Nobody likes a rat: On the willingness to report lies and the consequences thereof

Ernesto Reuben and Matt Stephenson

ABSTRACT

We investigate the intrinsic motivation of individuals to report, and thereby sanction, fellow group members who lie for personal gain. We further explore the changes in lying and reporting behavior that result from giving individuals a say in who joins their group. We find that enough individuals are willing to report lies such that in fixed groups lying is unprofitable. However, we also find that when groups can select their members, individuals who report lies are generally shunned, even by groups where lying is absent. This facilitates the formation of dishonest groups where lying is prevalent and reporting is nonexistent.

1. Introduction

Reporting on another's deceptive behavior is an act that arouses diverse and conflicting opinions. Children are sometimes scolded for being "tattle-tales" and "snitch" is a derogatory term used to describe those who report on others. Yet this act can also be deemed praiseworthy, as in the case of whistleblowers or crime informants. One stereotype portrays a whistleblower as an ungrateful coworker eyeing a lucrative book deal, while another brings to mind a self-sacrificing hero. In this paper, we study the intrinsic motives of people to report on others' lies and evaluate the consequences of this act when it can be rewarded by others.

A growing body of research in economics focuses on deception and the inclination of some people to tell the truth despite it being in their material interest to lie (Ellingsen and Johannesson, 2004; Gneezy, 2005).[1] For instance, Gibson et al. (in press) demonstrate that individuals are averse to lying (to varying degrees), but they can be tempted to lie when doing so is profitable enough.[2] We extend the work on lying aversion by studying whether the predilection to tell the truth also extends to the willingness to uphold a truth-telling norm by punishing and disassociating oneself from people who lie. If this is the case, the cost of deceiving others could be even higher than with lying aversion alone, which has important implications for, among other things, the transmission of information in asymmetric information (Kartik, 2009; Bolton et al., 2012) and the design of institutions that monitor the actions of economic agents (see discussion in Dyck et al., 2010).

---

[1] In this study, we do not distinguish between a preference for truth-telling and an aversion to lying. This distinction is important only if it is possible to not tell the truth and not lie at the same time, e.g., by remaining silent (see Sanchez-Pages and Vorsatz, 2009). We also do not differentiate between an aversion to lying and an aversion to being observed lying (Hao and Houser, 2011).

[2] Other experimental studies on lying include Charness and Dufwenberg (2006, 2010), Cai and Wang (2006), Fischbacher and Heusi (2008), Vanberg (2008), Hurkens and Kartik (2009), Lundquist et al. (2009), Sutter (2009), Rode (2010), Eisenkopf et al. (2011), Lopez-Perez and Spiegelman (2012), Erat and Gneezy (in press), and Gibson et al. (in press).

We run a laboratory experiment in which subjects play a repeated "whistleblowing" game. In each repetition of the game, subjects receive a random number that corresponds to their "real" earnings. Subsequently, they are given the opportunity to lie by overstating their earnings in order to receive a higher payoff. Importantly, subjects are divided into groups within which they can observe each other's real and stated earnings. Upon observing whether lying occurred, subjects have the opportunity to report others in their group. Subjects who are reported for overstating their earnings are sanctioned by a factor proportional to the size of their lie. However, subjects who report others do not receive any monetary benefits from their action. This game captures basic elements of many situations of interest. In particular, it explores situations where lying is individually profitable but, if discovered by a central authority, it is heavily sanctioned. Critically, the central authority does not monitor behavior—perhaps because it is prohibitively expensive to do so—and instead relies on individuals within the organization to report lying.[3] We are interested in whether enough people disapprove of and are willing to sanction lying (by reporting it) such that overstatements do not occur.

Some evidence suggests that people are willing to sanction those who tell them lies (Brandts and Charness, 2003; Croson et al., 2003; Sanchez-Pages and Vorsatz, 2007, 2009), even after controling for the economic damage the lie might have imposed (Eisenkopf et al., 2011). This evidence is consistent with the enforcement of a truth-telling norm. However, besides being in the pecuniary interest of the person who lies, in these studies a lie is to the detriment of the person being lied to. Therefore, lying also conveys an intention to hurt the person who is subsequently administering the punishment.[4] In this study, we test the willingness to punish liars even when lies do not affect the pecuniary interest of, and are not directed at, the person administering the punishment. If the punishment of lies occurs in

---

[3] A common factor for most whistleblowers is a low cost of obtaining information regarding the malfeasance and the ease with which it can be reported to the appropriate authority (Dyck et al., 2010).

[4] For evidence that unkind intentions trigger punishment see Falk et al. (2008).

this setting, it indicates that individuals consider lying *per se* as behavior that deserves to be sanctioned.[5]

In addition to studying motivations to report lies, we also evaluate the consequences of reporting. One could reasonably expect that sanctioning behavior that is generally disapproved of would be welcomed. However, some empirical evidence indicates that this is not always the case. For example, Dyck et al. (2010) report that the career prospects of employees who report corporate malfeasance are so dismal that it is actually surprising that people blow the whistle at all. Indeed, the existence of strong community norms against reporting others has been documented by both investigative journalists and academics and is well epitomized by the phrase "snitches get stitches" (Brown, 2007; Kahn, 2007). These reports point to fear of ostracism and punishment by their peers as one of the main reasons why people do not report others' wrongdoings (Whitman and Davis, 2007).

To introduce these types of peer effects in the experiment, we give subjects a say in who gets to join their group. Specifically, every three repetitions of the game, some subjects are randomly removed from their group and have to rejoin a group to continue to earn money. In order to join a group, however, displaced subjects must be accepted by a unanimous vote of the group's current members. To cast their vote, current group members are informed (i) whether displaced subjects reported others for lying, (ii) whether they were sanctioned for lying, and (iii) their mean stated earnings. This design allows us to determine whether subjects welcome people who report lies or whether they prefer to avoid them. Moreover, we can observe whether subjects who do not lie are the only ones who vote in favor of people who report lying or whether reporting lies is well regarded more generally. Similarly, we can observe whether subjects who regularly lie are the only ones who vote against those who report lying or whether reporting lies is generally poorly regarded. Finally, to determine the importance of these effects in the overall amount of lying and reporting, we run another treatment without voting where displaced subjects rejoin groups

---

[5] Erat and Gneezy (in press) show that some individuals are not willing to lie even when doing so is in the pecuniary interest of the person doing the lying and the person that is being lied to, which is further evidence that lying *per se* is the act that is disapproved of.

at random. These comparisons can give us insights regarding the formation and impact of anti-reporting groups.

Field research that explores the causes and motivations for reporting lies faces a complex task. For one, lies are difficult to observe, and even where lying is evident, extrinsic incentives may exist that can influence the reporting of lies beyond just an aversion to dishonesty. An important example is whistleblowing. In their review, Bowen et al. (2010) report that whistleblowing is affected by factors that are plausibly unrelated to reporting dishonest behavior, such as *qui tam* statutes that financially reward whistleblowers. Furthermore, as Schmidt (2003) points out, numerous selection effects make the evaluation of whistleblowing *per se* difficult. For instance, if unproductive employees are more likely to be whistleblowers then we would overestimate the impact of whistleblowing on employability if the job prospects of non-whistleblowers with average productivity are used as the counterfactual. Due to the difficulties posed by field research, an experiment provides the ideal setting to isolate the reasons and consequences for lying and reporting lies. In our experiment, there are no financial incentives for reporting lies and there is no externality associated with lying. Thus, we can be assured that reporting someone for lying reflects intrinsic motives such as upholding a norm of truth telling. Moreover, we avoid the aforementioned selection effects because all subjects are equally productive, they randomly leave groups, and they do not have a say in which groups they want to rejoin. In fact, by comparing the treatment with random reassignment to the treatment where subjects get to choose who joins their group, we can identify the importance of this type of selection on the prevalence of lying.

Our main findings are that, in addition to restraining themselves from lying, enough subjects report those who lie that in the average group lying is largely unprofitable. However, we also find that when group members have a say in who joins their group, subjects who report lies are generally shunned, making this behavior very costly. In fact, the subjects who have better prospects of rejoining (some) groups are those who lie. Accordingly, reporting lies is less common in this treatment and lying is more frequent, particularly in groups that develop a rule of no reporting lies.

