Title: Selection, Tournaments, and Dishonesty

Author: Marco Faravelli Lana Friesen Lata Gangadharan

PII: S0167-2681(14)00279-0
DOI: http://dx.doi.org/doi:10.1016/j.jebo.2014.10.019
Reference: JEBO 3454

To appear in: *Journal of Economic Behavior & Organization*

Received date: 26-6-2014
Revised date: 29-10-2014
Accepted date: 30-10-2014

Please cite this article as: Faravelli, M., Friesen, L., Gangadharan, L.,Selection, Tournaments, and Dishonesty, *Journal of Economic Behavior and Organization* (2014), http://dx.doi.org/10.1016/j.jebo.2014.10.019

Highlights

We examine the relationship between dishonesty and selection into competition.

Do honest or dishonest people enter competitive schemes more often?

Competition increases dishonesty when participants are randomly assigned to schemes.

With selection, those who are more dishonest select competition more often.

Selection, Tournaments, and Dishonesty[1]

Marco Faravelli
School of Economics
University of Queensland
m.faravelli@uq.edu.au

Lana Friesen
School of Economics
University of Queensland
l.friesen@uq.edu.au

Lata Gangadharan
Department of Economics
Monash University
Lata.Gangadharan@monash.edu

October 2014

Abstract

We conduct a real effort experiment in which performance is not monitored and participants are paid according to their reported performance. Participants are paid according to a piece rate and a winner-take-all tournament and then select between the two schemes before performing the task one more time. Competition increases dishonesty and lowers output when the payment scheme is exogenously determined. Participants with a higher propensity to be dishonest are more likely to select into competition. However after selection, we find no output difference between piece rate and tournament. This is attributable to a handful of honest individuals who select competition.

Keywords: Dishonesty, Selection, Tournament, Piece rate, Real effort experiment

JEL Codes: C91, K42

---

[1] Acknowledgements: We are grateful for valuable comments provided by Tim Cason, Christine Harbring, and by conference participants at the International Meeting of the Economic Science Association (2013), the Australia New Zealand Workshop on Experimental Economics (2013) and the Asia-Pacific ESA Meeting 2014. Funding was provided by University of Queensland and Monash University. This research project has been approved by the Behavioural and Social Science Ethical Review Committee, at the University of Queensland. We retain responsibility for any errors.

1. Introduction

One of the most important lessons from economics is that competition is beneficial because it is efficiency enhancing. A labor tournament, a democratic election or a marathon, all share the same goal: to maximize performance by rewarding the most productive worker, the most talented politician or the fastest runner. However, reality is not that simple. Firstly, it is often difficult, if not impossible, to measure actual output and contestants might be able to misrepresent their performance. Thus an employee may exaggerate her accomplishments, a politician may lie about her achievements and a sportsman may take performance-enhancing drugs. Secondly, participation in contests is endogenous. Some people choose to enter competitive careers, become politicians or professional sportsmen while others shy away from competition (e.g. Niederle and Vesterlund, 2007). The possibility of being dishonest, on the one hand, and the presence of selection on the other raise a number of relevant and interesting questions which we will address in this paper. First, are people more or less dishonest in competitive situations? Second, does competition actually improve performance when contestants can lie? Third, who is more likely to select into a competitive scheme, honest or dishonest people, and how does this selection impact on the effectiveness of the competitive scheme?

Economists have shown theoretically that tournaments are beneficial because they increase production (e.g. Lazear and Rosen, 1981), and some have also modeled sabotage and dishonesty in tournaments theoretically (e.g. Konrad, 2000; Curry and Mongrain, 2009). Empirical evidence supports some of these theoretical predictions (e.g. Ehrenberg and Bognanno, 1990). In this paper we use an experimental approach as it allows us to observe dishonest actions empirically, which can be difficult using field data. We present a real effort experiment in which participants perform a similar task under different payment schemes. Actual performance is not monitored and payment depends on self-reported performance.

Each participant performs the task three times. In the first two tasks subjects are paid according to either a piece rate scheme or a winner-take-all tournament in which they are randomly matched with another participant. After having experienced both mechanisms, each participant selects either the tournament or the piece rate as a payment scheme for the final task.

We find that the tournament incentive scheme results in significantly higher dishonesty, but also substantially lower output as compared to the piece rate scheme. In our experiment, for a given performance, the tournament winner receives a higher payoff than in the piece rate scheme; while, on the contrary, the loser receives nothing. Thus, the higher opportunity cost of being honest in the tournament might explain the increase in dishonesty. This is despite the fact that there is evidence that individuals face a psychological cost of lying (e.g. Gneezy, 2005; Mazar et al., 2008; Gneezy et al., 2013). We refer to this as the increase in dishonesty due to the *competition effect*. This in turn can explain the decrease in output. Once an individual decides to lie, or to increase her level of dishonesty, the relative benefit of honestly produced output suddenly falls. Moreover, such benefit decreases even further if one expects others to lie. Thus we find that, when monitoring is not available, not only does competition increase dishonesty, but it can also damage performance.

We then turn to the issue of selection, which represents the main research question of our paper. Who is more likely to choose the tournament over the piece rate? And what is the effect of letting people select the payment scheme? We find that dishonest subjects are more likely to select into tournaments. As a consequence, when participation in the contest is endogenous we have a doubly perverse effect. On the one hand, dishonesty increases as a consequence of the winner-take-all nature of the tournament, which we described earlier as the *competition effect*. A politician may lie to her electorate or a sportsman may use performance enhancing drugs because of the pressure of the winner-take all contest. On the other hand, individuals with a higher propensity to lie are more likely to enter the contest, causing an even more pronounced level of dishonesty. This can be termed the *selection effect*. As they chose to compete, the politician and the sportsman are more likely than others to be dishonest.

