cheating games,"* American Economic Journal: Microeconomics, *11, 79–110. Cited on pages 2, 5, and 20.*

Lichter, A., M. Löffler, and S. Siegloch (2021): *"The long-term costs of government surveillance: Insights from stasi spying in East Germany,"* Journal of the European Economic Association, *19, 741–789. Cited on page 5.*

Lilleholt, L., C. Schild, and I. Zettler (2020): *"Not all computerized cheating tasks are equal: a comparison of computerized and non-computerized versions of a cheating task,"* Journal of Economic Psychology, *78, 102270. Cited on page 3.*

Lohse, T. and S. Qari (2021): *"Gender differences in face-to-face deceptive behavior,"* Journal of Economic Behavior & Organization, *187, 1–15. Cited on pages 9, 21, and 22.*

Mol, J. M., E. C. van der Heijden, and J. J. Potters (2020): *"(Not) alone in the world: Cheating in the presence of a virtual observer,"* Experimental Economics, *23, 961–978. Cited on page 3.*

Ockenfels, A. and J. Weimann (1999): *"Types and patterns: an experimental East-West-German comparison of cooperation and solidarity,"* Journal of Public Economics, *71, 275–287. Cited on page 5.*

Parra, D. (2024): *"Eliciting dishonesty in online experiments: The observed vs. mind cheating game,"* Journal of Economic Psychology, *102715. Cited on page 3.*

Richter, D. and J. Schupp (2015): *"The soep innovation sample (soep is),"* Journal of Contextual Economics–Schmollers Jahrbuch, *135, 389–399. Cited on page 6.*

Rubin, D. B. (1976): *"Inference and missing data,"* Biometrika, *63, 581–592. Cited on page 14.*

Schudy, S., S. Grundmann, and L. Spantig (2024): *"Individual Preferences for Truth-Telling,"* CESifo Working Paper Series N 11521. *Cited on pages 5 and 21.*

Song, M., X. Zhou, M. Pazaris, and D. Spiegelman (2021): *"The missing covariate indicator method is nearly valid almost always,"* arXiv preprint arXiv:2111.00138. *Cited on page 14.*

A Instructions

Below we transcribe the instructions that participants saw on their screens. We also include in *italics* interviewers instructions. Importantly, participants did not see the instructions for the interviewer.

All of these instructions come from the module documentation which can be found online at the SOEP-IS webpage.

Intro

Now I would like to invite you to participate in a short and simple game. The task consists of two parts. What the two parts are about, you will learn immediately before. In both parts, you have a chance to win real money. But only one of the two parts will be selected for the final payout. You will flip a coin for it. If the coin shows "tails" after the toss, you will get the result for part 1. If the coin shows "heads", you will get the result for part 2. All information from the game, like all other information from this interview, will be evaluated anonymously only and will not be associated with your name. This is the end of the introduction.

Interviewer: Please do not inform the target person in advance about the content of the two games! As the interviewer, make sure that the target person does not see the dice cup and dice before the dice game begins. Please do not give the dice cup and dice to the target person until the target person asks you for it!

Q490

Would you like to play the game?

Interviewer: Before starting the module: As the interviewer, please try to make sure that no other interviewee is within earshot. As the interviewer, also try to leave the room or turn away so that the target person can answer the questions alone. It is important that the target person can complete the following tasks in complete privacy. If the target person agrees to play the game, please hand over the laptop to the target person for self-completion. After the experiment, the target person will be asked to hand over the laptop to you and continue the interview as regular.

- Yes

- No

Q491

Interviewer: Please hand over the laptop to the interviewee for self-completion.

[assume the randomization has participants start with the unobservable task]

Q493

The interviewer has just given you a 10-sided die and a dice cup. Now please roll the dice as many times as you like using the dice cup. Caution: Make sure that the interviewer does not see the result of your roll. Please remember the result of your first roll and enter it on the screen. How much you can earn in this part, if this part of the task is selected for payout, depends on the number you write down here. A higher number means a higher payout. You can see the exact payout for each reported number in the table below. The table will remain on the screen until you enter your number.

Note: Please have your interviewer hand you the dice cup and dice cup. Please make sure that no one can observe you playing the game.