2. Experimental design and procedures

2.1. The whistleblowing game

Here we describe the game used in the experiment in more detail. For simplicity, we describe the game already with the parameters used in the experiment. Consider a "society" composed of $i = 1, \dots, 12$ individuals and $g = 1, \dots, 3$ organizations. Each organization g is staffed by $n_g \in \{2,3\}$ individuals. The game is played repeatedly for nine periods and each period is divided into two stages.[6] In the first stage, each individual first observes her "true" earnings t_i, which are independently drawn from a uniform distribution with support $[0, t_g^{max}]$ where t_g^{max} are the maximum earnings in i's organization g. The value of t_g^{max} is increasing with the size of the organization. Specifically, we set $t_g^{max} = 300$ points for organizations of $n_g = 3$ individuals and $t_g^{max} = 225$ points for organizations of $n_g = 2$ individuals. After observing t_i, each individual simultaneously decides on the earnings she wishes to state s_i. Individuals are free to state any feasible earnings, i.e., $s_i \in [0, t_g^{max}]$. Barring any sanctions in the second stage, an individual's payoff is equal to her stated earnings and not her true earnings. In the second stage, all individuals observe both the true and stated earnings of others in their organization. Thereafter, they simultaneously decide whether they report their organization to a central authority. If at least one individual reports her organization, the organization is inspected and all individuals who overstated their earnings, i.e., chose $s_i > t_i$, are sanctioned by three times the overstated amount.[7] Hence, the payoff of individual i of organization g in a period is given by:

$$\pi_i = \begin{cases} s_i - 3(s_i - t_i) & \text{if } g \text{ is inspected and } s_i > t_i \\ s_i & \text{otherwise} \end{cases}.$$

At the end of the second stage, individuals are informed of the payoff and actions of all individuals in their organization. Hence, as in other studies of lying behavior, in our game

---

[6] The first stage is inspired by the design of Fischbacher and Heusi (2008).

[7] To avoid confusion in the experiment, subjects were given the opportunity to report only if one or more individuals overstated their earnings.

individuals have a monetary incentive to lie by overstating their true earnings, but in our case lying is profitable only if individuals believe that others will not report their actions.

Now we describe how organizational membership is determined. We refer to individuals who are part of an organization as being active and to those who are not as being inactive. At the beginning of the game, all individuals in the society are randomly assigned to one of the three organizations and are therefore active individuals. However, every third period (i.e., after periods 3 and 6 are played), one individual in each organization of $n_g = 3$ is selected at random to be separated from it and hence become an inactive individual.[8] Moreover, everyone in the society gets to observe the following information of each inactive individual: (i) their mean stated earnings over the last three periods, (ii) whether they reported their organization in the last three periods, and (iii) whether they were sanctioned for overstating their earnings in the last three periods. Before play resumes, inactive individuals have the opportunity to rejoin an organization. Individuals who remain inactive do not receive or state earnings and obtain a payoff of $\pi_i = 0$ points per period. In the experiment, we implement two treatments, each using a different procedure to convert inactive individuals into active individuals. In the *Random* treatment, all inactive individuals are randomly reassigned to organizations.[9] In the *Selection* treatment, inactive individuals must be accepted into organizations using the unanimity rule. Specifically, active individuals first indicate whether they accept or veto each inactive individual. Thereafter, inactive individuals are randomly assigned to an organization where they were unanimously accepted. If no such organization exists, the individual remains inactive for the next three periods (until the next vote or until the game ends). Meanwhile organizations with $n_g = 2$ individuals continue to play but with reduced maximum earnings of $t_g^{max} = 225$ points.

---

[8] Organizations with $n_g = 2$ individuals do not lose members. This was done to avoid the added complexity of disappearing organizations.

[9] Note that in Random nobody remains inactive and individuals are simply being randomly reassigned in a similar way as in experiments with a strangers matching protocol. Unlike these experiments, here only some individuals are reassigned and there are information asymmetries between individuals who join a group and those who are already in it (unless an individual is reassigned to the same group she used to belong to).

2.2. Theoretical predictions

We now briefly discuss the theoretical predictions of the game. We start with the assumption that all individuals are risk neutral and own-earnings maximizers. In this case, in both treatments and during all periods, all individuals always state the maximum earnings and they never report others because they are indifferent to the outcomes of other players and reporting carries the opportunity cost of lying.[10] Finally, in the Selection treatment, all inactive individuals are accepted because organizations with $n_g = 3$ individuals allow for higher stated earnings than organizations of $n_g = 2$ individuals and vetoing someone carries the risk of ending up with a smaller organization.[11]

We next discuss how the above predictions change if some individuals have an aversion to lying. Namely, they incur disutility if they tell a lie (examples of this approach are Ellingsen and Johannesson, 2004; Demichelis and Weibull, 2008; López-Pérez, 2012).[12] By assumption, models of lying aversion predict less overstating in the whistleblowing game as some individuals state their earnings truthfully. None of these models, however, consider motivations for individuals to report others for overstating. It is not our aim to develop such a model here. Instead, we hope to contribute to an understanding of reporting behavior so that such models can be developed in the future. For now, we simply provide an informal discussion of what we could expect if we assume that, in addition to *honest* individuals who incur disutility if they lie, there are also *indignant* individuals who incur a utility loss when they or others in their organization lie, but whose utility loss is smaller if liars are sanctioned. If the parameters were such that indignant individuals are willing to report overstatements, lying behavior would be further reduced or even eliminated. Introducing

---

[10] If an organization is inspected then all members who overstated, including the reporter, get sanctioned.

[11] More precisely, strategies where active individuals coordinate their voting such that all organizations end up with three players are also equilibria of the game. However, accepting all inactive individuals weakly dominates these voting strategies and requires no coordination, making it the safer option.

[12] Another approach to lying aversion is to assume there are two types of individuals: an "honest" type who never lies and an "economic" type who lies if it is beneficial to do so (see Koford and Penno, 1992; Hurkens and Kartik, 2009). Predictions using this approach are similar.

indignant individuals also allows for differences in behavior between the Random and Selection treatments. In particular, active individuals who are neither honest nor indignant might find it optimal to veto inactive individuals who reported others. In fact, we chose the parameters of the game so that two individuals who are willing to lie are better off overstating maximally in an organization of $n_g = 2$ compared to telling the truth in an organization of $n_g = 3$ (in the former they guarantee themselves 225 points and in the latter their expected payoff equals 150 points). In a similar way indignant individuals might find it optimal to veto inactive individuals who (are likely to have) overstated. This introduces an incentive for individuals to sort themselves into honest and dishonest organizations much like the anti-snitching groups described in the introduction. Depending on the frequency of each type, the risk of being left without an organization makes one's behavior relatively more costly. For instance, if indignant individuals are relatively rare then they would face a high probability of rejection by organizations in Selection treatment, which might be enough to make reporting less common or nonexistent. In this regard, it is interesting to note that indignant individuals might face rejection also by honest individuals. The intuition behind this is that even individuals who have a high cost of lying understand that they might be tempted to lie if the gain from doing so is truly large. Consequently, honest individuals might prefer not to have indignant types around.

We would like to point out that in addition to indignation at observing lies, other motivations for reporting overstatements exist. In particular, reporting could be motivated by lying averse individuals who are also envious (i.e., they dislike disadvantageous inequality, Fehr and Schmidt, 1999; Bolton and Ockenfels, 2000) or possess competitive preferences (i.e., they like advantageous earnings disparities, Frank, 1985; Fershtman and Weiss, 1998).[13] In our statistical analysis, we test whether these motivations are present by

---

[13] Yet another reason to report overstatements is a pure desire to harm others. However, evidence from other experiments suggests that this type of motivation is an unlikely explanation for reporting others. For instance, Abbink and Sadrieh (2009) find little to no evidence of destruction when subjects know their behavior is observable by others and there are opportunities to retaliate, which is the case in our game.

looking for a positive correlation between reporting others and the difference in the reporter's and liar's stated earnings.

2.3 Experimental procedures

The computerized experiment was conducted using the typical procedures of anonymity, neutrally worded instructions, and monetary incentives. In total, 192 students participated in the one-hour experiment. Each session of the experiment consisted of 24 subjects who were randomly assigned to one of two societies. We employed a between-subjects treatment design so that in each session subjects in one society played the Random treatment and those in the other played the Selection treatment. This gives us eight independent observations (societies) per treatment. Points were converted to US dollars at a rate of 150 points = $1 and subjects received a show-up fee of $15. Average earnings equaled $22.84.

Upon arrival, each subject drew a card to be randomly assigned to a seat in the laboratory. Once everyone was seated, subjects read the instructions for the experiment (a copy of the instructions is available in the online appendix) and answered a few questions to ensure their understanding of the game. When all subjects had correctly answered the questions, the computerized experiment (programmed in z-Tree, Fischbacher, 2007) started. After the game ended, subjects were confidentially paid their earnings in cash.