Finally, what is the effect of selection on output? Interestingly, at the selection stage, we find no difference in output between the piece rate and the tournament. This is attributable to a handful of honest individuals. While most honest subjects select the piece rate scheme, a few of them choose the tournament. Being honest, they do not consider lying as an option and thus, having opted for the contest, they work hard and produce high output. Hence, despite the low performance of dishonest individuals, overall output does not differ across the two schemes. This is an intriguing and encouraging result. While competition increases the amount of dishonesty it does not necessarily harm performance. This result however is subject to an important caveat, as it depends on the initial composition of the population and the fraction of honest people who are attracted by competition.

A number of experiments on dishonesty and deception have been conducted in recent years.[2] Gneezy's (2005) seminal contribution considered a sender-receiver game in which a player can lie and thus increase her payoff at the expense of her opponent. His main finding is that, on average, individuals are averse to lying when this leads to a small increase in their own payoffs but a substantial reduction in others' payoffs.[3] Mazar et al. (2008) and Fischbacher and Heusi (2013) also find evidence of aversion to lying but in individual decision-making tasks (although neither observe individual lying behavior but instead rely on

[2] See, for instance, Cappelen et al., 2013; Battigalli et al., 2013; Gino et al., 2013; Ploner and Regner, 2013, among others.

[3] Several subsequent studies have also used the sender-receiver framework to study dishonesty. See, for example, Ellingsen and Johannesson (2004), Dreber and Johannesson (2008), Lundquist et al. (2009), Maas and Rinsum (2013), and Innes and Mitra (2013).

session level data). While lying does occur it is less than the full extent possible, which can be attributed to aversion to lying.[4]

The relation between dishonesty and payment schemes is the subject of only a limited number of recent experiments. Some studies focus on target-based schemes (Schweitzer et al., 2004; Cadsby et al., 2010)[5] and team incentives (Conrads et al., 2011; Danilov et al., 2013). Closer in spirit to us is a paper by Schwieren and Weichselbaumer (2010). Using the real effort task first employed by Gneezy et al. (2003), they compare a piece rate and a tournament with a focus on gender and ability. They find no difference in the overall proportion of those who cheat. Competition, however has an impact on women's cheating behavior, though this is mainly explained by differing ability levels between the genders in the real effort task.[6] Belot and Schroeder (2013) conduct a real effort experiment comparing a winner-take-all contest, a piece rate and a flat rate scheme with a focus on whether subjects simply misreport their performance or steal money. As no evidence of theft is found, the authors conclude that subjects' behavior can be explained by social norms, but not by other regarding preferences. Finally, Rigdon and D'Esterre (2012) also employ the matrix task introduced by Mazar et al. (2008). They compare a piece rate scheme with self-grading, a tournament with self-grading and a tournament in which subjects grade their opponent. Their main finding is that people are more willing to inflate their result than to down grade others.

Last but not least, our experiment is related to the literature on sabotage (e.g., Harbring and Irlenbush, 2005, 2008, 2013; Carpenter, 2010), in which researchers have found that sabotage is more common in tournaments than in piece rate situations, leading to less

[4]These studies typically use students as subjects. In contrast, Abeler et al. (2014) use a representative sample and find that aggregate reporting behavior is close to the expected truthful distribution, suggesting that people rarely lie.
[5] See also Gill et al. (2013) as an example of target-based schemes where bonuses are random.
[6] Gender differences in dishonest behavior have been reported by several researchers. Friesen and Gangadharan (2012), for example, find that in the matrix game men are not only more likely to be dishonest than women, but they are also likely to be more dishonest.

output. While sabotage is aimed at reducing the output of the other contestants, subjects in our experiment, in contrast, can only inflate their own performance. Sabotage is interesting and important to study, however there may be limited opportunities to harm others in the field. Inflating one's own performance on the other hand is easier to engage in and is commonly observed. We focus on this aspect in our research.

None of the above papers explore the issue of entry into tournaments, which is instead the object of a vast literature on gender differences (e.g. Niederle and Vesterlund, 2007; 2013; Gneezy et al., 2009) and contest design (e.g., Cason et al., 2010). Our paper builds on the above research on dishonesty and is unique in addressing selection into tournaments when there is an opportunity for individuals to be dishonest. We present an experiment with both within and between subject design features and measure dishonesty at the individual level. A within subject design allows us to identify whether incentive schemes alter participants' honesty and explain who selects into tournaments. By varying the sequence of the payment schemes in a between subject design we eliminate the potential confounds created due to learning or behavioral spillovers. Our findings have implications regarding how financial incentives should be designed for managers, employees, politicians, and people in many other professions.

The paper proceeds as follows. Section 2 describes the experimental design and procedures. Section 3 reports the results from the analysis of the data. Section 4 explains our findings and the distinction between the *competition effect* and the *selection effect* arising due to dishonesty. Section 5 discusses the implications of our results and concludes by suggesting avenues for future research.

