DISCUSSION PAPER

// BENOÎT LE MAUX, DAVID MASCLET, AND SARAH NECKER

Monetary Incentives and the Contagion of Unethical Behavior

Leibniz
Association

ZEW

Monetary Incentives and the Contagion of Unethical Behavior[*]

Benoît Le Maux[†]
David Masclet[‡]
Sarah Necker[§]

February 2021

Abstract

We analyze both theoretically and empirically how monetary incentives and information about others' behavior affect dishonesty. We run a laboratory experiment with 560 participants, each of whom observes a number from one to six with there being a payoff associated with each number. They can either truthfully report the number they see or lie about it in order to increase their payoff. We vary both the size of the payoff (Low, High, and Very High) and the amount of information about others' dishonesty (With and Without Information). We first find that dishonesty falls in the Very High treatment. Second, while social information has on average at most a weak positive effect, there is a strong effect if the accuracy of individuals' beliefs is accounted for. Third, social information and payoffs do not interact with each other.

[†]University of Rennes 1, CREM-CNRS, Condorcet Center; benoit.le-maux@univ-rennes1.fr.

[‡]University of Rennes 1, CREM-CNRS, Condorcet Center, Cirano/Montreal; david.masclet@univ-rennes1.fr.

[§]Walter Eucken Institute; University of Freiburg; ZEW - Leibniz-Zentrum für Europäische Wirtschaftsforschung GmbH Mannheim; necker@eucken.de.

1. Introduction

Dishonesty is a major concern in modern societies. Cheating in exams, fare-dodging, doping in sport, curriculum-vitae inflation, questionable research practices, tax evasion, misrepresentation of output at work, the misuse of public funds and other fraud generate large societal costs. Dishonesty is at the heart of the principal-agent problem, and can cause substantial damage by eroding trust, creating uncertainty, reducing efficiency, and being harmful for collaboration. Given its significant economic consequences, understanding what lies behind dishonest behavior is central.

Economists have worked on the determinants of unethical behavior for decades. It is traditionally assumed that dishonesty results from the comparison of the expected associated pecuniary benefits and costs (Becker, 1968; Allingham and Sandmo, 1972). However, more-recent work has shown that there is less dishonesty than that predicted by the standard model (for overviews see, e.g., Gerlach et al., 2019; Abeler et al., 2019). One common explanation is that individuals incur intrinsic costs by lying, although whether and how these costs respond to external factors is still subject to debate.

While the standard economic model of crime predicts that dishonesty should rise with the payoffs, the empirical evidence is not so clear-cut. A number of recent contributions in the deception game have shown that senders are more likely to lie to receivers as the incentives to do so rise (e.g., Gneezy, 2005; Sutter, 2008; Dreber and Johannesson, 2008; Erat and Gneezy, 2012; Kajackaite and Gneezy, 2017), but work on incentives in non-strategic games mostly finds that dishonesty does not vary with payoffs (Mazar et al., 2008; Wiltermuth, 2011; Fischbacher and Föllmi-Heusi, 2013; Gino et al., 2013; Hugh-Jones, 2016; Andersen et al., 2018; Abeler et al., 2019).[1] Balasubramanian et al. (2017) find a hump-shaped relationship between dishonesty and the stakes. In Cohn et al. (2019), a lost wallet is more likely to be returned if it contains more money. Last, Kajackaite and Gneezy (2017) show that dishonesty

---

[1]The finding that monetary incentives increase dishonesty in strategic but not in non-strategic games is confirmed in the meta-analysis of Gerlach et al. (2019). We should bear in mind that strategic games such as the deception game conflate distributional preferences and lying costs, as lying affects not only the absolute payoffs but also the relative payoff as compared to other individuals.

increases with payoffs when concerns about being exposed as a liar are removed.

It might in addition be thought that individuals are more responsive to payoffs when they can justify their own lie by a larger fraction of others being dishonest. Only little is known about these social-norm effects. While informing subjects about others' behavior has been shown to have at most a weak effect on average dishonesty (Diekmann et al., 2015; Kroher and Wolbring, 2015; Abeler et al., 2019; Akin, 2019), this effect seems to be stronger when individual beliefs are taken into account (Rauhut, 2013).

We here contribute to the existing literature by investigating the impact of both social information and monetary incentives on dishonesty. We ask the following questions. Is cheating influenced by the size of payoffs in the context of a repeated setting? Does information about others' behavior matter, and produce contagion in behavior? Do the effects of information and payoffs interact with each other? We address these questions in a laboratory experiment, and develop a theoretical model of the decision to be dishonest.

Our experiment is based on a modified version of the non-strategic cheating game developed by Kajackaite and Gneezy (2017). Participants observe a number from one to six where a payoff is associated with each number. They can either truthfully report the number they see or lie about it in order to increase their payoff (see also Gneezy et al., 2018; Abeler et al., 2019). We vary both the size of the payoff (Low, High, and Very High) and information about others' dishonesty (With and Without Information). The variable of interest is the extensive margin: the fraction of subjects who cheat. We predict the decisions taken by the participants in our cheating game from a dynamic model of lying choices in which agents have heterogeneous beliefs about the true fraction of liars.

To the best of our knowledge, there is no existing work on the interaction between social information and payoffs in shaping dishonesty. Another advantage of our work here is that we can observe cheating at the individual level, and so do not have to rely on statistical distributions to infer dishonesty. This is important, as Hilbig and Thielmann (2017) suggest that lying is heterogeneous, with some agents not responding to payoffs, while others lie more as payoffs rise and a third group lies less. Last, the participants play 20 rounds of our cheating

game, allowing us to observe possible contagion or convergence over time.

We show that dishonesty is sharply lower with Very High payoffs (where participants can earn from 4 to 24 Euros), as compared to High (1 to 6 Euros) and Low (0.10 to 0.60 cents) payoffs, so that subjects are less likely to cheat only for the highest stakes. In line with our theoretical predictions, information about others has, on average, a relatively small or no effect on lying. Participants are responsive to social information if the accuracy of their beliefs is taken into account. Those who overestimate (underestimate) the fraction of liars are less (more) likely to be dishonest when they receive this information, so that the two adjustment effects cancel each other out. Last, we find no evidence of any interaction between social information and payoffs. Dishonesty increases significantly over time in all payoff conditions, yet with a similar trend with and without information.

The remainder of the paper is structured as follows. Section 2 describes the experimental design, and Section 3 sets out the theoretical predictions. The empirical results then appear in Section 4. Last, Section 5 concludes.

2. Experimental design

2.1. Cheating game

Our game is a modified version of the "observed game" developed by Kajackaite and Gneezy (2017). It is played for 20 periods. Figure 1 provides an overview of the experimental procedure. In each period, subjects are privately shown six boxes labeled "a, b, c, d, e and f" on their screen. The numbers from one to six are randomly assigned to these boxes, but participants do not know which number is in which box. Participants reveal the numbers by clicking on the boxes and report anonymously the number associated with the first box that they opened (e.g., Shalvi et al., 2011; Fischbacher and Föllmi-Heusi, 2013). A higher reported number is associated with a higher payoff, as described in Section 2.2.[2]

---

[2]After our repeated cheating game, we ran a separate variant using the *strategy method*. Participants were asked to report a number for each possible outcome of the box and each possible fraction of others cheating. We do not show the results from this game, as participants had difficulty in understanding how to fill in the lengthy table (containing $6 \times 5 = 30$ numbers).

Figure 1: Experimental flow
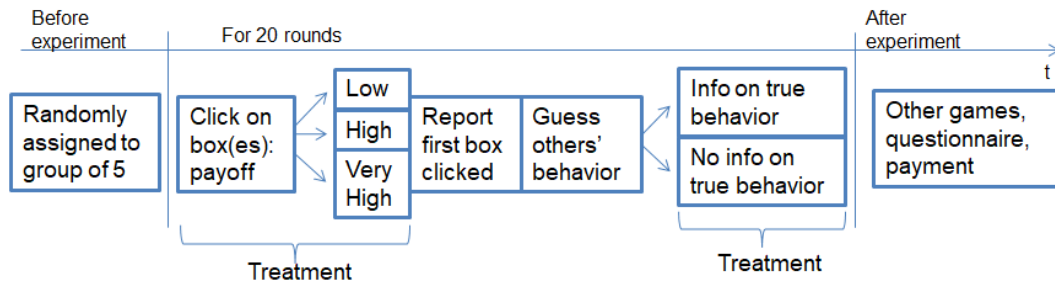
Participants can either truthfully report the number they observe or lie and report a higher number in order to increase their payoff. In line with an expanding body of literature (see the recent review by Abeler et al., 2019) our controlled laboratory experiment has no penalties, and participants are told that their decisions are anonymous and private (see the instructions in Appendix A.1). The purpose is to test how and whether dishonesty is influenced by payoffs and information in the absence of any material costs. While most real-life situations involve a positive probability of detection and punishment, this simplification is an asset here, as only intrinsic lying costs will matter for the honesty decision. Note that Potters and Stoop (2016) and Dai et al. (2017) both show that simple games like that used here predict behavior in the field. Furthermore, the participants' answers to the post-experiment questionnaire confirm that they did not worry about penalties or punishment of any kind.