[Participants saw a table reflecting the payoffs for all values they could report]

Q494

In this part, you will see 10 black boxes on your screen. There are numbers behind the black boxes. The numbers behind the black boxes are 0, 1, 2, 3, 4, 5, 6, 7, 8, and 9. The numbers are in random order, so the first box does not necessarily contain 1, the second does not necessarily contain 2, and so on. As soon as you click on a black box, the number hidden behind it is displayed. You can click on as many black boxes as you like. The laptop records how many and which boxes you click. Please remember the number behind the first black box you clicked and enter it on the screen. How much you earn in this part, if this part of the task is selected for payout, depends on the number you enter. A higher number means a higher payout. You can see the exact payout for each number reported in the table below. The table will remain on the screen until you enter your number.

[Participants saw a table reflecting the payoffs for all values they could report]

Q495

You have reached the end of the task. Please hand the laptop back to your interviewer(s) and continue the survey as usual.

Q496_0

Now we will determine whether the first or second part of this task will be paid out. To do this, I ask you to flip a coin in a clearly visible position so that I can note down the result. Ask the target person to toss a coin in plain sight. Note whether "tails" or "heads" is on top.

Interviewer: Ask the target person to toss a coin in plain sight. Note whether "tails" or "heads" is on top.

Q496

As you can see, [PROG: Show result from IZAHLAUSZ01] is on top. Therefore, you will now receive the payoff from [PROG: please show: Part 1 (if IZAHLAUSZ01=1); Part 2 (if IZAHLAUSZ01=2)]. You have indicated that in [PROG: please show: Part 1 (if IZAHLAUSZ01=1); Part 2 (if IZAHLAUSZ01=2)] the first number was [PROG: please enter the number the participant entered based on their split group (SPLITZAHL) in either Part 1 or Part 2]. Therefore, you will now receive euros from me [PROG: reported number = euro amount].

Interviewer: Please pay the target person the appropriate amount.

B Extra Tables

Table B.1: Probability of accepting to participate in the lying experiments

	Unrestricted Participate	**Restricted** Participate	**Missing Indicator** Participate
Age	-0.003*	-0.003***	-0.003***
	(0.002)	(0.001)	(0.001)
Female	-0.006	0.011	0.009
	(0.035)	(0.020)	(0.020)
Education	0.000	0.003	-0.000
	(0.007)	(0.004)	(0.004)
Net income	0.002	-0.009*	-0.007
	(0.011)	(0.005)	(0.005)
Unemployed	0.030	-0.007	0.009
	(0.043)	(0.032)	(0.032)
Other Employment	-0.053	-0.004	0.009
	(0.056)	(0.028)	(0.028)
Retired	-0.037	-0.024	-0.006
	(0.056)	(0.032)	(0.033)
GDR	0.070	-0.038	-0.031
	(0.058)	(0.034)	(0.034)
Risk willing	0.003	0.014***	0.017***
	(0.007)	(0.004)	(0.004)
Religious	-0.005		-0.037
	(0.036)		(0.024)
Moderate	0.072*		0.011
	(0.039)		(0.021)
Extreme Left	0.125*		0.032
	(0.067)		(0.050)
Extreme Right	0.115***		0.082**
	(0.037)		(0.039)
CRT	0.001		0.068**
	(0.021)		(0.028)
Trust	0.029		0.070**
	(0.032)		(0.032)
Patience	-0.011*		0.071*
	(0.006)		(0.037)
Openness	-0.006		-0.011
	(0.015)		(0.016)
Conscientiousness	0.025		-0.002
	(0.021)		(0.004)
Extraversion	-0.003		0.007
	(0.018)		(0.009)
Agreeableness	0.007		0.007
	(0.019)		(0.011)
Neuroticism	0.004		-0.001
	(0.017)		(0.009)
Constant	0.921***	1.052***	1.040***
	(0.259)	(0.075)	(0.135)
Observations	398	1474	1603
R^2	0.151	0.071	0.111
Land Fixed Effects	Yes	Yes	Yes

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Linear probability model. The dependent variable equals 1 if the subject was offered the module and accepted and 0 if they were offered the module and declined. Column 1 uses the "Unrestricted Covariates Set" approach, and column 2 the "Restricted Covariates Set" approach, and column 3 the "Missing Indicator" method. See section 3.4 in the main text for detail.