2. Experimental Design

All subjects participated in four tasks. The first task was an investment task, similar to the one discussed in Gneezy and Potters (1997) and Gneezy et al. (2009). In this task each participant decided how much of their endowment of $AUD 5.00 to put into an investment. There was an equal chance that the investment would yield triple the amount invested or nothing. The outcome was decided at the end of the experiment by a flip of a coin. This task was conducted to elicit participants' attitudes towards risk, which could influence their propensity to be dishonest or tournament selection decision.

In Tasks 2-4 we use the matrix task devised by Mazar et al. (2008) to examine if participants are dishonest when they have an opportunity. In contrast to Mazar et al. (2008), we collect individual level data.[7] One of these three matrix tasks was randomly chosen for payment at the end of the session. In each of the tasks, there were two stages. In the matrix stage, subjects were given a sheet of 20 matrices, where each matrix contained 12 three-digit numbers (e.g. 5.34). The task was to find a pair of numbers in each matrix that add up exactly to 10.00. A sample matrix is shown in Figure 1.[8] However, the task was made more difficult because not all of the matrices have solutions, of which subjects were made aware. Subjects were given five minutes to solve as many matrices as possible. Once the five minutes were up, subjects were asked to count the matrices they have solved and take note of it. They then folded the matrix sheet and put it in a box, which would not be opened until everyone had left the laboratory. After this, in the reporting stage, subjects were asked to write down on a collection slip the number of solved matrices they would like to report. It was specified that the number reported would be used to determine their earnings in case that particular task was selected for payment.

[7] Mazar et al. (2008) measure dishonesty at the group or the session level as they do not collect subjects' matrix sheets. Friesen and Gangadharan (2012, 2013) also use this task to examine individual level differences in dishonesty.

[8] Experimental instructions are contained in the Appendix.

This task allows us to measure the degree and existence of dishonesty at an individual level, in a situation where there is no plausible way of being detected or punished. There are several reasons why this task is useful to measure dishonesty (Mazar et al., 2008). First, subjects consider that the outcome is predominantly effort related rather than IQ related. Second, subjects can readily evaluate their own performance – they know if they have the correct answer or not. This means that any dishonest gain can be reasonably interpreted as cheating, rather than as a genuine mistake.

Tasks 2 and 3 differ only in the way the subjects were to be paid. In Task 2, the Piece Rate task, participants were told that they would earn $1 for each matrix they reported having correctly solved. In Task 3, the Tournament task, each participant was randomly paired with another person in the same session and their payment depended on their performance relative to that of their partner. If they reported solving more matrices than their partner, they would receive $2 per reported matrix, else they receive zero.[9] In order to maintain the same incentives across payment schemes, the tournament task has an element of piece rate as well, as in Niederle and Vesterlund (2007). Task 4 asked subjects to first select a payment option before they participated in the matrix task. They could choose between the Piece Rate scheme or the Tournament scheme, which would be applied to the number of matrices they reported solving in this task. If they chose the Tournament scheme, they would be randomly paired with another person who had also chosen this scheme in this task and, as in Task 3, they would be paid $2 per reported matrix if they reported solving more than their partner and zero otherwise. Participants therefore compete against the competitive performances of others,

[9] In the case of a tie, the experimenter tossed a coin in front of the subject to determine the winner.

who had also selected competition.[10,11] We chose to have a higher payoff in the Tournament as our main aim is to study selection into the two schemes and without this feature there would be no monetary motivation to enter the Tournament. Choosing double the payoff in the Tournament helps keep the expected value similar in the two payment schemes (as in Niederle and Vesterlund, 2007).[12]

After Task 4, subjects completed a questionnaire which collected both standard demographic measures as well as attitudes towards lying. They were then paid. All subjects received a $5 show up fee, payment from the investment task, and payment from one of the randomly chosen Tasks 2-4. Earnings ranged from AUD $5 to AUD $60 with an average of AUD $21.14 (approximately US $20).

We conducted the tasks in two different orders. In the first order, subjects participated in the four tasks in the sequence explained above. In the second order, subjects participated in the Tournament task before the Piece Rate task. We varied the order of the payment mechanism to examine if subjects who were exposed to the Tournament first had different behavioral responses towards dishonesty than subjects who were exposed to the Piece Rate

[10] This is different from the design adopted by Niederle and Vesterlund (2007) in their seminal paper that has inspired several selection experiments. In their study, subjects who choose the tournament compete against the performance of other participants in the (previously run) piece rate scheme. While their design presents a number of advantages, we wanted to represent a more realistic situation in which, when selecting a tournament, an individual competes against the performance of others who also chose to enter the contest.

[11] If Task 4 was randomly chosen for payment the subjects were matched as follows. The subject drew a card from a box containing the personal ID numbers of all the other participants who had chosen the Tournament in the session. Her performance was matched to that of the person whose ID was drawn. The matched participant's payoff remained unaffected. The subjects were informed of this process beforehand. This ensured that a random matching process could be maintained irrespective of an odd or even number of subjects choosing to enter the Tournament. A similar matching process was used if Task 3 was chosen for payment, except that the box contained the personal ID numbers of all participants in the session.

[12] This equivalence in expected values assumes that subjects are a priori identical, which in the absence of any other information is the most natural assumption.

first.[13] Moreover we were interested in documenting and controlling for whether exposure to one of them leads to spillover effects on dishonesty in the other task.