Unlike in the basic coin-tossing or die-rolling games, we measure dishonesty at the individual level, and thus do not have to rely on statistical distributions in order to infer it. This is an advantage, as we expect heterogeneity in behavior (Hilbig and Thielmann, 2017) and beliefs about what others are doing (Rauhut, 2013). Observing dishonesty at the individual level allows the classification of individuals into cheating types based on their behavior across periods (see Section 4.1). We can also create groups based on the observed gap between beliefs and the true fraction of liars (see Section 4.2). In addition, this approach allows us to inform individuals about the observed fraction of dishonest subjects in the previous period. To minimize the feeling of being observed, the instructions make clear that mistakes are not

punished and could be unintentional.

Note that a number of contributions have used cheating games in which dishonesty is observable (e.g., Gibson et al., 2013; Faravelli et al., 2015; Gneezy et al., 2018; Kocher et al., 2018; Abeler et al., 2019). Observing dishonesty at the individual level is also possible in deception games (e.g., Gneezy, 2005) and tax-evasion games (see, e.g., Malezieux, 2018). Gneezy et al. (2018) and Abeler et al. (2019) compare behavior in a basic cheating game to that in an observed cheating game. Both find that dishonesty falls with observability, suggesting that individuals care about their reputations.

2.2. Treatments

We use a 2×3 between-subjects design, as shown in Table 1. The first dimension is the level of monetary incentives, i.e. the payoff per reported number (see Table 2). In the treatments with low payoffs (hereafter Low), participants can earn between 10 cents (if they report a one) and 60 cents (if they report a six). In the high-payoff treatment (High), these payoffs are multiplied by 10, so that participants can earn between 1 and 6 Euros. Last, in the very-high payoff treatment (Very High), the low payoffs are multiplied by 40, with earnings of between 4 and 24 Euros.

The second dimension relates to the amount of social information. At the beginning of the experiment, participants are randomly assigned to a group with five members. The composition of the groups remains unchanged over the 20 periods. In the treatments without information, subjects play 20 rounds of the game without any information about the behavior of the other four group members. With information, participants are informed at the end of each period about the percentage of the others in the group who correctly reported their first click. Comparing the With and Without Information treatments allows us to investigate the effect of social information on dishonesty.

In all six treatments, we elicit subjects' beliefs about others' over-reporting after the subject reports the number that they saw. Following Rauhut (2013), we are thus able to establish if beliefs are important in explaining the reaction to social information. This

6

Table 1: Treatments

Reward / Information	Low	High	Very High	N
Without	Without/Low N=70	Without/High N=95	Without/Very High N=95	260
With	With/Low N=90	With/High N=120	With/Very High N=90	300
N	160	215	185	560

Table 2: Payoffs (in Euros)

Reported number / Treatment	1	2	3	4	5	6
Low	0.10	0.20	0.30	0.40	0.50	0.60
High	1	2	3	4	5	6
Very High	4	8	12	16	20	24

elicitation of beliefs was incentivized. In the With Information treatments, the belief was measured before the receipt of information about others' behavior.

2.3. Procedure

The laboratory experiment took place at the LABoratory of EXperiments in Economics and Management LABEX-EM (CREM-University of Rennes 1, France) in September 2017, March to May 2018, and September 2018. Ztree was used to program the experiment, and participant recruitment was organized via Orsee. Before the experiment started, participants had to sign a consent form. They were then randomly allocated to the desks. Instructions were given on paper and read aloud before the game started.

At the end of the experiment, risk aversion was elicited using the task devised by Holt and Laury (2002). Furthermore, to control for possible cognitive biases, participants were asked to play a memory game designed to measure difficulties in memorizing numbers. Last, participants completed a post-experimental questionnaire on their socio-demographic characteristics, attitudes, and perception of the experiment (see Appendix A.2 for the data description). This post-experimental questionnaire allows us to test whether participants understood the experiment and in particular that mistakes would not be punished. The average payment was 19.94 Euros. Appendix A.2 contains the summary statistics, showing that the

sample was balanced across treatments.

3. Theoretical predictions

To illustrate the decision that agents face in our cheating game and derive testable hypotheses, we here set out a theoretical model of the decision to be dishonest. In our theory, as in the experiment, we focus on a setting in which dishonesty is not punished.

In each period of our cheating game, individuals observe a random number. They can either truthfully report the number seen and be paid m or report a higher number and be paid n (where n depends on the size of the lie). The payment scheme in the Low, High, and Very High treatments is designed such that m and n are multiplied by the same factor x as payoffs rise. The pecuniary advantage from lying is $U(xn) - U(xm)$, where U denotes the utility of wealth.

Following the existing literature, we assume that individuals face intrinsic costs of lying. These include (1) a fixed Kantian cost k (as in Figuières et al., 2013; Kajackaite and Gneezy, 2017; Abeler et al., 2019), (2) a variable cost V that rises with xn, i.e., agents worry more about their reputation as the payoff obtained by dishonesty rises (Kajackaite and Gneezy, 2017), and (3) norm-based costs that vary with the fraction $\hat{\beta}_t$ of other agents who are thought to be dishonest in period t (Figuières et al., 2013; Le Maux et al., 2019; Abeler et al., 2019).

For simplicity, we assume that the marginal cost of deviating from the social norm is one, so that beliefs are directly translated into lying costs. Beliefs $\hat{\beta}_t$ depend on two factors: any information received about the true fraction β_{t-1} of liars in period $t-1$, and the individual type of either optimistic or pessimistic beliefs. Pessimistic individuals have a higher reference point r^H and expect a greater fraction of agents to be dishonest, as compared to optimists who have a lower reference point of r^L:

$$\hat{\beta}_t^j = \alpha \beta_{t-1} + (1-\alpha)r^j, \quad j = L, H. \tag{1}$$

The parameter α allows us to distinguish between the With and Without Information

treatments. Without information, α is constrained to be zero, so that beliefs are the reference points; with information, any value of $\alpha \in (0,1]$ may hold and reflects the degree to which agents react to the information about the true fraction of liars. Even when information is provided, agents may not adjust their beliefs perfectly (i.e., $\alpha < 1$, so that the agent type continues to play a role, with a higher α implying that agents are more responsive to the information). Note that we have $\hat{\beta}_t^j = \beta_{t-1} + (1-\alpha)(r^j - \beta_{t-1})$. If r_j is above (below) β_{t-1}, then type j overestimates (underestimates) the true fraction.

For each agent type, $j = L$ or H, the fixed cost of cheating k_i is uniformly distributed over the unit interval. Agent i of type $j = L$ or H with parameter k_i will cheat in period t if:

$$\underbrace{U(xn) - U(xm)}_{\text{Pecuniary advantages}} > \underbrace{k_i + V(xn) - \hat{\beta}_t^j(\beta_{t-1}, r^j, \alpha)}_{\text{Lying costs}}, \quad j = L, H. \tag{2}$$

Let k_t^j denote the cheating cost at which agents are indifferent between lying and honesty:

$$k_t^j = U(xn) - U(xm) - V(xn) + \hat{\beta}_t^j. \tag{3}$$

Any j-type agent with Kantian morality under k_t^j will lie about his/her true state. Integrating over all agents yields the fraction of agents who are dishonest in period t:

$$\begin{aligned} \beta_t &= pk_t^L + (1-p)k_t^H \\ &= U(xn) - U(xm) - V(xn) + \alpha\beta_{t-1} + (1-\alpha)\bar{r}, \end{aligned} \tag{4}$$

where p and $(1-p)$ are the population shares of low and high types, and $\bar{r} = pr^L + (1-p)r^H$ is the average expectation of others' behavior when no information is received. From Equation (4), we obtain the partial derivative of β_t with respect to the payoff level x:

$$\frac{\partial \beta_t}{\partial x} = nU'(xn) - mU'(xm) - nV'(xn). \tag{5}$$

Proposition 1 illustrates the possible cases under the assumption of risk neutrality.[3] While dishonesty is more profitable as payoffs rise (through the first term on the right-hand side of Equation 5), the utility of the honest payoff (the second term) and lying costs (the third term) also increase. With a low constant marginal cost of lying ($V'' = 0$ and $V' < \frac{n-m}{n}U'$), lying becomes more prevalent as x rises (Proposition 1a). With a high constant marginal cost of lying ($V'' = 0$ and $V' > \frac{n-m}{n}U'$), lying will instead fall (Proposition 1b). If the marginal cost of lying is not constant ($V'' \neq 0$), we obtain either a maximum (V">0) or a minimum (V"<0). As such, two different values of x can yield the same fraction of liars (Proposition 1c).