Table B.2: Binomial Test Results

	Unobservable			**Observable**		
report	Count	$p-val$	Corrected	Count	$p-val$	Corrected
1	85	< 0.001	< 0.001	94	< 0.001	0.003
2	86	< 0.001	< 0.001	116	0.154	1.000
3	80	< 0.001	< 0.001	117	0.183	1.000
4	120	0.291	1.000	125	0.581	1.000
5	119	0.251	1.000	133	0.890	1.000
6	165	0.003	0.032	126	0.646	1.000
7	148	0.141	1.000	122	0.383	1.000
8	160	0.011	0.114	150	0.098	0.982
9	175	< 0.001	< 0.001	161	0.008	0.088
10	180	< 0.001	< 0.001	174	< 0.001	0.001

Note: For each treatment, the number of observations and the p-value resulting from a binomial against the theoretical prediction. The Corrected column presents the corrected *p-value* using a Bonferroni approach.

Table B.3: Covariate effects using reports for both tasks

	report	**Unrestricted** report	**Restricted** report	**Missing Indicator** report
Unobservable First	0.080	0.002	0.092	0.112
	(0.114)	(0.231)	(0.120)	(0.113)
Observable	-0.313***	-0.611***	-0.255**	-0.313***
	(0.105)	(0.219)	(0.111)	(0.106)
Age		-0.003	0.001	-0.001
		(0.011)	(0.005)	(0.005)
Female		-0.383	-0.201	-0.227*
		(0.244)	(0.130)	(0.127)
Education		-0.073	-0.053**	-0.075***
		(0.056)	(0.026)	(0.025)
Net income		-0.072	-0.054*	-0.033
		(0.078)	(0.033)	(0.029)
Unemployed		-0.623	-0.253	-0.171
		(0.390)	(0.224)	(0.209)
Other Employment		0.486	0.144	0.126
		(0.371)	(0.188)	(0.180)
Retired		-0.001	-0.158	-0.061
		(0.417)	(0.211)	(0.204)
GDR		-1.289***	0.447*	0.307
		(0.442)	(0.242)	(0.226)
Risk willing		0.002	0.006	-0.004
		(0.049)	(0.026)	(0.026)
Moderate		-0.473*	0.038	0.040
		(0.254)	(0.132)	(0.129)
Extreme Left		-0.316	0.052	0.120
		(0.552)	(0.350)	(0.357)
Extreme Right		-0.405	-0.031	0.092
		(0.392)	(0.303)	(0.290)
Religious		0.363		0.061
		(0.265)		(0.160)
CRT		0.138		0.056
		(0.146)		(0.174)
Trust		-0.102		-0.042
		(0.230)		(0.097)
Patience		-0.020		-0.005
		(0.049)		(0.025)
Openness		-0.020		-0.027
		(0.099)		(0.055)
Conscientiousness		0.197		0.186***
		(0.145)		(0.071)
Extraversion		-0.028		0.067
		(0.118)		(0.057)
Agreeableness		0.276**		0.048
		(0.132)		(0.068)
Neuroticism		0.120		0.084*
		(0.105)		(0.050)
Constant	5.711***	3.981**	6.537***	4.818***
	(0.306)	(1.769)	(0.560)	(0.907)
Observations	2636	700	2366	2636
Adjusted R^2	0.012	0.034	0.012	0.018
Land Fixed Effects	Yes	Yes	Yes	Yes

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: OLS regression where the dependent variable is the reported value $report_{i,t}$ in each task t. Robust standard errors clustered at the individual level. Column 2 use the "Unrestricted Covariates Set" approach, column 3 the "Restricted Covariates Set" approach, and column 4 the "Missing Indicator" method. See section 3.4 in the main text for detail.