All sessions were conducted at the University of Queensland. One hundred and nineteen subjects participated in the experiment, with approximately 30 in each session and half in each order. Each subject participated only once. Subjects were predominately undergraduate students and were drawn from different academic disciplines and broadly recruited across the university by email using ORSEE (Greiner, 2004). Although some had participated in other economics experiments, all were inexperienced in the sense that they had never participated in a similar experiment with any opportunity for dishonesty or one having identical payment schemes. Of the 119 participants, 55% were male.

Upon arrival, subjects were randomly allocated a seat in the experimental laboratory. Even though our procedures were non-computerized, the laboratory provided a convenient setting as the cubicles enhanced privacy and anonymity. Each of the subjects had a set of the general instructions on their desk, a faced down personal ID card and four sealed envelopes. The personal ID was reproduced on all matrix sheets, reporting collection slips and forms. At the beginning of each session an experimenter read the instructions aloud while subjects followed along on their own copy. Instructions for the next task were only given once the previous task was completed. There was no feedback given to subjects between tasks. There was a box for the matrix sheets kept at one end of the laboratory and subjects were asked to fold and put their solved matrix sheets in that box for Tasks 2-4. Three reporting boxes were also placed at the opposite end of the laboratory from the matrix box, one for each of the Tasks 2-4, in which subjects were asked to put their collection slips reporting how many

[13] Harbring (2010) finds that being exposed to different incentive schemes can change behavior, with participants sending more in a trust game after being exposed to a cooperative team setting as compared to a tournament setting.

matrices they had solved. The experimenters made it clear that they would not be touching the matrix box during the session and that only the relevant reporting box would be opened up depending on which task was selected for payment. Separating out the matrix solving and reporting stage (subjects were instructed to return to their desks to complete the reporting form after placing their completed matrix sheet in the box) and creating different boxes for the sheets helped reinforce the fact that the probability of subjects' dishonesty being detected and their being punished or frowned upon was zero. Including the instruction and payment distribution time, sessions lasted between 90 and 120 minutes.

3. Results

In this section we present results that address our three research questions. Summary statistics for all the variables used in the analysis are provided in Table 1.

3.1 Which Incentive Scheme Results in Greater Dishonesty?

Recall that subjects had the opportunity to be dishonest by reporting to have solved more matrices than they had in fact completed. After all the experimental sessions were completed, we opened the matrix box and assessed how many matrices were correctly solved by each subject in each task. We then matched this with the relevant reporting slip to construct a measure of dishonesty in each task equal to the difference between the amount reported and the actual amount solved.[14] For each subject we have three such measures; one for each of Tasks 2, 3, and 4.

In addition to this measure of the magnitude of dishonesty in each task, we constructed two additional measures. First, a binary indicator of whether or not the subject

[14] As mentioned before an ID number was written on each reporting slip and matrix sheet enabling them to be matched later. To enhance anonymity we randomly distributed these across desks so that no participant could be easily identified with a particular ID number.

was dishonest in a particular task. Second, a measure of the magnitude of dishonesty as a percentage of the maximum degree of dishonesty possible given the number of matrices solved (i.e., 20 – number of matrices solved).

It is worth pointing out that the nature of the task is such that subjects *should* be able to assess the correctness of their own performance. Thus we are interpreting excess reporting as dishonesty rather than as a genuine mistake.[15] Later in this section we examine the robustness of our conclusions to using an alternative measure which categorizes "small" lies as mistakes rather than deliberate falsification. Also note that because not all matrices have a solution (only 10 in fact did), even the most proficient participant had ample incentive to be dishonest.

In Table 2 we report these three measures of dishonesty for Tasks 2 and 3, both overall and disaggregated by treatment and order. We also report the magnitude of dishonesty conditional on it occurring. The table also reports p-values from non-parametric paired Wilcoxon signed rank tests of differences between dishonesty in the Piece Rate versus the Tournament. Dishonesty in Task 4 is considered in the next section.

The first section of Table 2 reports results pooled over order. For all measures of dishonesty, the Tournament leads to more dishonesty than the Piece Rate. In particular, there is a substantial increase in the proportion of liars from 45% in the Piece Rate to 63% in the Tournament. Conditional on lying, the size of dishonesty is approximately the same, leading to an overall increase in the magnitude of lying (4.22 versus 3.29).

Disaggregating by order, the table shows that these results are stronger in the sessions that started with the Piece Rate, with significant increases in all three dishonesty measures when the Tournament is employed (Wilcoxon signed rank p-values<0.01). While the

[15] Only two people erred on the negative side each reporting one fewer matrix than they had actually solved correctly.

direction of the effect is the same in the sessions that started with the Tournament, significant differences are only detected for two of the dishonesty measures (p-values<0.10) and not when dishonesty is measured as a percentage of the maximum possible lie (p-value=0.586).

We find evidence of a substantial order effect whereby subjects who experience the Tournament first lie more in all categories for both incentive schemes. The difference is strongly significant for the Piece Rate, using a Mann-Whitney rank-sum test to compare dishonesty when the Piece Rate comes first versus when it follows the Tournament task (p-values of 0.004, 0.034, and 0.003 respectively, for size of lie, proportion of dishonest subjects, and percentage of the maximum possible lie). In the Tournament task, however, while experiencing the Tournament first leads to greater dishonesty, this difference is not significant (p-values of 0.146, 0.228, 0.151 respectively for size of lie, proportion of dishonest subjects, and percentage of the maximum possible lie, using Mann-Whitney rank-sum test). Hence, the greater subsequent lying in the Piece Rate task when it follows the Tournament makes it harder to detect any differences across the payment mechanisms. The order effect suggests that the motivations for being dishonest, once evoked, can have spillover effects. This could be an important consideration in choosing payment mechanisms in organizations.