3. Results

In this section, we present the experimental results. We start by analyzing the subjects' overstating behavior (3.1). Thereafter, we examine their reporting behavior and the ensuing sanctions (3.2). Lastly, we discuss the subjects' voting behavior in the Selection treatment (3.3) and how it affects the profitability of overstating earnings (3.4). Note that, throughout this section, we use regression analysis to test the statistical significance of our findings. All regressions use robust standard errors clustered on societies and report *p*-values based on two-tailed tests. Moreover, unless it is otherwise noted, we include subject
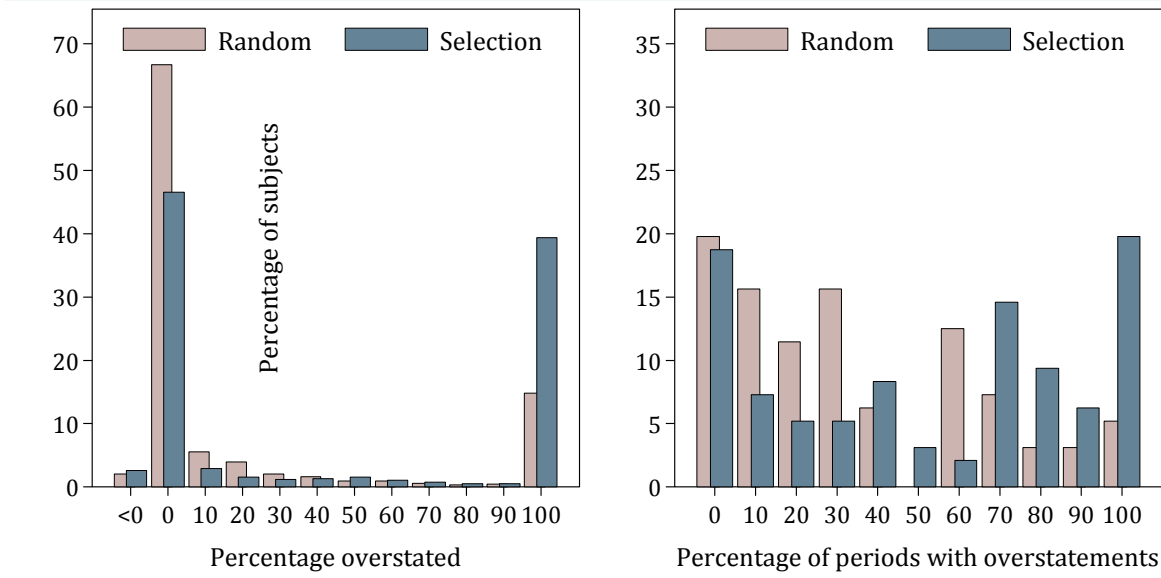
Figure 1 – Distributions of overstating behavior

Note: The left panel shows the amount overstated as a fraction of the maximum that could have been overstated. The right panel shows the percentage of periods in which subjects overstate.

random effects. This method allows us to utilize fully the panel structure of our data.[14] Table A1 in the appendix provides summary statistics for the key variables.

3.1. Overstating

To describe the degree to which subjects overstate their earnings, we look at the number of points a subject overstates in a given period as a fraction of the points she could have overstated, i.e., $(s_i - t_i)/(t_g^{max} - t_i)$. The distribution of this variable is depicted in the left panel of Figure 1. In both treatments, the modal behavior is to be honest: subjects choose $s_i = t_i$ 62 percent of the time in Random and 42 percent of the time in Selection. However, there are plenty of cases where subjects state higher earnings than the ones they received.[15] In fact, when subjects overstate their earnings they usually choose the maximum earnings they could possibly state: subjects choose $s_i = t_g^{max}$ 15 percent of the time in Random and

---

[14] In addition, we make basic treatment comparisons with non-parametric tests using society averages as units. The results of these tests are qualitatively similar and are available in the online appendix.

[15] There are cases where subjects report lower earnings than the ones they received. However, they are not very common: subjects choose $s_i < t_i$ 2 percent of the time in Random and 3 percent of the time in Selection. See Hao and Houser (2011) for more evidence of individuals lying by underreporting earnings.

39 percent of the time in Selection (see Figure 1). Thus, conditional on overstating, subjects lie as much as they can 61 percent of the time across the two treatments. In the right panel of Figure 1, we depict the distribution of the percentage of periods each subject overstates. In both treatments, we can see that the most common behavior is to overstate earnings in some but not all periods. In Random, 20 percent of subjects never overstate and 5 percent always overstate, which leaves 75 percent who sometimes overstate. In Selection, 19 percent never overstate, 20 percent always overstate, and 61 percent sometimes overstate. Hence, overstating behavior in our game is consistent with studies that report that people are willing to lie, but they do not do so maximally (e.g., Fischbacher and Heusi, 2008).

We run a series of regressions with a treatment dummy variable as the independent variable to test whether overstating behavior differs between Random and Selection. The dependent variable in the first regression is the amount overstated as a fraction of the highest feasible overstatement, i.e., $(s_i - t_i)/(t_g^{max} - t_i)$. We use Tobit estimates and censor the variable at 1 and 0. The dependent variable in the second regression is simply a dummy variable indicating whether a subject overstates, i.e., chooses $s_i > t_i$. Since the dependent variable is binary, we use logit estimates. We find that both the amount overstated and the frequency of overstatements are significantly higher in Selection ($p = 0.010$ and $p = 0.046$, respectively).

Interestingly, the difference between treatments occurs from the beginning. This is illustrated in Figure 2, which plots the means of these two variables over time. As we can see, the amount overstated and the frequency of overstatements is higher in Selection already in the first three periods.[16] Hence, even before the selection of inactive subjects takes place, subjects expect that the selection process allows them to overstate more often.

Next, we take a closer look at how subjects adjust their overstating behavior. To do so, we run two logit regressions per treatment. As the dependent variable we use a dummy variable indicating whether subject i overstates her earnings in period z. In all cases, we use subject fixed effects, period fixed effects, and robust standard errors clustered on societies.

---

[16] Using the same regressions to test for treatment differences but restricting them to the first three periods gives $p = 0.008$ for the amount overstated and $p = 0.046$ for the frequency of overstatements.
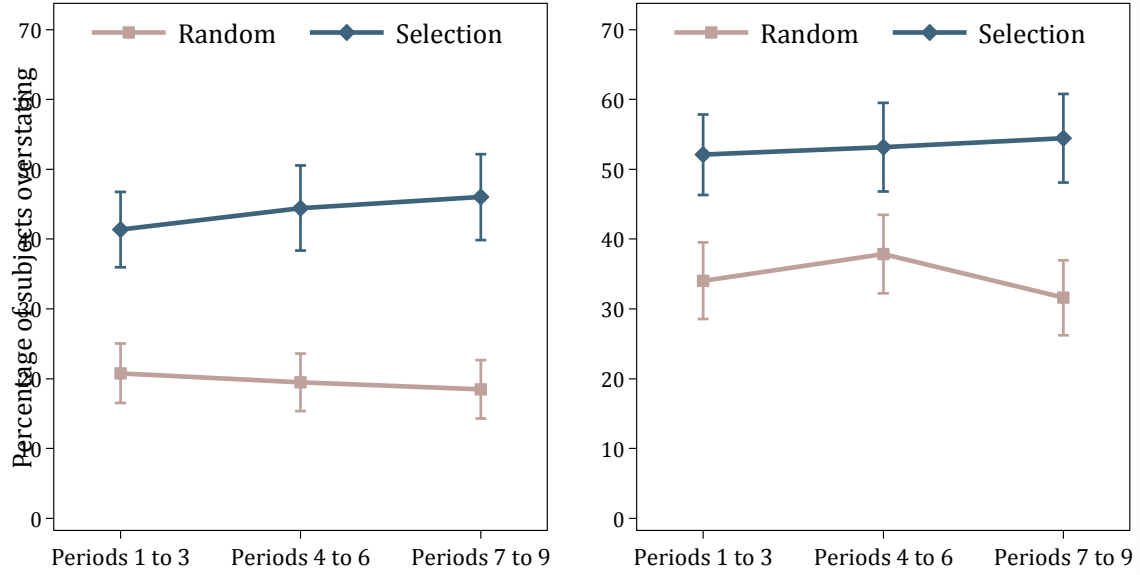
Figure 2 – Overstating behavior over time

Note: The left panel shows the mean amount overstated as a fraction of the maximum that could have been overstated. The right panel shows the mean fraction of subjects that overstate. Error bars correspond to 95 percent confidence intervals.