Table B.4: OLS Regression Missing Indicator - full output

	report
Unobservable First	0.110
	(0.113)
Observable	-1.062
	(0.785)
Female	-0.444***
	(0.168)
Observable × Female	0.435
	(0.231)
Age	-0.003
	(0.007)
Observable × Age	0.006
	(0.010)
GDR	0.558**
	(0.256)
Observable × GDR	-0.508
	(0.262)
Education	-0.064**
	(0.031)
Observable × Education	-0.020
	(0.036)
Education Missing	-0.802
	(0.503)
Net income	-0.035
	(0.042)
Observable × Net income	0.003
	(0.055)
Net income Missing	0.227
	(0.281)
Unemployed	-0.047
	(0.269)
Other Employment	0.241
	(0.238)
Retired	-0.293
	(0.273)
Observable × Unemployed	-0.235
	(0.395)
Observable × Other Employment	-0.216
	(0.350)
Observable × Retired	0.469
	(0.375)
Religious	0.139
	(0.205)
Observable × Religious	-0.160
	(0.260)
Religious Missing	0.218
	(0.156)
Moderate	0.030
	(0.168)
Extreme Left	-0.298
	(0.491)
Extreme Right	-0.270
	(0.438)
Observable × Moderate	0.016
	(0.231)
Observable × Extreme Left	0.842
	(0.678)
Observable × Extreme Right	0.698
	(0.649)
Risk willing	-0.003
	(0.035)
Observable × Risk willing	-0.002
	(0.050)
CRT	0.111
	(0.092)
Observable × CRT	0.041
	(0.119)
CRT Missing	0.466***
	(0.164)
Trust	-0.013
	(0.117)
Observable × Trust	-0.066
	(0.134)
Trust Missing	-1.460*
	(0.825)
Patience	0.000
	(0.031)
Observable × Patience	-0.012
	(0.040)
Patience Missing	1.461*
	(0.854)
Extraversion	0.029
	(0.074)
Observable × Extraversion	0.073
	(0.101)
Extraversion Missing	2.307***
	(0.857)
Conscientiousness	0.202**
	(0.087)
Observable × Conscientiousness	-0.032
	(0.113)
Conscientiousness Missing	1.807*
	(1.030)
Openness	-0.053
	(0.071)
Observable × Openness	0.052
	(0.096)
Openness Missing	-1.988***
	(0.696)
Neuroticism	0.025
	(0.064)
Observable × Neuroticism	0.119
	(0.086)
Neuroticism Missing	-1.883***
	(0.498)
Agreeableness	0.068
	(0.083)
Observable × Agreeableness	-0.043
	(0.108)
Agreeableness Missing	1.680***
	(0.594)
Constant	5.192***
	(0.979)
Observations	2636
R^2	0.047
Land Fixed Effects	Yes

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: OLS regression where the dependent variable is the reported value in each task t ($report_{i,t}$). Robust standard errors clustered at the individual level. "Missing Indicator" method approach for handling missing data.

C Pre-registration

Pre-analysis Plan: A Study on Truth-telling Using a Representative Sample

Ciril Bosch-Rosa[1,2], Levent Neyse[3,4,5], and Daniele Nosenzo[6]

[1] Technische Universität Berlin, Germany
[2] Colegio Universitario de Estudios Financieros, Madrid, Spain
[3] WZB, Berlin, Germany
[4] DIW, Berlin, Germany
[5] IZA, Bonn, Germany
[6] Department of Economics and Business Economics, Aarhus University, Denmark

August, 2021

Abstract

In this study, we will investigate the relationship between honesty and observability using a large sample of the German population. For this purpose, we included in the SOEP-IS 2020 wave a within-subject experiment consisting of the lying paradigms from Fischbacher and Föllmi-Heusi (2013) and Gneezy et al. (2018). Our main research question is whether observability has an impact on the lying behavior of respondents. Additionally, we will investigate the individual and socio-economic correlates of lying behavior under observable and unobservable conditions. The correlation between lying and observability, along with the micro and personality data of SOEP, will allow us to shed light on the behavioral correlates of intrinsic and image-related motives for lying. Our results should be of broad interest in the social sciences as well as to policy makers willing to understand better the impact of their interventions.

1. Aim of the study and background

The lying paradigm, developed by Fischbacher and Föllmi-Heusi (2013) (FF2013 hereafter), has played a particularly important role in the growth of recent literature studying the lying behavior of economic agents. However, its implementation is mostly limited to laboratory settings (For detailed literature reviews, see Abeler et al., 2019; Gerlach et al.; 2019: Jacobsen et al. 2018), with some exceptions (e.g. Cohn et al., 2015; Dai et al., 2017; Cohn et al., 2017, Abeler et al., 2014).