To avoid the confounding implications of the order effect and present a cleaner comparison of the two incentive schemes we conduct between-subject tests of dishonesty in the first (matrix) task encountered. The results, reported in the first panel of Table 3, clearly show that not only do more people lie in the Tournament (nearly double the proportion), they also lie by a significantly larger amount (by nearly twice as much, and three times more for the percentage of the maximum possible lie measure). These results are illustrated in Figure 2, which shows the distribution of lying in the first task encountered.

Figure 3 illustrates these results in a different way. The figure plots the number of correctly solved matrices on the horizontal axis with the number of matrices reported as solved on the vertical axis. Thus accurate reporting occurs along the diagonal. The size of the bubble reflects the number of observations at each coordinate. As is evident from the graph, in the Tournament many more observations occur above the diagonal line, including a number at the maximum report possible. There is also no apparent correlation between lying and being successful in the task.

Finally we explore the robustness of our conclusion by adjusting our dishonesty measures for the possibility of genuine mistakes. As mentioned in the previous section, our measure of dishonesty assumes deliberate over-reporting on the part of subjects when the number of reported matrices exceeds the number actually solved. An alternative interpretation is that many of the "small" lies are in fact (unintentional) errors.[16] In fact, the most common category, after a lie of zero, is a lie of one (16% of subjects in the Piece Rate; 17% in the Tournament) thus this adjustment might affect our conclusions. To examine this, we construct an alternative error-adjusted measure of dishonesty where we recode lies of size one as errors if the subject never lied by more than one in either task. That is, only if a subject overstates by one in one task and zero in the other, is this classified an error. (No one lied by one in both tasks.) The rationale is that if a person lies by one in one task but by seven in the other task, it seems less likely to be a mistake. Of the 20 people who lied by one in the Tournament, four lied by more than one in the Piece Rate, and of the 19 people who lied by one in the Piece Rate ten lied by more than one in the Tournament.

---

[16] A number of factors in fact mitigate this view. First, if it is an error, we would observe some *under*-reporting as well. However *under*-reporting is virtually non-existent (2 out of 119 subjects underreported in one task). Second, the task we use allows for self-checking and thus the line between deliberate carelessness and deliberate falsification is a murky one. Third, in some pilot sessions we actually checked the matrix sheets during the session and found very few errors (1 out of 16 subjects). Nevertheless since small lies are commonplace we examine the robustness of our results.

First consider the overall (pooled) results, as well as disaggregated by order (i.e., as in Table 2). Using this error-adjusted measure, differences in the overall results and when the Piece Rate came first remain strongly significant (i.e., the Tournament leads to greater dishonesty than the Piece Rate). However, when the Tournament comes first the size of lie and proportion of dishonest subjects in the Tournament is no longer significantly different than in the Piece Rate (p-values of 0.17 and 0.16 respectively, Wilcoxon signed rank test).

Next consider whether the cleaner, first task, results are also robust to this alternative definition of dishonesty. In their first task, 21 people over-reported by one matrix, but nine of these lied by two or more in their second task, thus only 12 were counted as errors in the adjusted measure. The error-adjusted averages are shown in the bottom panel of Table 3 along with the relevant p-values. They show that the main result that dishonesty is higher in the Tournament is robust to a different interpretation of small lies.

To further analyze dishonest behavior, we classify subjects into one of the four following categories, based on their behavior in the two incentive schemes: 1): subjects who lie in both the Piece Rate and the Tournament (*AlwaysLie)*; 2) subjects who lie only in the Piece Rate *(LiePRonly)*; 3) subjects who lie only in the Tournament *(LieTonly)*; and 4): subjects who lie in neither task (*NeverLie)*. Table 4 shows the proportion of subjects falling into each category using the raw measure of lying and then the measure that adjusts for "small" lies as errors. The table shows that around one-third of subjects always lie, while very few people lie only in the Piece Rate. A larger proportion lies only in the Tournament (27% using the raw data, 13% using the error-adjusted measure). Subjects in this category may feel compelled to lie to compete against partner's dishonesty. This motivation is not relevant for people in the *LiePRonly* category. Hence subjects in the *AlwaysLie* and *LiePRonly* category can be thought to be relatively more dishonest while those in the *NeverLie* and *LieTonly* categories can be classified as relatively more honest.

We also conducted regressions to control for possible differences in demographic factors such as gender and risk preferences, both of which may affect the decision to be dishonest. Using dishonesty in the first task as the dependent variable we ran several different models controlling for demographic differences. While the estimated coefficient on the Piece Rate indicator is significantly negative in all specifications, confirming the non-parametric results of less lying in the Piece Rate, none of the demographic controls are significant, so we do not report the regressions here.

In summary, our first result can be stated as follows:

Result 1: *Tournament incentives lead to more dishonesty than Piece Rate incentives: not only do more people lie, they also lie by a greater amount when they are dishonest. This is observed irrespective of whether we pool the two order sequences, examine them separately, only focus on the first task participants experience, or if we adjust the measure of dishonesty for possible errors.*

3.2 Does Competition Actually Improve Performance When Participants Can Lie?

We have demonstrated how the Tournament payment scheme has the perverse effect of increasing dishonesty. Nevertheless this scheme may still be beneficial if output is increased. After all, greater output is one of the advantages of competitive incentive schemes. Recall that each matrix sheet contained 20 matrices but only 10 were in fact solvable. In both tasks, the number of correctly solved matrices ranged from zero to nine. Figure 4 shows the distribution of output in the first (matrix) task disaggregated by incentive scheme. The figure shows a clear right shift in the distribution in the Piece Rate scheme. The modal output in the Tournament is two correctly solved matrices versus four in the Piece Rate. Also, zero output is more common in the Tournament.