Proposition 1. The effect of payoffs under risk neutrality:

(a) *if $V'' = 0$ and $V' < \frac{(n-m)}{n}U'$, then $\frac{\partial \beta_t}{\partial x} > 0$,*

(b) *if $V'' = 0$ and $V' > \frac{(n-m)}{n}U'$, then $\frac{\partial \beta_t}{\partial x} < 0$,*

(c) *if $V'' \gtrless 0$, then $\frac{\partial^2 \beta_t}{\partial x^2} \lessgtr 0$.*

Proof. Assuming constant marginal utility U', Equation (5) becomes $(n-m)U' - nV' > 0$. The second derivative determines whether we have a maximum ($V'' > 0$), a minimum ($V'' < 0$), or a constant marginal cost of lying ($V'' = 0$). □

To analyze the effect of social information, we appeal to the concept of a stationary point. Equation (4) shows that the fraction of liars rises with past dishonesty under the treatment With Information ($\alpha \neq 0$). The resulting long-run equilibrium is:

$$\beta_{\text{With}} = \frac{U(xn) - U(xm) - V(xn) + (1-\alpha)\bar{r}}{1 - \alpha}. \tag{6}$$

On the contrary, without information ($\alpha = 0$) we have:

$$\beta_{\text{Without}} = U(xn) - U(xm) - V(xn) + \bar{r}. \tag{7}$$

The effect of social information, defined as the gap between β_{With} and β_{Without}, hence depends

---

[3]In the empirical results there is no significant association between dishonesty and risk-aversion. Risk-neutrality does not thus seem to be a restrictive hypothesis. See Gneezy et al. (2018) for a similar assumption.

on two factors: (1) whether information is exogenously provided (whether α is different from zero) and (2) whether agents actually process the information about the true fraction under the With Information treatment (the value of α in Equation (6)). Moreover, as shown in Proposition 2 below, there is a third factor: (3) the position of the average reference point $\bar{r}$ relative to the true fraction β_{Without}. It is only when agents erroneously estimate the true fraction on average that the social information has an effect.

Proposition 2. The average effect of information: *if $\bar{r} \lesseqgtr \beta_{\text{Without}}$, then $\beta_{\text{With}} \gtreqless \beta_{\text{Without}}$.*

Proof. We have $\beta_{\text{With}} > \beta_{\text{Without}}$ iff $[\beta_{\text{Without}} - \bar{r} + (1-\alpha)\bar{r}]/(1-\alpha) > \beta_{\text{Without}}$, which is equivalent to $\bar{r} < \beta_{\text{Without}}$. $\qquad\square$

Replacing β_{t-1} with β_{With} and β_{Without} in Equation (3) yields the equilibrium share of liars among j-types, with and without information:

$$\beta_{\text{With}}^{j} = \beta_{\text{With}} + (1-\alpha)(r^{j} - \bar{r}) \quad \text{and} \quad \beta_{\text{Without}}^{j} = \beta_{\text{Without}} + (r^{j} - \bar{r}), \quad j = L, H. \qquad (8)$$

Since $r^{L} < \bar{r} < r^{H}$, we expect differences in the fraction of liars by type. However, these differences will be smaller when there is information, as stated in Proposition 3. H-types, who expect a large fraction of others to lie, are on average more dishonest and push the equilibrium upwards; L-types, who are optimistic about others' honesty, pull the equilibrium downwards. These effects counteract each other and yield convergence in behavior.

Proposition 3. The effect of information on types: $|\beta_{\text{With}}^{j} - \beta_{\text{With}}| < |\beta_{\text{Without}}^{j} - \beta_{\text{Without}}|$, $j = L, H$.

Proof. We have $|\beta_{\text{With}}^{j} - \beta_{\text{With}}| = (1-\alpha)|r^{j} - \bar{r}|$ and $|\beta_{\text{Without}}^{j} - \beta_{\text{Without}}| = |r^{j} - \bar{r}|$. As $\alpha \in (0,1]$ in the With Information treatment, the first term is necessarily smaller than the second. $\qquad\square$

Last, payoffs and social information are shown to interact with each other. Let Δ stand

for the average treatment effect of social information:

$$\Delta = \beta_{\text{With}} - \beta_{\text{Without}} = \frac{\alpha}{1-\alpha}(\beta_{\text{Without}} - \bar{r}). \qquad (9)$$

The partial derivative of Δ with respect to x is:

$$\frac{\partial \Delta}{\partial x} = \frac{\alpha}{1-\alpha}\frac{\partial \beta_{\text{Without}}}{\partial x}. \qquad (10)$$

As $\alpha/(1-\alpha)$ is positive, any change in payoffs that affects β_{Without} is exacerbated if agents are informed about the true extent of norm violations. There are multiplier effects. As demonstrated in Proposition 4, the magnitude of the interaction, however, depends on how elastic beliefs are to the true fraction. If α is under 0.5, we expect little interaction between payoffs and information. We will see convergence in types, as stated in Proposition 3, and the equilibrium β_{With} towards which the outcome converges is stable. If instead α is above 0.5, any deviation from the norm by a small number of individuals will be greatly amplified.

Proposition 4. The interaction of information and payoffs: *We have $\frac{\partial \Delta}{\partial x} \gtreqless \frac{\partial \beta_{Without}}{\partial x}$ if $\alpha \gtreqless$ 0.5.*

Proof. From Equation (10), $\alpha/(1-\alpha)$ is above one iff $\alpha > 0.5$. $\qquad\qquad\square$

4. Empirical results

The extensive margin of dishonesty (the fraction of liars) comes from the comparison of the number observed (the first box clicked on) to the number reported. We create a dummy for the subject reporting a number above that which he or she first clicked on.[4] We find that – even though we are able to observe dishonesty – a substantial fraction of individuals take the opportunity to cheat. Over the whole sample, 29% report a number above that they first clicked on. This figure is remarkably similar to that in Gneezy et al. (2018) in their observed

---

[4]The 67 observations in which subjects report a number below that on which they first clicked are not counted as dishonesty (downward lying).

Figure 2: The percentage of dishonest subjects over time by payoff condition

(a) Without Information   **(b)** With Information
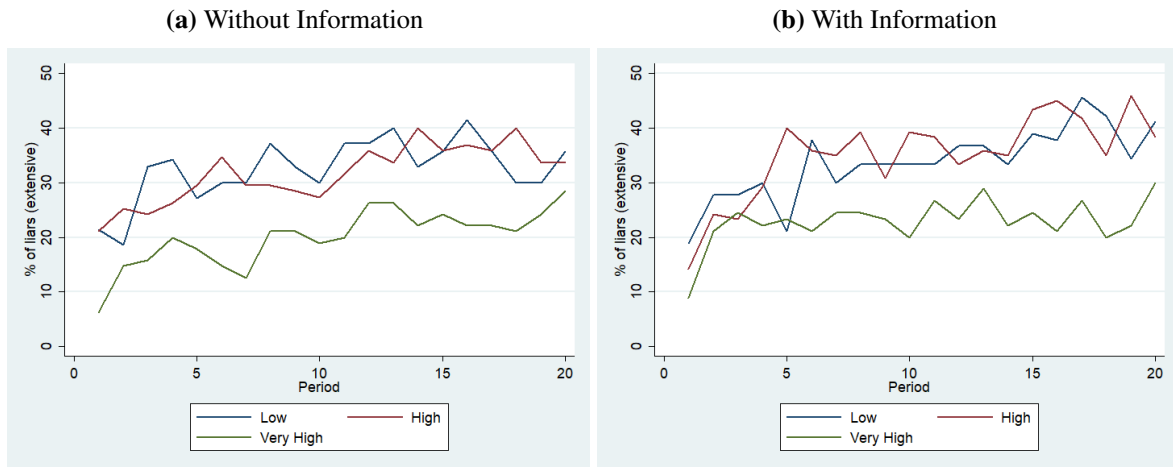
Note: Without Information covers 5,200 observations, and With Information 6,000 observations.

game. Those who lie predominantly do so to the maximum extent, with 85% reporting a six. This result is in line with Gneezy et al. (2013) and Andersen et al. (2018). The analysis of the extent of the lie, i.e. the gap between the observed and reported numbers (the intensive margin of dishonesty), appears in Appendix A.3.

4.1. The effects of payoffs

Panels (a) and (b) of Figure 2 show the percentage of liars over time across all payoff conditions, with and without social information. In both panels, dishonesty is lower for very-high payoffs. Without information, the average percentage of liars over the 20 periods is 32.5% in the Low condition and 31.6% in the High condition (Δ High-Low=0.9 ppts, p-value from a two-tailed Mann-Whitney test=0.597). This figure is much lower (20%) in the Very High treatment (Δ Very High-Low=12.5 ppts, p-value=0.000). A similar pattern occurs with information, as shown in panel (b). The average percentage of liars is 33.6% in Low, 35.1% in High (Δ High-Low=1.5 ppts, p-value=0.325), and 22.9% in Very High (Δ Very High-Low=10.7 ppts, p-value=0.000).