In the first two regressions we concentrate on the effect of i's previous actions and on factors that directly affect i's monetary payoff. Specifically, we use the following independent variables: (i) i's temptation to overstate in period z, which we define as i's monetary gain of reporting the highest possible earnings, $t_g^{max} - t_i$; (ii) a dummy variable equal to one if i overstated her earnings in period $z - 1$; (iii) a dummy variable equal to one if i was sanctioned for overstating her earnings in period $z - 1$ (because another subject chose to report); and (iv) a dummy variable equal to one if i reported another subject in period $z - 1$. In the following two regressions, we also include events observed by i that have no direct effect on her payoff. Namely, we add (v) a dummy variable equal to one if i observed at least one other subject overstate their earnings in period $z - 1$, and (vi) a dummy variable equal to one if i was not sanctioned but observed another subject being

Table 1 – Determinants of overstating

Treatment	Specification I				Specification II			
	Random		Selection		Random		Selection	
Independent variables	o.r.	s.e.	o.r.	s.e.	o.r.	s.e.	o.r.	s.e.
Temptation to overstate	1.93***	(0.22)	2.29***	(0.34)	2.00***	(0.19)	2.35***	(0.32)
Overstated in $z-1$	4.68***	(0.99)	2.66***	(0.96)	2.99***	(0.74)	1.62	(0.50)
Sanctioned in $z-1$	0.06***	(0.03)	0.37**	(0.15)	0.06***	(0.03)	0.43**	(0.17)
Reported in $z-1$	0.35***	(0.08)	0.48**	(0.14)	0.15***	(0.05)	0.25***	(0.09)
Others overstated in $z-1$					2.79***	(0.73)	2.17***	(0.43)
Others sanctioned in $z-1$					0.27**	(0.18)	0.12***	(0.06)
Period fixed effects	Yes		Yes		Yes		Yes	
Subject fixed effects	Yes		Yes		Yes		Yes	
Pseudo R^2	0.21		0.16		0.23		0.19	
# of obs./subj./societies	648/72/8		462/59/8		648/72/8		462/59/8	

Note: Logit regressions with a dummy variable indicating whether subjects overstate their earnings as the dependent variable. Clustered standard errors allow for correlation within societies. Asterisks indicate significance at the 1 percent (***), 5 percent (**), and 10 percent (*) level.

sanctioned in period $z-1$ because a subject other than i chose to report. The estimated odds ratios and respective standard errors are presented in Table 1.[17]

Consistent with models where lying behavior is sensitive to the benefits of lying (e.g., Ellingsen and Johannesson, 2004), Table 1 shows that subjects are significantly more likely to overstate their earnings as the gain from doing so increases. Specifically, an increase of one standard deviation in the temptation to overstate roughly doubles the odds of overstating. As is common in laboratory experiments, subjects tend to repeat actions that resulted in a high payoff and avoid actions that gave them a low payoff. Specifically, subjects who overstated and were not sanctioned have increased odds of overstating whereas subjects who overstated and were sanctioned have decreased odds of overstating.

With the second specification, we see that subjects are learning from observing the actions of others. That is, observing other subjects overstate increases the odds of overstating and observing other subjects be sanctioned for overstating decreases the odds

---

[17] To facilitate the interpretation of the coefficient of (i), we normalize this variable to have a mean of zero and a standard deviation of one. Moreover, since we use subject fixed effects and there are subjects who never or always overstate, the coefficients are estimated using less than 96 subjects and 864 observations.

of overstating.[18] Finally, note that by and large the dynamics of overstating are similar in both treatments.[19] Hence, the effect of selecting who enters the organization does not seem to affect how subjects react to each other's actions.

We summarize these findings as our first result.

Result 1 (Overstating earnings): *Subjects regularly overstate their earnings, albeit, most do not do so in every period. The possibility of selecting who enters their organization facilitates overstating even before selection takes place. Overstating is more likely if the gain from lying is large and subjects observe or experience unpunished lying in the past.*

3.2. Reporting behavior

In both treatments, the overall rate at which subjects report others is not high, which results in few instances in which subjects are sanctioned. In Random, subjects report others 17 percent of the time, resulting in 19 percent of choices being sanctioned, while in Selection they report others 11 percent of the time resulting in 13 percent of choices being sanctioned. Logit regressions confirm that reporting others for overstating and the ensuing sanctions are significantly more common in Random compared to Selection ($p < 0.032$). However, it is more informative to look at reporting behavior conditional on there being a reason to report.

Figure 3 plots the fraction of subjects who report others given that at least one other subject overstated (left panel) and the fraction of subjects who are sanctioned conditional on having had overstated (right panel). Subjects report others when they see them overstate 32 percent of the time in Random and 17 percent of the time in Selection (significantly different with a logit regression, $p = 0.028$). We also see an important treatment difference in the fraction of subjects who never report: 28 percent in Random

---

[18] We also observe that subjects are less likely to overstate if they reported someone in the previous period. It is not clear why this might be the case, but it could be due to subjects expecting retaliation from subjects who got sanctioned. For evidence of this type of retaliation see Nikiforakis (2008) and Hopfensitz and Reuben (2009).

[19] If we run one regression where we interact all variables with a treatment dummy variable, we obtain a significant interaction term only for 'sanctioned in $z - 1$' ($p = 0.001$, for all others $p > 0.150$).
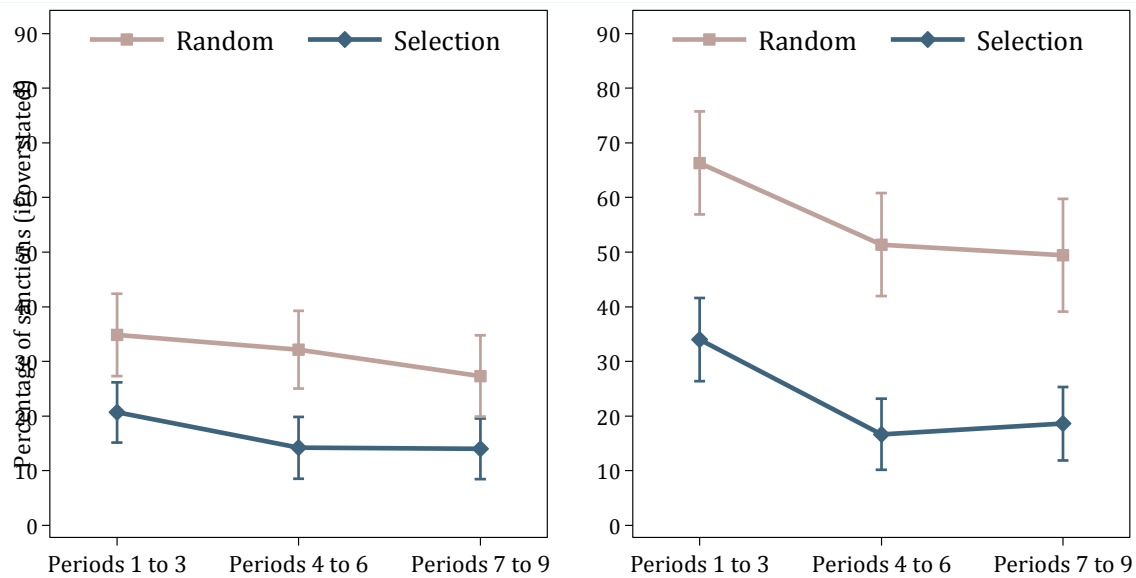
Figure 3 – Reporting behavior and sanctions over time

Note: The left panel shows the mean fraction of subjects who report others for overstating their earnings conditional on having the opportunity to report. The right panel shows the mean fraction of subjects who are sanctioned conditional on having had overstated. Error bars correspond to 95 percent confidence intervals.

and 53 percent, a majority, in Selection. The fact that the reporting rate varies between treatments suggests that the willingness to report is sensitive to the incentives introduced by the selection process. We see an even larger difference between treatments in the fraction of overstatements that are sanctioned: a sizable 56 percent of overstatements are sanctioned in Random but only 24 percent in Selection (significantly different with a logit regression, $p = 0.006$).

To evaluate the determinants of reporting we run two logit regressions per treatment. The dependent variable is a dummy variable indicating whether subject i reports other subjects for overstating their earnings in period z. In all cases, we use subject fixed effects, period fixed effects, and robust standard errors clustered on societies. In the first two regressions we concentrate on the effect of i's actions and others' actions in period z. We use the following independent variables: (i) the amount by which others overstate their earnings in period z, $\sum_{j \neq i}(s_j - t_j)$, which measures the extent to which others lie; (ii) the difference in stated earnings between i and others in period z, $\sum_{j \neq i}(s_j - s_i)$, which captures incentives to report that are unrelated to lying but might be important such as inequity

Table 2 – Determinants of reporting overstatements

	Specification I				Specification II			
Treatment	Random		Selection		Random		Selection	
Independent variables	o.r.	s.e.	o.r.	s.e.	o.r.	s.e.	o.r.	s.e.
Deviation from true earnings	9.53***	(4.79)	10.02***	(5.59)	9.18***	(4.49)	10.17***	(5.31)
Difference in earnings	0.93	(0.11)	1.32**	(0.18)	0.94	(0.11)	1.39**	(0.19)
Overstated	0.05***	(0.03)	0.05***	(0.04)	0.05***	(0.03)	0.06***	(0.06)
Overstated × dev. true earnings	0.18***	(0.11)	1.16	(0.63)	0.18***	(0.11)	1.11	(0.60)
Overstated × diff. earnings	1.54	(0.84)	2.95	(2.61)	1.60	(0.86)	2.44	(1.91)
Sanctioned in $z-1$					0.57	(0.27)	3.01	(2.96)
Reported in $z-1$					0.76	(0.42)	0.89	(0.48)
Others reported in $z-1$					0.98	(0.79)	2.36	(2.26)
Period fixed effects	Yes		Yes		Yes		Yes	
Subject fixed effects	Yes		Yes		Yes		Yes	
Pseudo R^2	0.33		0.53		0.33		0.54	
# of obs./subj./societies	585/65/8		318/41/8		585/65/8		318/41/8	