In Table 5 we report the average number of matrices solved. As is clear from the table, output is significantly lower in the Tournament across all measures (pooled, by order, and also focusing on the first task). Thus not only does the Tournament lead to greater dishonesty but it also results in significantly less output. Regression analysis, using the output in the first (matrix) task as the dependent variable corroborates that the number of matrices solved correctly is significantly higher in the piece rate task. None of the demographic variables have any consistent explanatory power.[17]

Table 6 reports the number of matrices solved by subjects in the first task they participate in, disaggregated by subject type. The number of matrices solved is significantly higher in the Piece Rate for subjects who never lie and for those who only lie in the Tournament irrespective of whether subjects are classified using the raw or the error adjusted measure. The Tournament therefore induces the relatively more honest types (those who never lie or who lie only when compelled to in the Tournament) to produce less output.

Our second result can be summarized as follows:

Result 2: *Output levels are significantly lower in the Tournament as compared to the Piece Rate. Honest types are more likely to have lower output in the Tournament.*

3.3 Selection into Tournaments

In this section we examine behavior in Task 4, beginning with the analysis of the Tournament entry decision. In the previous tasks, the payment schemes are given to the participants, whereas in this task participants can choose the payment mechanism. We will refer to these as exogenous and endogenous mechanisms.

[17] These results are available from the authors, but not reported in the paper to save space.

Various individual characteristics are likely to affect selection, such as the psychological cost of lying (see Gneezy, 2005), ability, demographic characteristics, performance in the specific task, as well as beliefs about the honesty of others. While for each one of these features in isolation it may be possible to predict whether they encourage or discourage entry into tournaments, this task becomes exceedingly complicated when we consider all these factors jointly. In this section therefore we look for empirical evidence pertaining to the importance of these factors.

3.3.1 Determinants of Entry Decisions

We first examine whether dishonest people are more likely to select into the Tournament payment scheme. Of those who always lie 40% choose the Tournament scheme, compared with only 24% of those who never lie, providing suggestive evidence that the answer is yes. About two-thirds (66%) of subjects choose the Piece Rate, while one-third choose the Tournament. Men are significantly more likely to choose the Tournament (46% of men versus 21% of women; p-value=0.004, Mann-Whitney rank sum test). Those choosing the Tournament are less risk averse as measured by a greater amount invested in Task 1 (3.94 versus 3.45; p-value=0.051, Mann-Whitney rank sum test). While these results are as expected, they suggest the importance of controlling for demographic factors when assessing how dishonesty affects the selection decision.

Thus we present results of Probit regressions where the dependent variable is an indicator of choosing the Tournament in Task 4 (*TChoice*). The regressors of primary interest are four measures of dishonesty (included one at a time): an indicator that equals one if a subject always lies (*AlwaysLie*), an indicator of lying at least once in Tasks 2 and 3 (*EverLie*), and indicators of lying in the Piece Rate (*LiePR*) or the Tournament (*LieT*) respectively.[18] We also include the following control variables: an indicator of participating in the Piece Rate task first (*PRFirst)*, gender (*Male*), the amount invested in Task 1 as a proxy for risk preferences (*Invested*), an indicator of whether the subject studies economics or business (*EconBus*), and an indicator if the subject is locally born (*LocalBorn*). Standard errors are clustered by session. The coefficient estimates for the four separate models are reported in Table 7.

The following results are apparent from the table. First, those who are dishonest in the earlier tasks are more likely to select the Tournament in Task 4. The coefficient on dishonesty is significantly positive for all four measures. The magnitude of the effect varies from 9 percentage points (*LiePR*) to 16 percentage points (*EverLie*). Given that on average 33% of subjects choose the Tournament, the effect is substantial: a 27%-48% increase depending on which measure of lying is used. Second, only one control variable is consistently significant: men are significantly (at the 10% level) less likely to choose the Piece Rate. No other variables are significant, although *Invested* (the measure of risk preference) has the expected positive sign. The variable capturing the order in which subjects participated in the tasks (*PRFirst*) is not significant.[19]

[18] Note that these four measures differ from the four mutually exclusive categories that we divided subjects into in Table 4 as our purposes are different. Here we construct different, plausible, measures of dishonesty and examine if they predict entry decisions. Clearly *AlwaysLie* is the same, while *EverLie* is the inverse of *NeverLie*.

[19] Alternative specifications in which the different type classifications are employed as dummy variables yield similar results, with subjects who always lie 17 percentage points more likely to enter the Tournament than those who never lie. These specifications are available from the authors. Regressions using the magnitude of dishonesty in either the piece rate or the tournament as the measure of dishonesty are statistically insignificant (though the point estimates have the expected sign), suggesting that the distinction between honest and dishonest is more important than how dishonest one is. We also included additional control variables such as age, GPA, year level, income, and performance in tasks 1 and 2 but they are never significant, thus we report only those regressors that seem most relevant. In additional specifications we examined whether performance in the previous tasks has a non-monotonic relationship with the tournament entry decision. The performance variables do not have a statistically significant impact.