There is a similar rising trend in dishonesty in all treatments. The fraction of cheaters increases sharply between periods one and two, which may be considered as a pure learning

13

Figure 3: The percentage of liars by payoff and first number seen
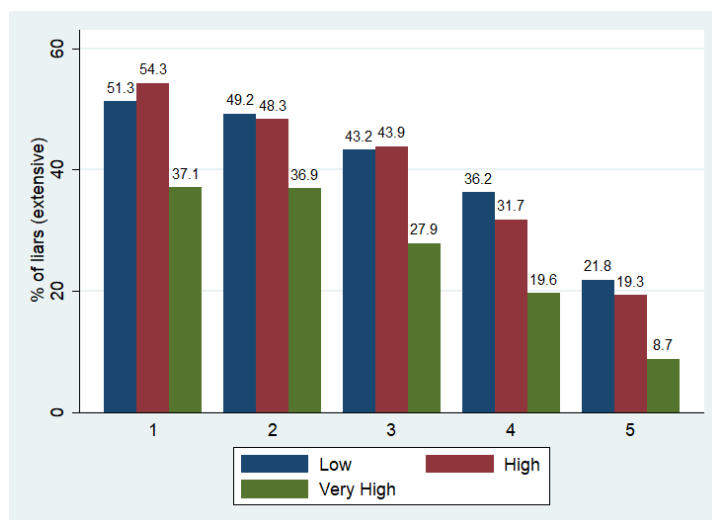
Note: Low covers 3,200 observations, High 4,300 observations, and Very High 3,700 observations. Box 6 does not appear as dishonesty is not possible when uncovering a six.

effect. However, the percentage of cheaters also continues to rise in subsequent periods.[5]

Result 1. *(a) The percentage of liars is significantly lower in the Very High compared to the Low and High treatments. (b) In all treatments, the percentage of liars rises over time.*

The lower dishonesty observed in the Very High treatment suggests that the greater honest payoff and/or the lying costs outweigh the larger dishonesty payoff (Proposition 1). To investigate which explanation fits best, we examine whether a higher first number seen, i.e., a higher honest payoff, has a differential effect across payoff treatments. Figure 3 shows the dishonesty levels in each of the true states by payoff; for ease of reading we pool the information and no-information treatments. Our results confirm previous findings that dishonesty falls with a higher true state (e.g., Gneezy et al., 2013; Gibson et al., 2013; Andersen et al., 2018). The relationship in Figure 3 is stable over the payoff treatments. The dishonesty gap between boxes 1 and 5 are 29.5 ppts in Low, 35 ppts in High, and 28.4 ppts in Very High. This stability suggests that participants care more about the box numbers (from

---

[5]Without information, the fraction of liars is significantly larger in the last five than in the first five periods in all three payoff conditions (Low Δ=8 ppts, p-value=0.027; High Δ=11 ppts, p-value=0.000; Very High Δ=8 ppts, p-value=0.000). The difference between Low and Very High is stable across periods (Period 1-5: Δ=11.9 ppts, p-value=0.000, Period 16-20: Δ=10.9 ppts, p-value=0.000). We see a similar trend in the treatments with information. Dishonesty rises from period 1-5 to period 16-20 in Low and High (Δ=15 ppts, p-value=0.000) but this increase is not significant in Very High (Δ=4 ppts, p-value=0.148).

Table 3: Cheating types by payoff and information

	N	Never cheater	Partial cheater	Always cheater
Panel A. By payoffs				
Low	160	35.6%	51.9%	12.5%
High	215	30.7%	57.7%	11.6%
Very high	185	48.1%	46.5%	5.4%
Panel B. By information				
Without	260	45.8%	44.2%	10%
With	300	31%	59.3%	9.7%

Notes: Based on behavior across the 20 rounds. A never cheater did not cheat in any of the periods in which the first number $\neq 6$, an always cheater cheated in all periods in which the first number $\neq 6$, and a partial cheater cheated in some but not all of the periods in which the first number $\neq 6$.

1 to 5), i.e., how far they are from box 6, and less about the Euro values associated with each box. In other words, the greater honest payoff in Very High may not explain the lower dishonesty in that treatment. The explanation would lie instead in a rise in intrinsic lying costs.

We also consider the extent to which the Very High result comes from a change in cheating types. The existing literature has categorized individuals according to their frequency of lying (e.g., Gneezy et al., 2013; Gibson et al., 2013; Hilbig and Thielmann, 2017). We take this approach and classify individuals as (1) "Always cheaters" (those who were dishonest in all periods by overreporting a figure), (2) "Partial cheaters" (those who were dishonest in at least one period), and (3) "Never cheaters" (those who were never dishonest). Table 3 shows the distribution of these three types in each treatment. The percentage of "Always cheaters" is relatively low, at around 10-12%, in all treatments.

We can see from panel (A) of Table 3 that the fraction of "Never cheaters" is significantly higher in Very High (Δ Very High-Low=12.5 ppts, p-value=0.019), and that of "Partial cheaters" (Δ Very High-Low=7.1 ppts, p-value=0.019) and "Always cheaters" (Δ Very High-Low=5.4 ppts, p-value=0.319) is lower. A closer look at "Partial cheaters" also shows that this type cheats less in Very High, with a percentage of 43.7% in Low, 41.5% in High, and 35.1% in Very High. The difference between Low/High and Very High is significant. The lower dishonesty in Very High thus partly comes from a shift from "Partial" or "Always

cheaters" to "Never cheaters", and partly from "Partial cheaters" reducing their dishonesty.

Result 2. *(a) In all treatments, the percentage of dishonest subjects falls with the true number. (b) The fraction of "Never cheaters" is higher in Very High compared to High and Low.*

*4.2. The effects of social information*

Figure 4 shows the fraction of cheaters over time by information conditions. It can be seen that social information has either no effect on dishonesty (Low treatment) or a small positive effect (High and Very High treatments). In Low, the average percentage of liars over the 20 periods is 32% without information and 34% with information (p-value on Δ with/without information=0.487). The analogous figures in the High and Very High treatments are 32% vs. 35% (p-value=0.016), and 20% vs. 23% (p-value=0.029).

Panel (B) of Table 3 shows that the percentage of Always cheaters is stable across the two treatments, while the percentage of "Never cheaters" falls by 15 percentage points, and that of "Partial cheaters" rises by the same amount (p-values=0.000). Information thus significantly increases the fraction of conditional cheaters. Our findings are summarized in Result 3.

Result 3. *(a) Information has either no or a small positive effect on dishonesty. (b) The percentage of "Partial cheaters" rises significantly with information.*

While we find a significant effect of information under High and Very High payoffs, the effect is only small in size on average. Our theory provides a number of explanations. It may first be that individuals without information form heterogeneous beliefs, but still correctly estimate the true fraction of liars on average (Proposition 2). Some might believe that individuals are more dishonest than they actually are (overestimators) while others think they are actually less dishonest (underestimators). If this were the case, on average, the belief could still be correct. Second, if under- and overestimators adjust their behavior to information in opposite ways, there may be no effect on average (Proposition 3). As we will see below, it is the second explanation that applies.