Note: Logit regressions with a dummy variable indicating whether subjects report others as the dependent variable. Clustered standard errors allow for correlation within societies. Asterisks indicate significance at the 1 percent (***), 5 percent (**), and 10 percent (*) level.

aversion and competitive preferences; (iii) a dummy variable equal to one if i overstated her earnings in period z; (iv) an interaction term between (i) and (iii); and (v) an interaction term between (ii) and (iii). In the other two regressions, we also control for events in the previous period. To this end, we use (vi) a dummy variable equal to one if i was sanctioned for overstating her earnings in period $z-1$; (vii) a dummy variable equal to one if i reported others in period $z-1$; and (viii) a dummy variable equal to one if i observed at least one other subject being sanctioned in period $z-1$. The estimated odds ratios and standard errors are presented in Table 2.[20]

Table 2 shows that observing others overstate is a powerful determinant of deciding to report. The coefficient for 'deviation from true earnings' is both large and statistically significant in both treatments. This is consistent with the assumption that some individuals feel indignation toward those who lie (and more so the larger the lie is) and are willing to

---

[20] As before, we normalize the continuous variables (i) and (ii) to have a mean of zero and a standard deviation of one. Moreover, since we use subject fixed effects and some subjects never or always report, the coefficients are estimated using less than 96 subjects and 864 observations.

sanction them for doing so. A second powerful determinant of the decision to report is having had overstated. Subjects who overstated are significantly less likely to report others, which is not very surprising as by reporting they also sanction themselves. By contrast, the effect of differences in earnings is significant only in Selection and is markedly smaller compared to the previous two. Lastly, note that the interaction term for having overstated times the deviation from true earnings is *not* significant in the Selection treatment, which shows there is some willingness to report others even when doing so implies incurring a sanction. We should note, however, that overstating and then reporting is a rare occurrence, so these latter results are based on few observations (11 out of 585 in Random and 10 out of 318 in Selection).

It is worth mentioning that the coefficients of the two most impactful variables (i.e., the deviation from true earnings and having had reported) are not significantly different across treatments.[21] This shows that the lower reporting rate in Selection is not due to honest subjects being less willing to report, but instead it is due to fewer instances where honest subjects see others overstate. In the next subsection, we analyze how the selection process facilitates this effect by separating subjects who report from those who overstate.

The following result summarizes the findings of this subsection.

Result 2 (Reporting overstatements): *In the absence of selection, enough subjects report others so that subjects who overstate are sanctioned relatively often. If it is possible for organizations to select their members, reporting is much less frequent and overstatements are sanctioned less often. If they are not lying themselves, subjects are more likely to report others the more egregious the lying in their organization is.*

3.3. Organization member selection

In this section, we restrict our analysis to the Selection treatment in order to assess how people decide whom to accept into their organization as well as how the composition of organizations changes over time. Figure 4 presents the fraction of active subjects who veto

---

[21] If we run one regression where we interact all variables with a treatment dummy variable, we obtain a significant interaction term only for 'sanctioned in $z - 1$' ($p = 0.001$, for all others $p > 0.150$).
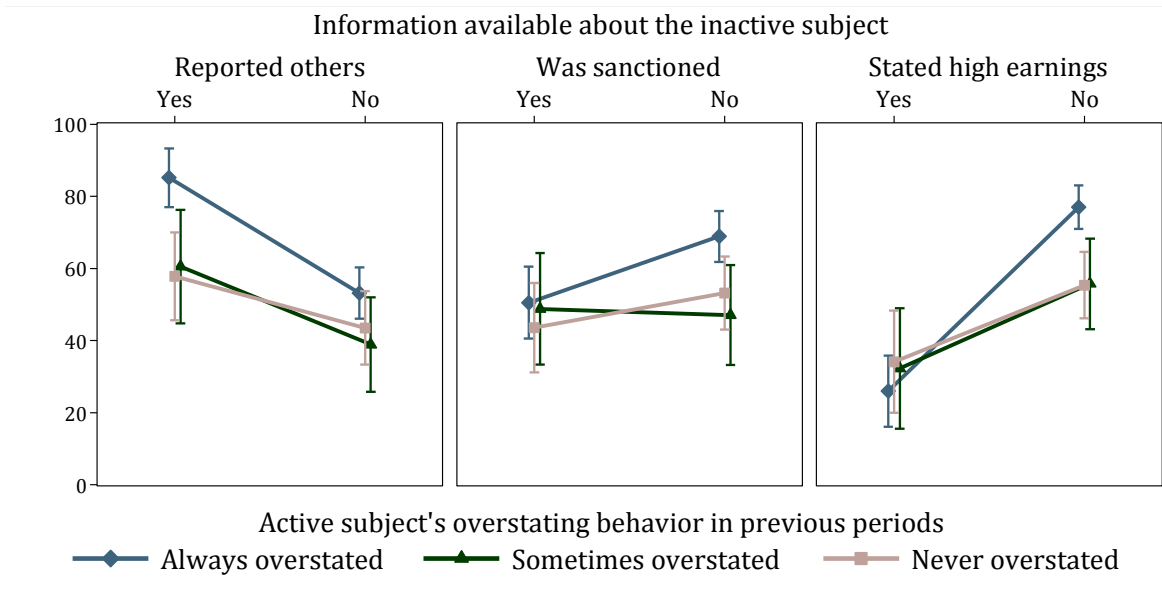
Information available about the inactive subject

Figure 4 – Vetoing depending on the subject's behavior

Note: From left to right, the panels show the fraction of active subjects who veto an inactive subject depending on whether the inactive subject: reported others, was sanctioned for overstating, or stated earnings above 233 points. Within panels, voting behavior is separated depending on how often the active subject overstated in the previous three periods. Error bars correspond to 95 percent confidence intervals.

an inactive subject depending on their behavior in the last three periods. Across panels, it displays the effect of the inactive subjects' behavior that is available to active subjects when they cast their vote. From left to right, the figure shows whether inactive subjects reported others, whether they were sanctioned for overstating, and whether their stated earnings were suspiciously high (earnings that occur with less than 5 percent probability by chance alone; 35 percent of inactive candidates fall in this category). Within panels, the figure plots separately the voting behavior of active subjects depending on how often they overstated in the previous three periods.

First, let us consider whether active subjects accept or reject inactive subjects who reported others for lying. As we can see, inactive subjects who never reported others fare much better than those who did (the former are vetoed by 48 percent of active subjects while the latter are vetoed by 70 percent). This pattern is expected for active subjects who always or sometimes overstated. What is more interesting is that the same pattern appears for active subjects who never overstated. However, as we pointed out in Section 2.2., this

Table 3 – Determinants of vetoing

	Specification I		Specification II	
Independent variables	o.r.	s.e.	o.r.	s.e.
j was sanctioned	1.19	(0.26)		
j's stated earnings	0.51***	(0.13)		
j reported others	2.47***	(0.40)		
i never overstated × *j* was sanctioned			1.25	(0.19)
i never overstated × *j*'s stated earnings			0.61**	(0.14)
i never overstated × *j* reported others			2.15***	(0.50)
i sometimes overstated × *j* was sanctioned			1.37	(0.48)
i sometimes overstated × *j*'s stated earnings			0.61**	(0.15)
i sometimes overstated × *j* reported others			3.11**	(1.72)
i always overstated × *j* was sanctioned			0.99	(0.48)
i always overstated × *j*'s stated earnings			0.38***	(0.13)
i always overstated × *j* reported others			2.82***	(0.73)
Period fixed effects	Yes		Yes	
Subject fixed effects	Yes		Yes	
Pseudo R^2	0.17		0.19	
# of obs./subj./societies	464/68/8		464/68/8	

Note: Logit regressions with a dummy variable indicating whether active subject *i* vetoes inactive subject *j* as the dependent variable. Clustered standard errors allow for correlation within societies. Asterisks indicate significance at the 1 percent (***), 5 percent (**), and 10 percent (*) level.

behavior is consistent with a model in which subjects who are generally averse to their own (but not others') lies understand that they might be tempted to lie in some future instances, and hence they prefer not to have individuals who report others in their organization. Next, we turn to inactive subjects who are known for, or can be suspected of, lying. We see that overstating actually facilitates rejoining organizations. Inactive subjects who display suspiciously high earnings or who were sanctioned for overstating are vetoed by a smaller fraction of active subjects (30 and 48 percent, respectively) than subjects with lower earnings or without sanctions (67 and 61 percent, respectively). Moreover, as before, this pattern is similar for active subjects who never, sometimes, and always lied.