In Table 8 the estimated coefficients from regressions using the error-adjusted measure of dishonesty are reported. For brevity, only the coefficient on the dishonesty variable is reported, with other control variables included as in Table 7 with similar results. As the table shows, selection into the Tournament is more common among those who always lie and among those who lie in the Piece Rate. Since very few lie only in the Piece Rate (3%), these two categories essentially coincide. We find a significant link between dishonesty in the Piece Rate task and selection into the Tournament as well as between subjects who always lie and the selection decision. Perhaps lying in the Piece Rate provides a more accurate measure of "inherent dishonesty" versus "feeling compelled" to lie in the Tournament (because of the pressure of competition). The relatively more dishonest subjects (classified as mentioned above, as those who always lie and those who lie in the Piece Rate) are more likely to enter the tournament.

At the completion of Task 4, subjects were asked an open-ended question about why they chose a particular incentive scheme. Two research assistants, both postgraduate students, independently coded these responses into the eight (non-mutually exclusive) categories shown in Table 9. Neither assistant was familiar with the purpose of the experiment or any decisions made by the subjects during the experiment. To assess the reliability of the coding we computed Cohen's Kappa (Cohen, 1960; Krippendorff, 2013), which accounts for the fact that agreement between the coders might occur by chance.[20] Landis and Koch (1977) suggest that values above 0.60 show "substantial agreement", while those between 0.40-0.60 show "moderate agreement". Thus for all categories, except for "dislike of competing" reliability appears high. Risk and return were mentioned the most often, followed by expressions of confidence or lack in one's task ability, and dishonesty inclinations.

[20] This measure is more informative than the more common procedure of reporting the agreement percentage across coders. Cason and Mui (2014) use the Kappa statistic when coding chat content in an experiment.

Table 9 also assesses the link between these stated reasons and the actual selection decisions in Task 4. For example, those who think others are dishonest never choose the Tournament, while those who report that they were planning to cheat were significantly more likely to choose the Tournament than those who did not mention this, providing support for our results.[21] Those who consider themselves smart or good at the task always choose the Tournament, while those who consider themselves weak or of low ability always choose the Piece Rate. The risk related to the Tournament is a significant determinant in subjects' choices. Those subjects who state that choosing the Tournament is not very risky and has a big payoff, always choose the Tournament. The percentages reported are for one coder only, but are similar across coders, while the p-values are conservative and report only the larger p-value of the two coders. There were no substantive differences across coders.

3.3.2 Task 4 Dishonesty and Output

Having demonstrated that dishonest people are more likely to select into the Tournament payment scheme in Task 4, we now determine what subjects do in Task 4 conditional on their selection decision. Two-thirds (66%) of those choosing the Tournament lie in Task 4, while only 59% of those choosing the Piece Rate do. The magnitude of lying is also higher in the Tournament (5.6 versus 4.2), but neither difference is statistically significant. Of those who chose the Tournament, the relatively dishonest subjects exhibit a dramatically higher magnitude of lying as compared to the relatively honest ones (9.2 versus 1.7, p-value= 0.0001, Mann-Whitney rank sum test)

Table 10 shows the estimates from two Tobit models examining the determinants of the size of the lie in Task 4. The first model only includes an indicator that equals one if the

[21] Of the 9% who stated that they would cheat, 80% chose the tournament. On the other hand, among the remaining 91% who did not state this reason only 30% chose the competitive scheme.

subject chooses the Tournament (*TChoice*) and an indicator that equals one if the subject participated in the Piece Rate task first in Tasks 2-3 (*PRFirst*). The second model specification controls for risk preferences and other individual specific characteristics as described earlier. The estimates from both models show that those who selected the Tournament payment mechanism in Task 4 lie by significantly more than those choosing the Piece Rate payment mechanism.[22] There is some evidence that the subjects who experience the Piece Rate first lie less in this task (p-value <0.10).

Now we turn our attention to the performance of subjects in this task. The average number of matrices solved by the 78 subjects who choose the Piece Rate is 4.26 as compared to 4.39 by the 41 subjects who choose the Tournament (this difference is not statistically significant). The standard deviation for the Piece Rate choices is lower than for the Tournament choices (2.12 versus 2.35), suggesting there is more variance of output in the Tournament.

Table 11 presents a Tobit regression of the number of matrices solved in Task 4 as a function of the choice of payment scheme (*TChoice*), the number of matrices solved in the first task they encountered (*Matrices2*), and demographic factors (as defined earlier). To explore how subject type influences Task 4 output contingent on their entry decision, we also include interaction terms between two subject types, *NeverLie* and *LieTonly*, and *TChoice*. We chose these two types to interact with *TChoice* because our earlier analysis (see Table 6) showed that these relatively honest types reduced their output when the Tournament was imposed upon them, thus we are interested to see if this effect continues here.[23] The results show that subjects who solved more matrices in the first task they encountered, also solved

[22] In alternative specifications we included an interaction term between *TChoice* and dishonest type and find support for the result from the non-parametric tests that the size of the lie is significantly higher for the relatively dishonest subjects who choose the Tournament.
[23] The order of the previous task (*PRFirst*) is not included in this regression as this variable is significantly correlated to the output in the first task (p-value = 0.0002).