An important design characteristic of FF2013 is that, due to the nature of the task, the experimenter does not know whether any individual subject is lying or not. Therefore, the interpretation of the data relies on the comparison between the empirical distribution of reports and the theoretical distribution under the assumption of truth-telling.

Gneezy et al. (2018) (GKS2018 hereafter) introduce a computerized task which allows experimenters to observe whether an individual subject is lying. By comparing their results obtained in their task to those of FF2013, GKS2018 show that subjects lie less when they know they are being observed. This result highlights the existence of image concerns beside an intrinsic preference for truth telling. The meta-analysis of Abeler et al. (2019) supports this differentiation in motives behind truth-telling.

The purpose of the current study is to uncover the socio-economic and personality correlates of the different motives for truth telling (intrinsic vs image motivated). For this purpose, we integrate a module in the German Socio Economic Panel Innovation Sample (SOEP-IS) containing the FF2013 and GKS2018 tasks in a within-subject design. The advantage of running this study in a household panel survey is the existence of a large set of variables that we can correlate with lying behavior, including socio-economic background, personality, social preferences and cognitive abilities. This will make our results relevant to a wide variety of disciplines, including sociology, political science, or psychology.

2. Methods

2.1. Data Collection

The experiment was run in the 2020-2021 wave of the German Socio-Economic Panel Innovation Sample (SOEP-IS). SOEP is a longitudinal survey study that has been active since 1984 and it has more than 30.000 respondents (Goebel et al., 2019). SOEP-IS, on the other hand, is a separate panel study that was established in 2012 and it has a sample of 5000 adult respondents, representative of the German population. The SOEP-IS survey involves not only standard socio-economic survey items but also annual innovative modules that researchers can use to integrate new questions.

The data collection of our truth-telling innovative module was performed in the 2020-2021 wave of the SOEP IS between September and December 2020. The dataset was planned to be delivered to us in May 2021. In accordance with the agreement with the SOEP-IS team, the data will not be given to us until this pre-analysis plan is completed and published online at an online repository.[1]

We asked to collect data on 1500 individuals, although the final number may vary depending on response rate. We will drop those participants who have incomplete data on either the GKS2018 or FF2013 tasks.

2.2.1. Experimental Design

We follow a within-subject design with two treatments (tasks). The first task (GKS2018), follows the design of Gneezy et al. (2018). The second task (FF2013) follows the design of Fischbacher and Föllmi-Heusi (2013).[2] The order of the tasks is randomized across participants.

[1] The questions can be addressed to the director of SOEP-IS, Prof. Dr. David Richter.

[2] Note that unlike in Fischbacher and Follmi-Heusi (2013), we use a 10-sided die instead of a 6-sided die. This was done to have more variance in the reported outcomes and for comparability with the Gneezy et al. (2018) task.

In the FF2013 task participants are asked to (privately) roll a 10-sided die [0,9] and to report the outcome without showing it to the experimenter. The reported number will be the payoff for this task, so if a participant reports a 4, her payoff is 4 euros, if she reports a 2 the payoff is 2 euros, etc.. Since the true outcome cannot be checked by the experimenter, the participant has an incentive to lie. The GKS2018 task, on the other hand, is a computerized version of the FF2013. In this case, participants are asked to click on one of the 10 black boxes on a computer screen. Each box contains an integer from 0 to 9 and participants are asked to report the number shown in their clicked box. Again, the payoff for this task will be a one to one transformation of the reported value into euros (if a 3 is reported, the payoff is 3 euros, if an 8 is reported the payoff is 8 euros, etc.) independent of the true value ofthe clicked box.

The crucial difference between the two tasks is that in FF2013 the researcher only observes the report of a participant, while in GKS2018 the researcher knows *both* the report and the outcome of the participant.

At the beginning of the module, participants are aware of the existence of two tasks but the instructions for each task are read immediately before each task. Participants are also told at the start of the experiment that they will be paid randomly for only one of the two tasks.

2.2. Statistical Analyses and Variables

Our analysis will be based on a type-I error probability of alpha = 0.05. Our analysis will be based on OLS regressions as well as univariate tests, as we specify in the hypotheses section below.