We test whether these differences are statistically significant with two logit regressions. In both cases, the dependent variable is a dummy variable equal to one if active subject *i* vetoes inactive subject *j*. In both cases, we use subject fixed effects, period fixed effects, and
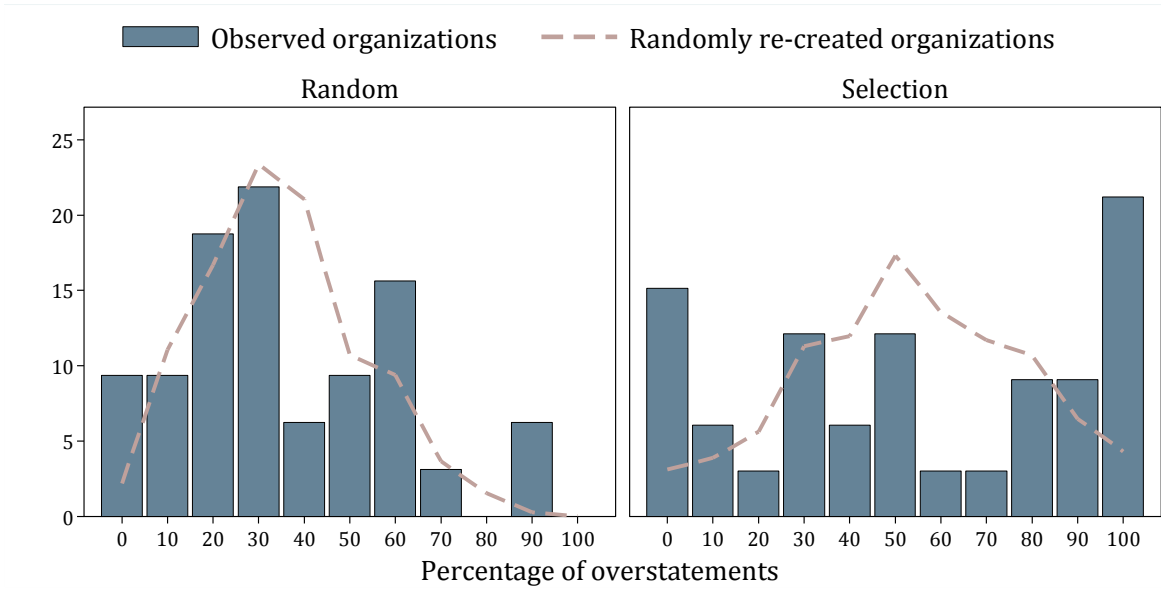
Figure 5 – Distribution of organizations according to the percentage of overstatements

Note: Bars show the fraction of organizations depending on the fraction of times subjects overstate in periods 4 to 9. Lines show the same distribution assuming that the subjects' propensity to overstate is unaffected by the organization they are in.

robust standard errors clustered on societies. In the first regression, the independent variables are: (i) a dummy variable equal to one if j was sanctioned for overstating; (ii) j's mean stated earnings; and (iii) a dummy variable equal to one if j reported others for overstating. In the second regression, each of these variables is interacted with three dummy variables: one equal to one if i never overstated, one equal to one if i sometimes overstated, and one equal to one if i always overstated. The results are displayed in Table 3.

We can see that the odds of vetoing significantly increase when the inactive subject is known to have reported others and the effect is significant irrespective of the overstating behavior of the active subjects. We also see a significant effect for the inactive subjects' stated earnings: the higher the stated earnings, and hence the probability of having had overstated, the higher the chance of acceptance. Finally, once we control for other variables we do not see that being sanctioned significantly influences vetoing.

We next evaluate the impact of the selection process on organizational composition. To this end, we look at the fraction of times subjects overstate in each organization. In the first three periods, before selection takes place, in both Random and Selection the vast majority

of organizations display a combination of truthful and untruthful statements. The fraction of fully honest organizations where no subject overstates in any period is 9 percent in Random and 6 percent in Selection. Similarly, the fraction of fully dishonest organizations where all subjects overstate in all periods is 3 percent in Random and 9 percent in Selection. After the third period, however, we see a noticeable change in the distribution of organizations in Selection. This is illustrated in Figure 5. The figure plots the fraction of organizations depending on the fraction of times subjects overstate in periods 4 to 9. As we can see, while in Random there are no fully dishonest organizations, in Selection the modal organization is in fact the fully dishonest one. By contrast, the fraction of organizations that are fully honest is still not that far apart: 9 percent in Random and 15 percent in Selection.

Figure 5 presents suggestive evidence that the selection process in the experiment is sufficient to create a substantial number of fully dishonest organizations. However, it is conceivable that the difference between the two distributions is simply the result of individual subjects having a higher propensity to overstate in Selection, and it is not due to selection and peer effects. To assess the importance of these effects, we simulate how the distributions in Figure 5 would appear if the subjects' propensity to overstate is independent of the organization they are in (plotted as dashed lines).[22] Compared to the simulated distributions, there is a larger fraction of organizations that are completely honest in both treatments. This suggests that observing each other's statements, which is the case in both treatments, is enough to produce some conformity in the direction of full honesty. In addition, in Selection we see an even larger difference for the fraction of organizations that are fully dishonest. More specifically, while in Selection a fifth of the organizations are fully dishonest, the probability of observing a fully dishonest organization by randomly regrouping individual overstatement rates is only 4 percent. This is compelling

---

[22] Specifically, we simulate the following scenario 50,000 times: first, we calculate the overstatement rate of each subject (in periods 4 to 9); second, we randomly reassign subjects to organizations (within each treatment); and third, we compute the fraction of overstatements that would occur in each organization if the reassignment does not change the subjects' overstatement rate.

evidence that granting subjects the ability to reject others from joining their organization is enough to create organizations with strong norms against reporting others.

We summarize the findings from this subsection in the following result.

Result 3 (Acceptance into organizations): *Inactive subjects who reported others face more resistance when rejoining organizations as they are rejected both by subjects who overstate and those who do not. By contrast, inactive subjects with high stated earnings, which suggest that they overstated, are generally welcomed. This selection process leads to the formation of a sizable number of highly dishonest organizations with strong norms against reporting others.*

3.4. Final payoffs

The subjects' final payoff is very similar in both treatments. On average, subjects make 129 points per period in Random and 133 points in Selection (not significantly different with a GLS regression, $p = 0.791$). Hence, the additional points subjects make in Selection due to more overstating and lower sanctions are canceled out by a smaller t_g^{max} in some instances and the low earnings of inactive subjects who do not manage to rejoin an organization (18 percent of subjects remain inactive after the third period). What is more interesting is to evaluate the effect of overstating and reporting on the subjects' final payoff.

To see whether overstating pays off in each treatment we calculate each subject's payoff over the nine periods and the fraction of times each subject overstated and reported others. For each treatment, the left panel in Figure 6 plots the relationship between final payoffs and the fraction of overstatements (divided into terciles). In Random, we see that the tercile of subjects who overstate the most have *lower* payoffs than subjects who overstate less. By contrast, in Selection the converse is true. Subjects who overstate the most have *higher* payoffs than subjects who overstate less. On the right panel of Figure 6, we plot the relationship between final payoffs and the fraction of times subjects reported others (also divided into terciles). In this case, we see that reporting others has no perceivable effect on payoffs in Random but it has a strong negative effect in Selection. To test whether these relations are statistically significant, we regress the subjects' payoff on the fraction of times each subject overstated or the fraction of times each subject reported others. We run the
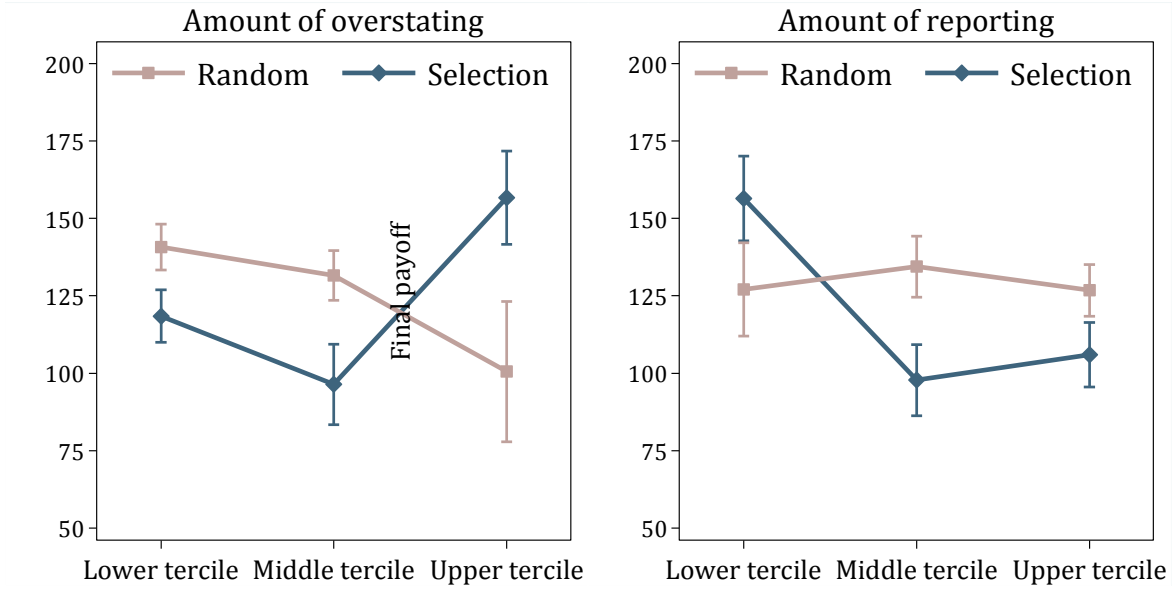
Figure 6 – Final payoff depending on the amount of overstating and reporting

Note: Error bars correspond to one standard error.