16

Figure 4: The percentage of dishonest individuals over time by information

(a) Low

(b) High
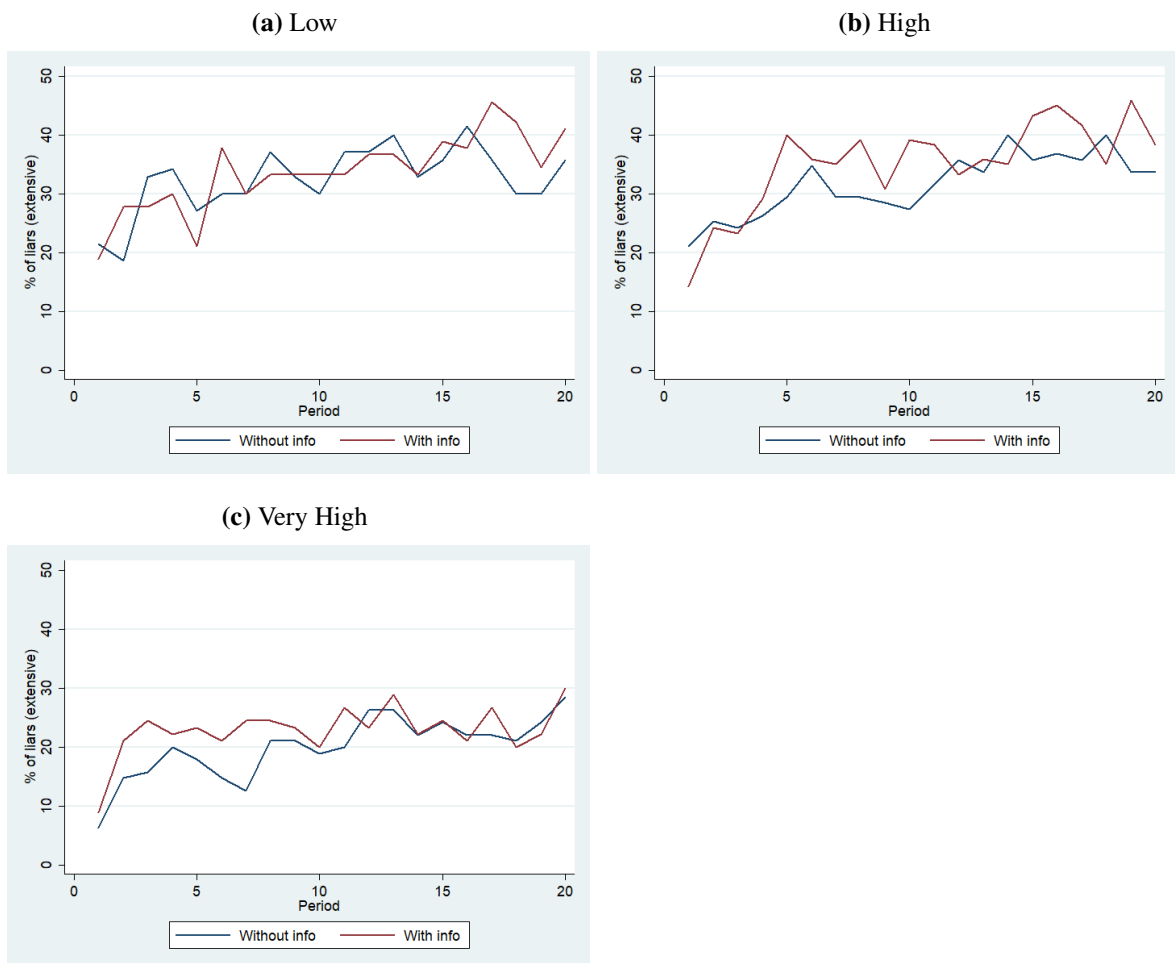
(c) Very High

Note: The Low, High and Very High treatments cover 3,200, 4,300 and 3,700 observations respectively.

Figure 5: Beliefs and the true percentage of others who are honest

(a) Without Information   **(b)** With Information
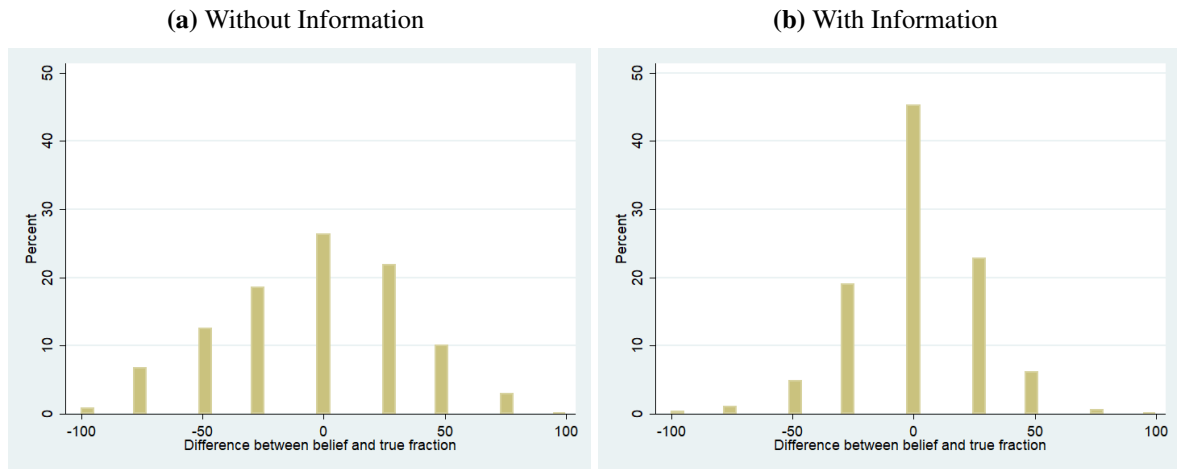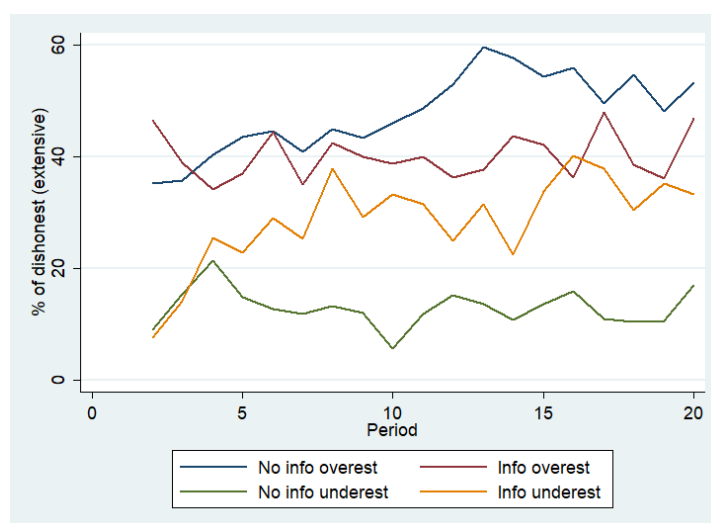

Notes: The Without Information treatment covers 5,200 observations and the With Information treatment 6,000 observations. Negative (positive) values imply that the true percentage honest is larger (smaller) than the percentage believed honest, i.e., the individual overestimates (underestimates) dishonesty.

Figure 5 plots the difference between beliefs about the percentage of others who are honest and the actual percentage. Panel (a) shows that beliefs deviate widely from the true percentage without information. On average, individuals significantly overestimate the true percentage of liars (average belief=68.1%; true fraction=72.1%; Δ=4.0 ppts, p-value=0.000). In other words, individuals misjudge the dishonesty of others at both the individual and average levels. In panel (b) of Figure 5, these differences narrow substantially when information is provided. However, significant average differences are still found (average belief=69.1%; true fraction=68.1%; Δ=-1.0 ppt, p-value=0.003).

Following our theory, we create two groups based on the observed gap between beliefs and truth. Underestimators think that the percentage of dishonest subjects is lower than it actually is, and overestimators believe that it is higher. Over- and underestimators represent 39% and 35% of the sample without information, and 25% and 30% with information. In the With Information treatment, the true percentage was provided at the end of each period (see Figure 1). We thus analyze the effect of information on the behavior of over- and underestimators in the subsequent period.

Figure 6 shows the level of dishonesty by information and over/underestimation. Across the 20 periods, overestimators are more dishonest than underestimators without information,

18

Figure 6: The percentage of liars by information and over-/underestimation


Note: The Without Information treatment includes 1,910 overestimators and 1,714 underestimators, and the With Information treatment includes 1,450 overestimators and 1,679 underestimators.

with figures of 48% and 13%, respectively (Δ=35 ppts, p-value=0.000). This gap narrows substantially with information, as the figure drops to 40% for overestimators and rises to 28% for underestimators (Δ=12 ppts, p-value=0.000). The evolution of this gap over time is also of interest. In the first five periods, the difference is 23% without information (p-value=0.000) and 22% with information (p-value=0.000); the corresponding figures in the last five periods are 39% (p-value=0.000) and 6% (p-value=0.090). Without information, we thus see divergence in the dishonesty of over- and underestimators, while on the contrary there is convergence with information. Looking at the size of the changes, underestimators react more strongly to information than do overestimators (overestimators: Δ=-8 ppts, p-value=0.000, and underestimators: Δ=16 ppts, p-value=0.000). This explains why we find on average a weakly positive effect of information on dishonesty.

Result 4. *(a) Individuals estimate dishonesty more correctly when information is provided. (b) Those who overestimated (underestimated) dishonesty in the previous period are more (less) likely to be dishonest themselves. (c) Underestimators react more strongly to information than do overestimators.*

19

Figure 7: The percentage of liars by payoff and information, conditional on over-/underestimation

(a) Overestimators
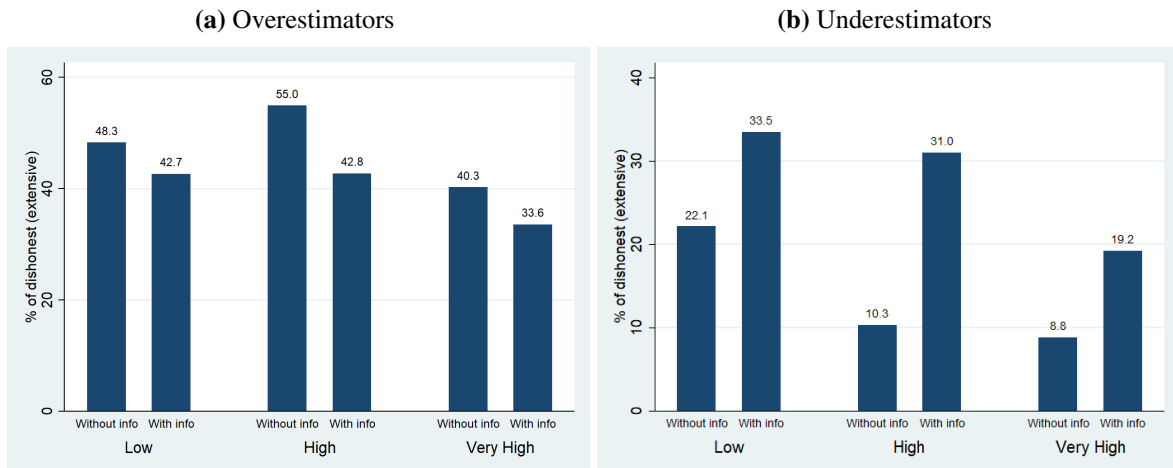
(b) Underestimators

Note: From left to right, the numbers of observations are as follows. Overestimators: 536, 403, 706, 636, 668, 411. Underestimators: 461, 490, 708, 752, 545, 437.