2.2.1. Dependent Variables:

Following FF2013 and GKS2018, our main dependent variable is the reported value in each task ($report_{ij}$).

2.2.2. Independent Variables:

Our analysis will include a list of variables which we believe can explain participants' differences in behavior across treatments. We present them in three groups: i) Treatment variables, ii) Demographic and socio-economic variables, iii) Personality and preference variables. Please note that including the mentioned variables is subject to not having too many missing values and having enough variance in the reported values.

2.2.2.1. Treatment Variables

Two variables will capture the type of task and the order in which tasks were played:

- $Task_i$ [0,1]: $task_i$ = 1 for the *GKS2018* task.
- $Order_i$ [0,1]: $order_i$ = 1 for those participants that were presented first with the *GKS2018* task.

2.2.2.2. Demographic and Socio-Economic Variables

We will include 9 demographic and socio-economic variables in our regression analyses.

These are:

- Gender
- Age
- Education
- Net monthly household income
- Employment status
- Religiosity
- Political orientation
- State of residence
- A dummy for whether the participant has grown up in the East Germany (DDR)

2.2.2.3. Personality and Preference Variables

Our third set of independent variables will measure psychological and behavioral traits of participants. These are:

- Risk preferences measured using the non-incentivized SOEP variable (Annually asked)
- Cognitive ability, measured using the Cognitive Reflection Test (Innovation module 2020)
- Interpersonal trust (Annually asked, self-reported)
- Big five (Asked irregularly- Available in the release 2019)
- Personal patience (Asked irregularly- Previously asked in 2018)

3. Primary Hypothesis Tests

Our analysis focuses on two questions:

1) First, we want to replicate the result that lying is more likely under unobservable than observable conditions. To this aim we will compare the distributions of reports under the FF2013 task (unobservable) and the GKS2018 task (observable). Additionally, we will compare the distribution of reports for each task to its respective theoretical distribution assuming no lying.

2) Second, we want to study whether individual characteristics (socio-economic and personality/preferences variables) correlate differently with reporting behavior across the two different conditions. Differences in correlations can shed light on the importance of specific individual characteristics for the intrinsic and image-related costs of lying. In particular, if a specific

characteristic correlates with behavior more strongly under GKS2018 than FF2013, this indicates that that characteristic is important for image-motivated lying.

To address the first question, we will follow the literature and use binomial tests to compare each empirical distribution with its truthful distribution and Kolmogorov–Smirnov tests to compare the two distributions with one another.

To address the second question we will use OLS regression analysis estimated with robust (Huber-White) standard errors. We will regress subjects' reports on the set of independent variables listed earlier. Since each subject will enter twice in the dataset (one observation is their report in the FF2013 task, the other is the report in the GKS2018 task), SE will be clustered at the subject level. We will run various types of nested models where we first introduce our set of experimental variables, demographic/socio-economic variables and personality/preference variables without interacting them with one another. Our model will be:

$Report_{ij}$ = $Task_i$+ $Order_i$ + Demographic/socio-economic $variables_i$ + Personality/preference $variables_i$ + state fixed effects

This model allows us to confirm the robustness of our univariate analysis (based on binomial and Kolmogorov–Smirnov tests) about differences between reports under the FF2013 and GKS2018 tasks (Task).

In a second model we will interact the Treatment variable with all Demographic/socio-economic variables + Personality/preference variables. Thus, our second model will be

$Report_{ij}$ = $Task_i$+ $Order_i$ + Demographic/socio-economic $variables_i$ + Personality/preference $variables_i$ + $Task_i$x Demographic/socio-economic $variables_i$ + $Task_i$x Personality/preference $variables_i$ + state fixed effects

We will not interact the $Task_i$ dummy with the state of residence dummies as we use the state dummies as fixed effects and we do not expect any differential association between state of residence and motives for lying.

4. Unregistered Exploratory Analyses

Our experiment was run with a subsample of SOEP-IS. So, while we can specify the hypotheses for those variables that were previously responded to by *the whole sample*, we cannot do the same for the modules which were asked to only a part of the sample (since at the time of pre-registering this study we do not know which of these variables will be available to us in terms of sufficient sample size).

Therefore, we might run exploratory analyses with variables that are not included in the list presented earlier. The manuscript will clearly state which variables are exploratory (i.e., unregistered) and which not.