GLS regressions using society fixed effects. In Random we obtain a significantly negative coefficient for the fraction of overstatements ($\beta = -67$ points, $p = 0.012$) and a positive but not significant coefficient for the fraction of reports ($\beta = 28$ points, $p = 0.519$). By contrast, in Selection the coefficient for fraction of overstatements is significantly positive ($\beta = 59$ points, $p = 0.010$) and the one for the fraction of reports is negative and weakly significant ($\beta = -76$ points, $p = 0.081$).

The findings from this subsection give us our last result.

Result 4 (Final payoffs): *In the absence of selection, dishonest subjects have lower payoffs than honest subjects. However, if it is possible for organizations to select their members, it is profitable to overstate one's earnings because dishonest subjects obtain higher payoffs than honest subjects do. In this situation, the lowest payoff is obtained by subjects who report others for overstating.*

4. Conclusions

In this paper, we study whether individuals are willing to sanction other people if they lie for personal gain by reporting them to a central authority. We find that, in randomly

assigned groups, enough people are willing to report lying to make lying unprofitable. We also investigate how the frequency of lying and reporting lies is affected when groups can select their members. Our results indicate that this type of selection is enough to increase the frequency of lying and decrease the amount of reporting by facilitating the formation of a significant number of dishonest groups where lying is prevalent and reporting is nonexistent.

The fact that some individuals report others for lying even when lies are not directed toward those doing the reporting and cause no obvious harm to third parties suggests that lying *per se* is considered by some as normatively undesirable behavior that ought to be punished. This fact calls for the modeling of lying as more than an individual cost and more as a social norm (e.g., López-Pérez, 2012). In such a model, whether individuals lie (or report others) would depend on not only the monetary benefit (or cost) of their action but also on the expectations of others regarding their behavior (Bicchieri, 2006; Charness and Dufwenberg, 2006). This approach to lying aversion helps explain why lying is accepted in some contexts, such as when playing poker, but not in others.

As in most experimental studies, our study abstracts away from many elements present in the field in order to cleanly identify specific effects or motivations. Hence, some caution ought to be exerted when thinking on the external validity of our findings. The presence of individuals who are intrinsically motivated to report dishonest behavior is encouraging and provides one answer to the question posed in the whistleblowing literature of why some individuals are willing to blow the whistle even though the consequences of doing so are dire. However, this does not mean that extrinsic motivations are not important or even the main factor behind many whistleblowing cases. Indeed, the whistleblowing game can be easily modified to evaluate the relative impact and interaction of intrinsic and extrinsic incentives to report on others' dishonesty. This type of research can potentially inform policies such as the granting of immunity to whistleblowers or rewarding them through *qui tam* statutes.

Our study also suggests that relying on reports from individuals within organizations to discourage dishonest behavior is a lost cause if others can identify the reporters and

subsequently avoid them. In fact, selection effects might be a lot stronger in the field as selection in our experiment is limited to rejecting individuals from joining organizations. An interesting avenue of research would be to observe the effect of other possible selection mechanisms. For instance, slight modifications to our game can be used to evaluate the effect of reporting others on the likelihood of being unjustly "fired". Alternatively, if individuals are given the option to leave then honest individuals might prefer to avoid dishonesty by leaving as opposed to reporting the organization.

Finally, our work also highlights that reporting or sanctioning others for being dishonest is not behavior that is always well received or sought after even by individuals who act honestly. This is an important finding because it implies that reporting dishonest actions is very costly as reporters can be ostracized even if most organizations are in fact honest. This helps explain the dismal career prospects of employees who blow the whistle (Dyck et al., 2010) and calls for caution when it comes to policies or actions that reveal the identity of whistleblowers.

References

Abbink, K., Sadrieh, A., 2009. The pleasure of being nasty. *Economics Letters* 105: 306-308.

Bicchieri, C., 2006. *The grammar of society: The nature and dynamics of social norms*. Cambridge University Press, New York.

Bolton, G.E., Ockenfels, A., 2000. ERC: A theory of equity, reciprocity, and competition. *American Economic Review* 90: 166-193.

Bolton, P., Freixas, X., Shapiro, J., 2012. The Credit Ratings Game. *The Journal of Finance* 67: 85-112.

Bowen, R.M., Call, A.C., Rajgopal, S., 2010. Whistle-blowing: Target firm characteristics and economic consequences. *The Accounting Review* 85: 1239-1271.

Brandts, J., Charness, G., 2003. Truth or consequences: An experiment. *Management Science* 49: 116-130.

Brown, E., 2007. *Snitch: Informants, Cooperators, and the Corruption of Justice*. New York: Public Affairs Press.

Cai, H., Wang, J. T.-Y., 2006. Overcommunication in strategic information transmission games. *Games and Economic Behavior* 56: 7-36.

Charness, G., Dufwenberg, M., 2006. Promises and partnership. *Econometrica* 74: 1579-1601.

Charness, G., Dufwenberg, M., 2010. Bare promises: An experiment. *Economics Letters* 107: 281-283.

Croson, R., Boles, T., Murnighan, K., 2003. Cheap talk in bargaining experiments: Lying and threats in ultimatum games. *Journal of Economic Behavior & Organization* 51: 143-159.

Demichelis, S., Weibull, J.W., 2008. Language, meaning, and games: A model of communication, coordination, and evolution. *American Economic Review* 98: 1292-1311.

Dyck, A., Morse, A., Zingales, L., 2010. Who blows the whistle on corporate fraud? *The Journal of Finance* 65: 2213-2253.

Eisenkopf, G., Gurtoviy, R., Utikal, V., 2011. Size matters—When it comes to lies. Working paper, University of Konstanz.

Ellingsen, T., Johannesson, M., 2004. Promises, threats, and fairness. *The Economic Journal* 114: 397-420.

Erat, S., Gneezy, U., in press. White lies. *Management Science* 58: 723-733.

Falk, A., Fehr, E., Fischbacher, U., 2008. Testing theories of fairness—Intentions matter. *Games and Economic Behavior* 62: 287-303.

Fehr, E., Schmidt, K., 1999. A theory of fairness, competition, and cooperation. *The Quarterly Journal of Economics* 114: 817-868.

Fershtman, C., Weiss, Y., 1998. Social status and economic performance: A survey. *European Economic Review* 42: 801-820.

Fischbacher, U., 2007. z-Tree: Zurich toolbox for readymade economic experiments. *Experimental Economics* 10: 171-178.

Fischbacher, U., Heusi, F., 2008. Lies in disguise. An experimental study on cheating. TWI Research Paper Series, Universität Konstanz.

Frank, R.H., 1985. Choosing the right pond: Human behavior and the quest for status. Oxford University Press, New York.

Gibson, R., Tanner, C., Wagner, A., in press. Preferences for truthfulness: Heterogeneity among and within individuals. *American Economic Review*.

Gneezy, U., 2005. Deception: The role of consequences. *American Economic Review* 95: 384-394.

Hao, Li, Houser, D., 2011. Honest Lies. Working Paper, George Mason University.

Hopfensitz, A., Reuben, E., 2009. The importance of emotions for the effectiveness of social punishment. *The Economic Journal* 119: 1534-1559.

Hurkens, S., Kartik, N., 2009. Would I lie to you? On social preferences and lying aversion. *Experimental Economics* 12: 180-192.

Kahn, J., 2007. The story of a snitch. *The Atlantic Monthly* 80: 86-88.

Kartik, N., 2009. Strategic communication with lying costs. *Review of Economic Studies* 76: 1359-1395.

Koford, K., Penno, M., 1992. Accounting, principal-agent theory, and self-interested behavior. In: Bowie, N.E., Freeman, E.R. (Eds.), *Ethics and agency theory: An introduction*. Oxford University Press, New York, 127-142.

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

López-Pérez, R., Spiegelman, E., 2012. Do economists lie more? Working paper, Universidad Autónoma de Madrid.

Lundquist, T., Ellingsen, T., Gribbe, E., Johannesson, M., 2009. The aversion to lying. *Journal of Economic Behavior & Organization* 70: 81-92.

Nikiforakis, N., 2008. Punishment and counter-punishment in public good games: Can we really govern ourselves. *Journal of Public Economics* 92: 91-112.

Rode, J., 2010. Truth and trust in communication: Experiments on the effect of a competitive context. *Games and Economic Behavior* 68: 325-338.

Sanchez-Pages, S., Vorsatz, M., 2007. An experimental study of truth-telling in a sender-receiver game. *Games and Economic Behavior* 61: 86-112.

Sánchez-Pagés, S., Vorsatz, M., 2009. Enjoy the silence: An experiment on truth-telling. *Experimental Economics* 12: 220-241.