4.3. The interaction between payoffs and social information

We now ask whether payoffs and information interact with each other (Proposition 4). Figure 4 shows that the effect of payoffs is similar across the two information treatments. We evaluate the statistical significance of the interaction from the empirical bootstrap distribution of the difference (1,000 replications, sampling with replacement). The difference in dishonesty levels with and without social information does not vary across payoff conditions (e.g., Δ Low/Very high=1.8 ppts, p-value=0.398).

Our theory and the empirical results in Section 4.2 show, however, that it is important to consider heterogeneity in beliefs at the individual level. We further investigate potential interactions by considering whether information and payoffs interact with respect to the convergence of the dishonesty of over- and underestimators. Figure 7 (a) plots the dishonesty of overestimators by information and payoffs. The difference between the With and Without Information treatments is 5% in Low, 12% in High, and 6% in Very High. There is a similar pattern in panel (b) for underestimators, but in the opposite direction. The difference is -11% in Low, -21% in High, and -10% in Very High. In other words, the gap between the dishonesty of over- and underestimators is lower with information, regardless of the payoffs.

Result 5. *On average, payoffs and information about others' behavior do not interact.*

4.4. Regression analysis

We provide further statistical evidence via a regression analysis. As each subject is observed 20 times, we can use panel methods, and estimate random-effect probit models in which the dependent variable is the binary decision to lie. Standard errors are clustered at the group level to control for interdependence within groups.

The results appear in Table 4. In column (1), we control only for the treatments and the period. The estimated coefficients confirm that the percentage of liars is significantly lower with Very High payoffs and that social information has a positive but small significant effect. The estimated period coefficients indicate that the percentage of liars constantly rises over time.

Column (2) adds controls for socio-demographic characteristics and attitudes towards dishonesty (see Appendix A.2 for the data description). While the effect of Very High is slightly larger, the effect of social information becomes insignificant. In line with previous work (e.g., Abeler et al., 2019), we find that women cheat less. The religious and those who consider cheating less justifiable are also less likely to cheat. Risk aversion, measured by a Holt and Laury task, is not significant, while a greater number of correct answers in our memory game is negatively related to dishonesty. This game was introduced to help separate lying from mistakes resulting from memory. The positive coefficient suggests that over-reporting cannot be attributed to mistakes reflecting memory issues.

In column (3), we add controls for the first number seen. The negative effect of Very High rises further. In line with earlier results, we find that a higher first number produces less dishonesty.

The regression in column (4) controls for both the belief about how many others cheated (Believes X/4 dishonest) and the actual fraction dishonest in the previous period (Actual X/4 dishonest). The significantly-negative effect of Very High is unchanged, as is the insignificant effect of information. A higher expected fraction of dishonest others is strongly positively

21

Table 4: The marginal effects from dishonesty regressions

	(1) b/se	(2) b/se	(3) b/se	(4) b/se
Info	0.062*	0.049	0.066	0.064
	(0.033)	(0.032)	(0.050)	(0.048)
Payoff: High	0.020	0.003	-0.009	-0.012
	(0.048)	(0.048)	(0.084)	(0.078)
Payoff: Very high	-0.105**	-0.131***	-0.180**	-0.178**
	(0.042)	(0.041)	(0.077)	(0.070)
Period	0.007***	0.008***	0.010***	0.006***
	(0.001)	(0.001)	(0.001)	(0.001)
Age		-0.003	-0.003	-0.003
		(0.003)	(0.003)	(0.003)
Female (binary)		-0.060**	-0.077	-0.074
		(0.030)	(0.049)	(0.050)
Economics (binary)		-0.007	-0.002	0.007
		(0.031)	(0.048)	(0.049)
Risk aversion (1=Low to 10=High)		0.009	0.012	0.011
		(0.009)	(0.014)	(0.014)
Morality (1=High to 10=Low)		0.039***	0.052**	0.054***
		(0.011)	(0.020)	(0.020)
Religious (1=No to 10=High)		-0.020***	-0.025***	-0.025***
		(0.005)	(0.008)	(0.007)
Political orientation (1=Left to 10=Right)		-0.012*	-0.014	-0.015
		(0.007)	(0.013)	(0.013)
Financial knowledge (1=None to 3=High)		0.025	0.029	0.031
		(0.029)	(0.045)	(0.047)
Memory Game (2=Low to 9=High)		0.024**	0.036**	0.034**
		(0.011)	(0.017)	(0.017)
First box: 2			-0.021	-0.029
			(0.017)	(0.018)
First box: 3			-0.144***	-0.154***
			(0.027)	(0.026)
First box: 4			-0.273***	-0.305***
			(0.046)	(0.043)
First box: 5			-0.360***	-0.404***
			(0.070)	(0.065)
First box: 6			-0.376***	-0.419***
			(0.077)	(0.070)
Believes 1/4 dishonest				0.053***
				(0.014)
Believes 1/2 dishonest				0.079***
				(0.021)
Believes 3/4 dishonest				0.088***
				(0.025)
Believes all dishonest				0.135***
				(0.042)
Actual 1/4 dishonest				0.015
				(0.012)
Actual 1/2 dishonest				0.017
				(0.013)
Actual 3/4 dishonest				0.031*
				(0.017)
Actual all dishonest				0.042
				(0.030)
N	11200	11200	11200	10640

Notes: These are marginal effects from random-effect probit models, calculated as the probability of a positive outcome, assuming that the random effect for that observation's panel is zero. Standard errors clustered at the group level appear in parentheses. Significance levels: $*p \leq 0.1$, $**p \leq 0.05, ***p \leq 0.01$.

correlated to the probability of lying. By way of contrast, the actual previous-period behavior of others has no effect on average. However, it should be remembered that subjects were only aware of others' behavior in the With Information treatments.

To see whether the effect of others' behavior differs by information conditions, we estimate the column-(4) regression including interactions of beliefs and actual behavior with an information-treatment dummy. As the marginal effects of interactions cannot be determined from the estimated coefficients in random-effect probit models, we list the marginal effects of beliefs/actual behavior conditional on the information provided (see panel (A) of Table 5). These reveal that, with information, more observed liars in the group produces a higher likelihood of own cheating. As expected, the effect of others' behavior is insignificant in the treatment without information. The marginal effects of the belief variable show that information reduces the importance of beliefs.

We last consider whether payoffs and social information interact, via interaction terms in the column-(3) regression. The resulting marginal effects are listed in Table 5, panel (B). In line with our previous findings, dishonesty is lower in the Very High treatments, both with and without social information. However, these effects are similar across the With Information treatments, suggesting no significant interactions.

5. Conclusion

The question of how monetary incentives affect lying, and whether lying depends on others' behavior, has long puzzled economists. This paper contributes to the literature by analyzing these effects in the particular context of a non-strategic cheating game. Our variable of interest is the percentage of liars in our sample, although similar but smaller effects were found regarding the size of the lie (see Appendix A.3).

In our model, the intrinsic lying costs include a norm-based component that varies with others' dishonesty. Two types of agents are considered: those who overestimate the true percentage of liars and those who underestimate it. Individual behaviors are heterogeneous and endogenous to others' behavior. Our variables also have different time

Table 5: The marginal effects from dishonesty regressions including interaction terms

Variable	Conditional on	ME	SE
Panel A. Marginal effects from column (4) incl. believes*info, actual*info			
Actual 0/4 dishonest	Without Information	Ref.	Ref.
	With Info	Ref.	Ref.
Actual 1/4 dishonest	Without Info	-0.004	0.011
	With Info	0.023*	0.013
Actual 1/2 dishonest	Without Info	0.003	0.011
	With Info	0.021	0.013
Actual 3/4 dishonest	Without Info	0.007	0.025
	With Info	0.034**	0.015
Actual all dishonest	Without Info	-0.016	0.023
	With Info	0.059**	0.029
Believes 0/4 dishonest	Without Info	Ref.	Ref.
	With Info	Ref.	Ref.
Believes 1/4 dishonest	Without Info	0.077***	0.02
	With Info	0.025**	0.012
Believes 1/2 dishonest	Without Info	0.093***	0.025
	With Info	0.045***	0.017
Believes 3/4 dishonest	Without Info	0.12***	0.03
	With Info	0.032	0.022
Believes all dishonest	Without Info	0.143***	0.039
	With Info	0.077	0.051
Panel B. Marginal effects from column (3) incl. payoff*info			
Payoff Low	Without Info	Ref.	Ref.
	With Info	Ref.	Ref.
Payoff High	Without Info	-0.057	0.098
	With Info	0.035	0.07
Payoff Very High	Without Info	-0.157*	0.091
	With Info	-0.139**	0.065

Notes: These are marginal effects from random-effect probit models, calculated as the probability of a positive outcome, assuming that the random effect for that observation's panel is zero. Standard errors clustered at the group level appear in parentheses. Significance levels: $*p \leq 0.1$, $**p \leq 0.05, ***p \leq 0.01$.

profiles. While the two types both converge in their beliefs and behaviors according to the true fraction (Proposition 3), the convergence cancels out on average, yielding little change in the equilibrium (Propositions 2 and 4). From an empirical point of view, these results underline the importance of analyzing dishonesty at both the average and type levels. Last, as far as we are aware, we are also the first to consider whether social information and payoffs interact.