References

Abeler, J., Becker, A., & Falk, A. (2014). Representative evidence on lying costs. *Journal of Public Economics*, 113, 96-104.

Abeler, J., Nosenzo, D., & Raymond, C. (2019). Preferences for truth-telling. *Econometrica*, *87*(4), 1115-1153.

Cappelen, A. W., Sørensen, E. Ø., & Tungodden, B. (2013). When do we lie?. *Journal of Economic Behavior & Organization*, *93*, 258-265.

Cohn, A., & Maréchal, M. A. (2018). Laboratory measure of cheating predicts school misconduct. *The Economic Journal*, 128(615), 2743-2754.

Cohn, A., Maréchal, M. A., & Noll, T. (2015). Bad boys: How criminal identity salience affects rule violation. *The Review of Economic Studies*, 82(4), 1289-1308.

Dai, Z., Galeotti, F., & Villeval, M. C. (2017). Cheating in the lab predicts fraud in the field: An experiment in public transportation. *Management Science*, 64(3), 1081-1100.

Erat, S., & Gneezy, U. (2012). White lies. *Management Science*, *58*(4), 723-733.

Fischbacher, U., & Föllmi-Heusi, F. (2013). Lies in disguise—an experimental study on cheating. *Journal of the European Economic Association*, *11*(3), 525-547.

Fosgaard, T. R. (2020). Students cheat more: Comparing dishonesty of a student and a representative sample in the laboratory. *The Scandinavian Journal of Economics*, 122(1), 257-279.

Gerlach, P., Teodorescu, K., & Hertwig, R. (2019). The truth about lies: A meta-analysis on dishonest behavior. *Psychological Bulletin*, 145(1), 1.

Gneezy, U., Kajackaite, A., & Sobel, J. (2018). Lying aversion and the size of the lie. *American Economic Review*, 108(2), 419-53.

Gneezy, U., & Kajackaite, A. (2020). Externalities, stakes, and lying. *Journal of Economic Behavior & Organization*, *178*, 629-643.

Jacobsen, C., Fosgaard, T. R., & Pascual-Ezama, D. (2018). Why do we lie? A practical guide to the dishonesty literature. *Journal of Economic Surveys*, 32(2), 357-387.

Kajackaite, A., & Gneezy, U. (2017). Incentives and cheating. *Games and Economic Behavior, 102*, 433-444.

D Correlation between propensity to have missing data and response to treatment

SOEP-IS randomized the modules that subjects are presented for all of our pre-registered covariates. However, for ethical and data quality reasons, participants are always allowed to not answer certain modules and/or questions. In this section, we examine whether selection in or out of these modules/questions is orthogonal to the response to our treatment.

In principle, orthogonality may not hold for a number of reasons. First, there may be a correlation between willingness to answer survey questions and image concerns. For instance, people who have stronger image concerns may find it harder to decline the request to complete a survey module. Second, SOEP-IS allows researchers to specify inclusion/exclusion criteria for the participants in the modules they propose. For instance, a researcher may specify that their survey module is only offered to men. This may generate imbalances in the covariates available between different groups of people, potentially correlating with the treatment of interest.

To study the extent to which this is an issue for our study, we create a dummy variable $complete_i$, which takes a value of one when participant i has *no missing values* across all the pre-registered covariates, and zero otherwise. This allows us to identify the participants who answered all modules/questions used in our regressions and test whether they differ in their response to observability. We run a regression of subjects' reports in the lying tasks on the $complete_i$ indicator interacted with the treatment variable *Observable*. Table D.1 shows a statistical difference at the 10% significance level for both models reported ($p - val = .096$ and $p - val = .097$, respectively) with observability having a stronger effect on those subjects that have no missing data. Moreover, the combined effect of Observable and the interaction term (Observable × Not Missing) yields a coefficient of -0.612 (SE = 0.259, $t = -2.363$, $p = 0.018$) in Column 1 and -0.612 (SE = 0.260, $t = -2.354$, $p = 0.019$) in Column 2, both significant at the 0.05 level. This indicates there may be some interdependence between the propensity to have missing data and the response to our treatment, which justifies our careful approach to handling missing data discussed in the main text.