Schmidt, M., 2005. "Whistle blowing" regulation and accounting standards enforcement in Germany and Europe—An economic perspective. *International Review of Law and Economics* 25: 143-168.

Sutter, M., 2009. Deception through telling the truth?! Experimental evidence from individuals and teams. *The Economic Journal* 119: 47-60.

Vanberg, C., 2008. Why do people keep their promises? An experimental test of two explanations. *Econometrica* 76: 1467-1480.

Whitman, J.L., Davis, R.C., 2007. *Snitches get stitches: Youth, gangs and witness intimidation in Massachusetts*. National Center for Victims of Crime, Washington, D.C.

Appendix – Descriptive and nonparametric statistics

For each treatment, Table A1 contains the mean and standard deviation of the following key variables: the stated earnings (s_i); the true earnings (t_i); the amount overstated $(s_i - t_i)$; the amount overstated as a percentage of the maximum possible overstatement, i.e., $(s_i - t_i)/(t_g^{max} - t_i)$; the fraction of subjects overstating; the fraction of subjects reporting others; the fraction of organizations inspected; the fraction of subjects sanctioned in those inspections; the fraction of inactive subjects, which is necessarily zero except for the later periods of the Selection treatment; and the final payoff.

Table A1 – Descriptive statistics

Treatment	Random				Selection			
Periods	1-3	4-6	7-9	All	1-3	4-6	7-9	All
Stated earnings	186.3	186.0	170.3	180.9	213.4	196.7	192.7	201.8
	(92.5)	(89.3)	(92.5)	(91.6)	(90.9)	(85.8)	(90.8)	(89.6)
True earnings	152.1	157.1	144.4	151.2	144.0	137.5	133.3	138.6
	(90.9)	(86.4)	(84.4)	(87.3)	(84.5)	(81.1)	(81.0)	(82.4)
Points overstated	34.2	28.9	25.8	29.7	69.4	59.2	59.5	63.1
	(77.0)	(66.3)	(63.9)	(69.3)	(98.2)	(85.9)	(84.8)	(90.5)
Fraction of points overstated (in percent)	20.8	19.5	18.5	19.6	41.3	44.4	46.0	43.8
	(36.8)	(35.6)	(36.3)	(36.2)	(46.6)	(48.1)	(48.3)	(47.6)
Fraction of subjects overstating (in percent)	34.0	37.8	31.6	34.5	52.1	53.2	54.4	53.1
	(47.5)	(48.6)	(46.6)	(47.6)	(50.0)	(50.0)	(49.9)	(49.9)
Fraction of subjects reporting (in percent)	18.8	18.8	13.2	16.9	14.9	8.9	8.9	11.1
	(39.1)	(39.1)	(33.9)	(37.5)	(35.7)	(28.5)	(28.5)	(31.5)
Fraction of inspections (in percent)	46.9	41.7	33.3	40.6	36.5	17.3	20.3	25.5
	(50.0)	(49.4)	(47.2)	(49.1)	(48.2)	(37.9)	(40.3)	(43.6)
Fraction of subjects sanctioned (in percent)	22.6	19.4	15.6	19.2	17.7	8.9	10.1	12.6
	(41.9)	(39.6)	(36.4)	(39.4)	(38.2)	(28.5)	(30.2)	(33.2)
Fraction of inactive subjects	0.0	0.0	0.0	0.0	0.0	17.7	17.7	11.8
	–	–	–	–	–	(38.2)	(38.2)	(32.3)
Final payoff	115.7	137.5	133.9	129.0	134.1	135.1	128.4	132.5
	(179.1)	(147.0)	(130.8)	(153.7)	(201.0)	(149.6)	(148.1)	(167.9)

Table A2 displays the exact *p*-values from two-sided Mann-Whitney U tests that evaluate whether key variables significantly differ between the two treatments. In all cases, tests are performed with society means as observations. The first four rows provide support for Result 1. Namely, the amount overstated as a fraction of the highest feasible overstatement is significantly higher in Selection than in Random, as is the fraction of subjects overstating. Moreover, the difference occurs already in the first three periods. The next four rows provide support for Result 2. Specifically, the overall fraction of subjects reporting others and of subjects being sanctioned is significantly smaller in Selection than in Random. Moreover, these significant differences also exist when we limit ourselves to instances where subjects have an opportunity to report (because others overstate) or be sanctioned (because they overstate). Finally, the last row provides support for the statement in section 3.4 that states that final payoffs do not differ statistically between treatments.

Table A2 – Treatment differences with Mann-Whitney U tests

Variable	*p*-value
Fraction of points overstated	0.010
Fraction of subjects overstating	0.038
Fraction of points overstated (periods 1 to 3)	0.021
Fraction of subjects overstating (periods 1 to 3)	0.083
Fraction of subjects reporting	0.038
Fraction of subjects reporting given others overstated	0.038
Fraction of subjects sanctioned	0.050
Fraction of subjects sanctioned given they overstated	0.010
Final payoff	0.959

Appendix – Instructions (Not for publication)

Here we provide the instructions for the Selection treatment (instructions for the Random treatment are available on request).

Instructions

You are now taking part in an economic study. You can earn money depending on your decisions and the decisions of other participants. How you earn money is described in these instructions. Please read them carefully.

During the study, you are not allowed to communicate with other participants. If you have a question, raise your hand. One of us will come to answer your question. During the study, your earnings will be calculated in points. At the end of the study, points will be converted to dollars at a rate of 150 points = 1 dollar.

In the study, you will be randomly assigned to a *market* of 12 participants (i.e. you and 11 other participants). The game is divided into *9 periods*. Your total earnings for the game will be the sum of your earnings over all periods. You will play with the same 12 participants during all the periods. Furthermore, within each market, you will be randomly divided into *groups*. Initially all groups have 3 participants (i.e. you and 2 other participants).

Your decisions in each period

You will observe a button called 'Get endowment' on the computer screen. Every time you click the button, it generates a random number from a *uniform* distribution. If you are part of a group of three participants, it will generate a number between 0 and 300 (each equally likely). If you are part of a group of two participants it will generate a number between 0 and 225 (each equally likely).

In every period, you will receive an endowment. Specifically, your endowment equals the random number you generate with your *first* click of the `Get endowment' button. Of course, you may click the button more than once to ensure the randomization algorithm is working properly. However, your endowment always corresponds to your first click.

Before making your decisions, you will be informed of the endowments of other participants in your group. You will not be informed of the endowment of participants in other groups. You make two decisions in each period.

For your first decision, you are asked to *state* what your endowment is. Your earnings in each period depend on the number you state. If you are part of a group of three participants, you can state a number between 0 and 300. If you are part of a group of two participants, you can state a number between 0 and 225. Note that the computer will not check whether the number you state is the same as the endowment you received. Before you make your second decision, you will be informed of numbers stated by all participants in your group.

Your second decision is to indicate whether you want your group to be *reviewed* or *not reviewed*. Your group will be reviewed if at least one of the participants in the group indicates they want a review. How this affects your and the earnings of others is explained below.

Earnings

Your earnings, in points, in each period depend on the number you state and on whether your group is reviewed or not. If your group is *not reviewed* then your earnings are equal to the number you state. If your group is *reviewed* then all the members of the group who stated a number that is bigger than their endowment will have their earnings reduced by *three times the difference between their stated number and their endowment*. There is no effect on the earnings of participants whose stated number is smaller than or equal to their endowment.

Example 1

If your endowment equals 200 points, you state 300 points, and your group is not reviewed then your earnings in this period equal 300 points.

Example 2:

If your endowment equals 200 points, you state 300 points, and your group is reviewed then your earnings in this period equal $300 - 3 \times (300 - 200) = 0$ points.

Example 3:

If your endowment equals 200 points, you state 200 points, and your group is reviewed then your earnings in this period equal 200 points.

Switching Groups

In addition, at the end of every *three periods* (i.e. after period three and six), in all groups of three participants, one participant will be randomly selected and separated from their group (groups with two participants do not become smaller). Participants who are separated from their group can be accepted into their old group or into a new group. The procedure for acceptance is described below.

First, all participants in the market receive information regarding the separated participants. Specifically, they will be able to see for each separated participant: (i) his/her average stated number over the last three periods; (ii) whether he/she requested at least one review in the last three periods; and (iii) whether he/she had his/her earnings reduced due to a review in the last three periods.

Second, participants in all groups indicate for each separated participant whether they would accept him/her in their group of not. If everyone in a group accepts a participant, he/she joins that group in the next period. If a participant is accepted by more than one group then he/she will be randomly assigned to one of the accepting groups (with equal probability). Similarly, if a group accepts more than one participant then it will be randomly assigned to one of the separated participants (with equal probability).

Note that participants who are not accepted to any group get earnings equal to 0 points for the next three periods. After three periods, they have another chance of being accepted into a group. Groups that do not accept any participants will remain as a group of two for the next three periods.

Click on 'Done' once you have finished reading the instructions.