In our experiment, we first find that individuals are less likely to lie when payoffs are Very High (40 times higher than in the Low payoff treatment). This partly reflects a shift from "Partial" and "Always" to "Never cheaters", and partly "Partial cheaters" becoming less dishonest. Possible explanations are that the cost of lying rises with payoffs (e.g., Cohn et al., 2019), or that reputational concerns are more pronounced in an observed than an unobserved cheating game (Kajackaite and Gneezy, 2017). Another explanation is that large payoffs could increase concerns about the cost that is imposed on the experimenter (Gneezy, 2005; Fischbacher and Föllmi-Heusi, 2013). Gratitude has also been shown to enhance honest behavior (DeSteno et al., 2019). Unfortunately, we cannot disentangle these explanations.

Second, in line with both our theory and previous work (Diekmann et al., 2015; Kroher and Wolbring, 2015; Rauhut, 2013), information on others' dishonesty has, at best, a small average effect. As empirically suggested in Rauhut (2013), heterogeneity in beliefs matters here. In our experiment, overestimators become less dishonest with information, while underestimators become more dishonest. These opposite reactions explain why, on average, information about others' dishonesty has only a subdued effect.

Third, according to our theory, any impact of payoffs should be exacerbated when agents are informed about the true extent of norm violations. Surprisingly, we find no evidence of any interaction between payoffs and information. While dishonesty is lower in the Very High treatment, there is no evidence of more honesty when information is provided. Our model offers an explanation: while information affects behavior at the type level, the aggregate percentage of liars is little affected, with not much change in dishonesty over time.

We find rising dishonesty over time in all payoff conditions. This cannot be explained

by contagion in behavior, as it also appears in the Without Information treatment. This trend could show learning, but our game is straightforward and was understood by the participants, as shown by their answers to our post-experimental questionnaire ("How did you perceive the game?"). An alternative is that their perception of being observed (the subjective probability of being detected as a cheater) evolved over time. However, precautions were taken in the instructions to make clear that under- or overreporting could be unintentional and was neither observed nor punished. Participants' lying costs may also fall over time, as they get used to being dishonest (e.g., Garrett et al., 2016). Last, prior honesty may provide a "license" to carry out morally-questionable behavior later on (Blanken et al., 2015). The investigation of the reasons behind this finding is left for future research.

Another result that deserves attention is the impact of the true state, i.e. how much honesty is rewarded. In accordance with previous work (e.g. Gneezy et al., 2013; Andersen et al., 2018), lying is less frequent with higher values of the true number. However, a similar pattern is found for all payoff treatments, suggesting that what matters for lying is how far participants are from their best situation (the distance to box 6 in our game) rather than the Euro values associated with each possible true state. In line with this intuition, most participants who lied did so to the maximum possible extent when being dishonest (see Appendix A.3).

To conclude, while our results are consistent with a number of findings in the existing literature, and reveal no interactions between payoffs and information, they also raise questions about the underlying mechanisms. How do observability and reputational concerns affect lying costs in our game? Why is there a rising trend in dishonesty in all treatments? Would we find contagion in lying in other circumstances? The understanding of these and other issues offers a formidable agenda for future research.

References

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

Akin, Z. (2019). Dishonesty, social information, and sorting. *Journal of Behavioral and Experimental Economics 80*, 199–210.

Allingham, M. G. and A. Sandmo (1972). Income tax evasion: a theoretical analysis. *Journal of Public Economics 1*(3), 323–338.

Andersen, S., U. Gneezy, A. Kajackaite, and J. Marx (2018). Allowing for reflection time does not change behavior in dictator and cheating games. *Journal of Economic Behavior & Organization 145*, 24–33.

Balasubramanian, P., V. M. Bennett, and L. Pierce (2017). The wages of dishonesty: The supply of cheating under high-powered incentives. *Journal of Economic Behavior & Organization 137*, 428–444.

Becker, G. S. (1968). Crime and Punishment: An Economic Approach. *Journal of Political Economy 76*(2), 169–217.

Blanken, I., N. van de Ven, and M. Zeelenberg (2015). A meta-analytic review of moral licensing. *Personality and Social Psychology Bulletin 41*(4), 540–558.

Cohn, A., M. A. Maréchal, D. Tannenbaum, and C. L. Zünd (2019). Civic honesty around the globe. *Science 365*(6448), 70–73.

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

DeSteno, D., F. Duong, D. Lim, and S. Kates (2019). The grateful don't cheat: Gratitude as a fount of virtue. *Psychological Science 30*(7), 979–988.

Diekmann, A., W. Przepiorka, and H. Rauhut (2015). Lifting the veil of ignorance: An experiment on the contagiousness of norm violations. *Rationality and Society 27*(3), 309–333.

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

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

Faravelli, M., L. Friesen, and L. Gangadharan (2015). Selection, tournaments, and dishonesty. *Journal of Economic Behavior & Organization 110*, 160–175.

Figuieres, C., D. Masclet, and M. Willinger (2013). Weak moral motivation leads to the decline of voluntary contributions. *Journal of Public Economic Theory 15*(5), 745–772.

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

Garrett, N., S. C. Lazzaro, D. Ariely, and T. Sharot (2016). The brain adapts to dishonesty. *Nature Neuroscience 19*(12), 1727.

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

Gibson, R., C. Tanner, and A. Wagner (2013). Preferences for truthfulness: Heterogeneity among and within individuals. *American Economic Review 103*, 532–548.

Gino, F., S. Ayal, and D. Ariely (2013). Self-serving altruism? The lure of unethical actions that benefit others. *Journal of Economic Behavior & Organization 93*, 285–292.

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

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

Gneezy, U., B. Rockenbach, and M. Serra-Garcia (2013). Measuring lying aversion. *Journal of Economic Behavior & Organization 93*, 293–300.

Hilbig, B. E. and I. Thielmann (2017). Does everyone have a price? On the role of payoff magnitude for ethical decision making. *Cognition 163*, 15–25.

Holt, C. A. and S. K. Laury (2002). Risk Aversion and Incentive Effects. *American Economic Review 92*(5), 1644–1655.

Hugh-Jones, D. (2016). Honesty, beliefs about honesty, and economic growth in 15 countries. *Journal of Economic Behavior & Organization 127*, 99–114.

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

Kocher, M. G., S. Schudy, and L. Spantig (2018). I lie? We lie! Why? Experimental evidence on a dishonesty shift in groups. *Management Science 64*(9), 3995–4008.

Kroher, M. and T. Wolbring (2015). Social control, social learning, and cheating: Evidence from lab and online experiments on dishonesty. *Social Science Research 53*, 311–324.

Le Maux, B., S. Necker, and Y. Rocaboy (2019). Cheat or perish? a theory of scientific customs. *Research Policy 48*(9), 103792.

Malezieux, A. (2018). A practical guide to setting up your tax evasion game. *Journal of Tax Administration 4*(1), 107–127.

Mazar, N., O. Amir, and D. Ariely (2008). The Dishonesty of Honest People: A Theory of Self-Concept Maintenance. *Journal of Marketing Research 45*, 633–644.

Potters, J. and J. Stoop (2016). Do cheaters in the lab also cheat in the field? *European Economic Review 87*, 26–33.

Rauhut, H. (2013). Beliefs about lying and spreading of dishonesty: Undetected lies and their constructive and destructive social dynamics in dice experiments. *PloS one 8*(11).

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

Sutter, M. (2008). Deception through telling the truth?! Experimental evidence from individuals and teams. *Economic Journal 119*(534), 47–60.

Wiltermuth, S. S. (2011). Cheating more when the spoils are split. *Organizational Behavior and Human Decision Processes 115*(2), 157–168.

A.1.   Experimental instructions (translated from French)

General information: Welcome. This experiment will last about an hour and a quarter. We ask you to not communicate with the other participants for the duration of this experiment. We also ask you not to write on the Instructions. For your participation, you will receive a payoff of 2.50 Euros, whatever the decisions you make during the experiment. You can in addition increase your earnings. The experiment is divided into a number of sections, for which you will receive separate instructions. Your earnings will be calculated at the end of the experiment. You will receive by check all of your earnings during the different sections, in addition to the 2.50 Euros.