Table D.1: Treatment interaction with non-missing covariate indicator

	(1) report	(2) report
Observable	-0.205*	-0.205*
	(0.119)	(0.120)
Not Missing	-0.0763	-0.0736
	(0.179)	(0.179)
Observable × Not Missing	-0.407*	-0.407*
	(0.244)	(0.245)
Constant	5.282***	5.763***
	(0.0863)	(0.299)
Observations	2636	2636
Adjusted R^2	0.005	0.014
Land Fixed Effects	No	Yes

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: OLS regression with standard errors clustered at the individual level.

E Extra Figures

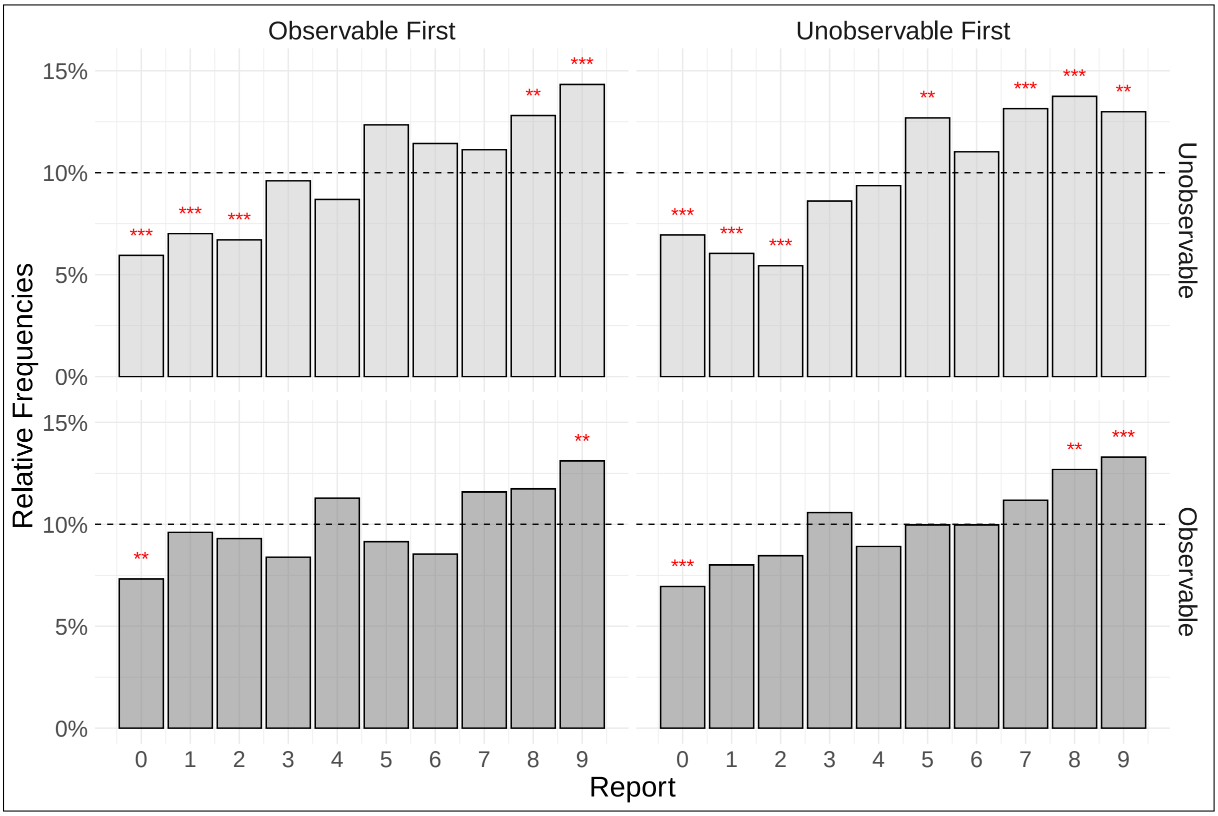

Figure E.1: Empirical distribution of reports in the unobservable task (top row) and observable task (lower row) for the Observable First subjects (left column) and Unobservable First (Right Column). The dotted horizontal bar marks the expected relative frequency of each report under the assumption of truthful reporting. The stars indicate significance levels based on binomial tests comparing the observed and expected frequency of each report (*** = 1% level; ** = 5% level).