Treatment-specific information (varying payoffs and the existence of information): Please read these instructions carefully. You can earn a certain amount of money depending on your decisions. You will be randomly assigned to a group of five participants for a series of twenty periods. The group will remain unchanged over these twenty periods. In each period, you will see on the screen six boxes marked a, b, c, d, e, and f with numbers hidden behind them (1, 2, 3, 4, 5, and 6). These numbers are placed in a random order that will change each time. In each period, you will have to click on a box. Once you have clicked on it, you will see a number appear in bold. For example, in the screen below, you clicked on box "c" and the hidden number is 4.

You can click on the other boxes but only the first box on which you clicked matters. You will be asked to report the number you saw at the first click, as described in the screen below:

The number reported has to be between 1 and 6. Your earnings will be determined by the number you entered according to the following table (Example: High):
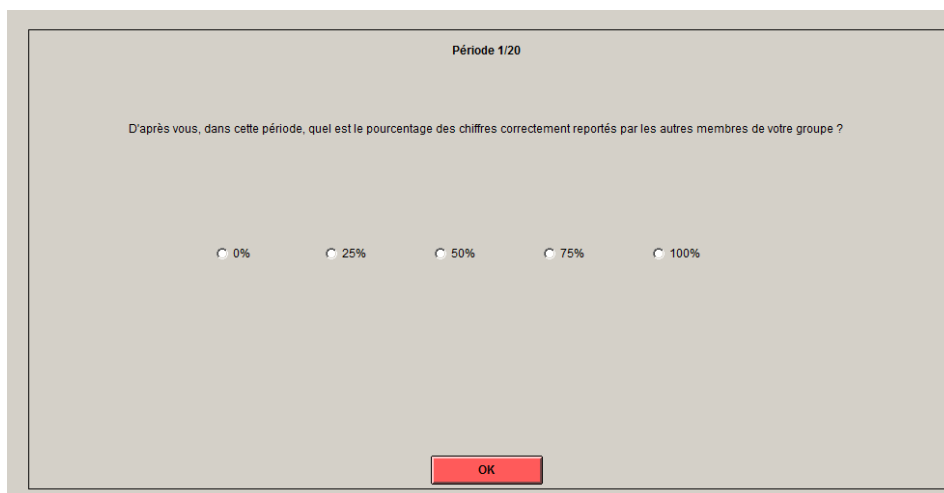
Reported number	1	2	3	4	5	6
Payoff	1 Euro	2 Euros	3 Euros	4 Euros	5 Euros	6 Euros

For example, if you wrote "1", you will receive 1 Euro. If you wrote "2", you will receive 2 Euros; if you wrote "3", you will receive 3 Euros, and so on. The higher the reported number, the higher your payoff.

In each period, six new boxes will appear on the screen. You will have to click on one of them and, again, fill in the input boxes. This part of the experiment will end after twenty periods.

At the end of the experiment one period will be drawn randomly to determine your payoff, and you will be paid according to the number you entered in this period.

In each period, we will also ask you to guess the percentage of those in your group who have correctly reported their first click (excluding yourself). For example, in the screen below, you are asked to choose between 0%, 25%, 50%, 75% or 100%.

[Next a screen will appear and inform you about the correct answer] *(text in brackets only in the information treatment).* One of the percentages you have entered will be randomly selected at the end of the experiment, and you will receive 50 Euro cents if you guessed correctly.

Important: your decisions are anonymous and private. The computer program does not allow us to link your identity to your answers. Errors will not be punished and the experiment will continue even in the case of mistakes. You will be paid according to what you report. If you have any questions, please raise your hand and we will come to see you!

Memory game: You will now participate in a memory game. Be careful: numbers will appear on the screen. You will have to memorize them and report the whole sequence. The list of numbers to remember will become longer and longer. It is forbidden to write them down or use a calculator. The participant in this session with the highest score in this memory game will receive two additional Euros. A random draw will be carried out in the case of a tie.

A.2. Descriptive statistics

Table A.1: Description and statistics of the control variables

Variable	Description	Total	Treatment 1	2	3	4	5	6
Age	Age of subject	21.50	22.67	22.57	22.96	20.18	19.97	21.34
Female	Is female (Y/N)	0.51	0.53	0.56	0.52	0.48	0.60	0.38
Economics	Is an Economics student (Y/N)	0.43	0.16	0.30	0.52	0.71	0.34	0.41
Risk aversion	Number of safe choices in Holt/Laury 2002 risk-aversion game: higher=more risk-averse	6.22	5.96	6.12	6.14	6.18	6.42	6.42
Cheating attitude	Average assessment of justifiability of 11 practices on a 1-10 scale (taken from the World Values Survey): higher=less moral	3.23	3.07	3.31	3.38	3.32	3.15	3.07
Religious	Importance of God in life: 1=No importance, 10=High importance	3.52	3.74	3.66	3.41	3.60	3.37	3.38
Political orientation	1 leaning to Left, 10 leaning to Right	4.61	4.59	4.26	4.36	5.08	4.75	4.48
Finance	Financial knowledge: 1=None, 2=Some, 3=High	2.33	2.29	2.36	2.20	2.45	2.29	2.34
Memory game	Points achieved in memory game: higher=better memory	6.23	5.99	6.20	6.11	6.36	6.28	6.38

A.3. Intensive margin: the determinants of the intensity of the lie

To analyze the size of the lie of those who are dishonest (the intensive margin of dishonesty), we create a binary variable for the subject lying to the maximum-possible extent.

As shown in Table A.2, a large majority cheat to the maximum possible extent by reporting a six when the true state was less than six. We find that the percentage of maximum cheaters falls with the payoff. Without information, the difference to Low is significant for High (p-value=0.001) and Very High (p-value=0.000); with information, the difference to Low is significant only for the Very High treatment (p=0.000). The percentage of maximum cheaters rises slightly but insignificantly over time.

Table A.2: Cheating to the maximum extent by payoff and information

| | Without Information | | | | With Information | | |
	N	All	Period 1-5	Period 16-20	N	All	Period 1-5	Period 16-20
Total	1436	0.866	0.852	0.873	1862	0.852	0.822	0.882
Low	455	0.934	0.914	0.95	606	0.87	0.858	0.883
High	601	0.874	0.841	0.906	843	0.871	0.841	0.914
Very High	380	0.774	0.788	0.741	413	0.787	0.744	0.806

We next look at the effect of over- and under-estimation on the probability of lying to the maximum extent. As shown in Figure A.1 (b), in the High and, in particular, Very High treatments, underestimators are more likely to lie to the maximum extent when information is provided, which is in line with the result from the extensive margin. However, in the Low treatment this same percentage falls. One possible explanation is that most underestimators already cheated to the maximum extent without information, precluding any further increase. Panel (a) shows that overestimators are slightly less likely to lie to the maximum extent under the With Information treatment, but that this difference is not significant for any of the payoff treatments.

Figure A.1: Maximum lie over-/underestimators by payoff and information

(a) Overestimators   **(b)** Underestimators
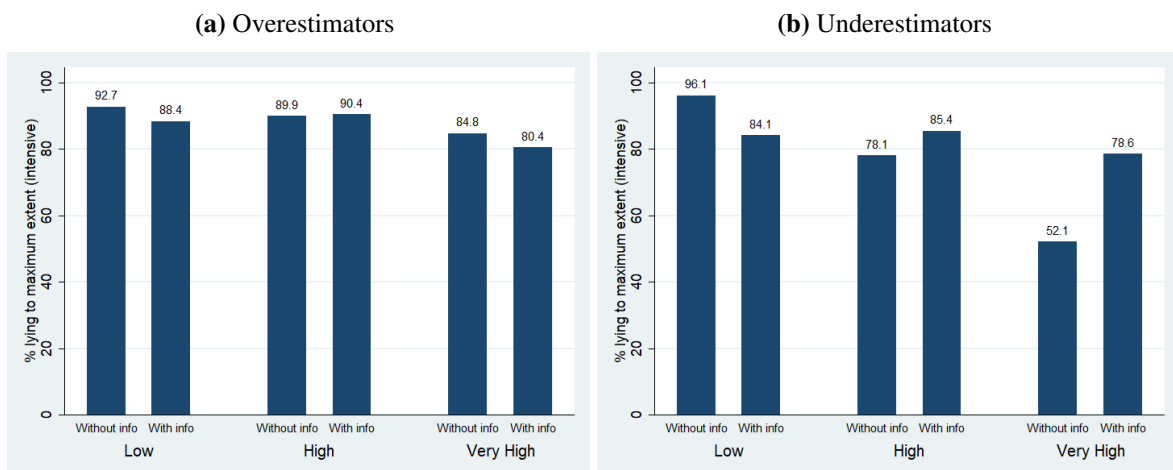
Note: Overestimators include 3,360 observations and underestimators 3,393 observations.

//

IMPRINT