more matrices in this task. Relatively honest types (*NeverLie, LieTonly*) who choose the Tournament in this task solve significantly more matrices than the more dishonest types (*AlwaysLie, LiePRonly*). This negative effect for the dishonest subjects is captured by the *Tchoice* variable, which has a coefficient that is marginally insignificant (p-value = 0.13). A joint test of significance indicates that the honest subjects who choose the Tournament have jointly a significantly higher output as compared to the dishonest subjects who also choose the Tournament (F-statistic (2,110) = 4.01, p-value = 0.02). Further, comparing the predicted output of honest subjects with dishonest ones shows that the linear combination of the output of the honest subjects is 2.08 matrices more than the dishonest ones (2.64 matrices from the linear combination of coefficients estimate minus 0.56 matrices solved by the dishonest subjects who chose the Tournament). The estimate from the linear combination of coefficients test is significantly higher, with a p-value = 0.006.

The finding that honest subjects produce higher output in the Tournament indicates that, when subjects are allowed to select into an incentive scheme, then some of the honest ones produce a higher output after choosing the Tournament. In contrast, in Section 3.2 we found that the relatively honest subjects produce significantly lower output when the Tournament is exogenous, resulting in lower output levels in the Tournament. Hence, choosing to compete can lead to different output levels as compared to being compelled to compete.

To explore this further, we compare the output of subjects when they are required to compete versus when they choose to compete. In aggregate, subjects who choose the Tournament significantly increase output from 3.37 matrices being solved in the exogenous Tournament to 4.39 being solved in the endogenous Tournament (p-value=0.04, Wilcoxon signed rank test). Subjects who never lie solve 2.4 matrices on an average when compelled to compete compared to 5.0 when they choose to compete, a statistically significant difference (p-value = 0.06, Wilcoxon signed rank test). The differences are not significant for the other types of subjects, although for *LieTonly* the difference in output is substantial in magnitude (3.8 when imposed and 5.1 when chosen, p-value = 0.21, Wilcoxon signed rank test). Hence this increase in output is driven by the relatively honest subjects who choose the Tournament in Task 4 and produce greater output.[24] Thus this change in output from honest subjects is the reason why output is not significantly different between the Piece Rate and Tournament in Task 4.

This leads us to Result 3.

Result 3: *Dishonest people are more likely to select into tournaments. Upon entry they lie more and their output is lower than honest people who enter the Tournament. However, due to the high output of honest people, overall output is not significantly different between the Tournament and the Piece Rate.*

4. Discussion

The main findings of our paper are the following. First, we find evidence that people lie more in tournaments than in piece rate schemes. Second, when the payment scheme is exogenously given, output is lower in tournaments. Third, dishonest people are more likely to enter tournaments; however, in spite of this, when the payment scheme is endogenous there is no overall difference in output between piece rate and tournament. In this section we aim to explain these results.

[24]On the other hand, subjects who choose the Piece Rate do not show a significant increase in output from the exogenous Piece Rate (4.45 matrices solved) to the endogenous Piece Rate (4.25 matrices solved).Thus there does not appear to be a general learning effect over time which might confound our interpretation.

In order to understand our findings we must abandon the *homo economicus* paradigm and assume that people face a cost of lying (be it moral or psychological). This notion allows us to explain why some subjects are honest and why a substantial proportion of those who lie do not lie fully.[25] Under this framework, the greater dishonesty encountered in tournaments can be interpreted as a result of the higher opportunity cost of being honest: while the tournament winner receives twice the reward per reported matrix as in the piece rate scheme, the loser receives nothing.[26] We call this the *competition effect*. Such an effect in turn can account for the lower productivity we find in tournaments. This is because conditional on having chosen to lie to a greater extent than in the piece rate scheme, the relative benefit of correctly solving an extra matrix suddenly drops. Moreover, if one expects the others to act dishonestly the benefit drops even further. Carpenter et al. (2010) find a similar effect in a different setting related to sabotage and competition: the possibility of sabotage more than offsets the incentives of a tournament, leading to a lower output than under a piece rate scheme. On the contrary, Belot and Schroeder (2013) find that competition does indeed lead to more cheating, but also increases productivity. The between subject design of their experiment however makes it difficult to identify possible factors that determine differences in productivity and whether this is attributable to a certain type of subject.

[25] Fischbacher and Heusi (2013) and Mazar et al. (2008) report similar results. In Fischbacher and Heusi (2013) only a minority of subjects are fully dishonest, around a third are honest and a large group choose an intermediate level of dishonesty. The authors interpret this as a desire to maintain a positive opinion of one self. Similarly, in Mazar et al. (2008) the magnitude of dishonesty is small, with most people cheating at 20% of the possible magnitude. Mazar et al. (2006) provide an explanation for this behavior based on the relationship between external and internal reward mechanisms. They speculate that below a certain threshold (i.e., for small lies) the internal reward mechanism may not be activated, thus the propensity to lie only depends on external cost-benefit considerations; beyond such threshold dishonesty becomes noticeable, the internal reward mechanism is activated and the propensity to be dishonest is independent of external considerations; finally, the material advantage from lying becomes so substantial that an individual acts in conformity to the *homo economicus* paradigm.

[26] In a related study, Rigdon and D'Esterre (2012) find no significant difference in cheating between the piece rate and the competitive scheme. However, this is most likely a consequence of a design difference, since, unlike us, Rigdon and D'Esterre pay the same amount per reported matrix in the piece rate as in the tournament.

Finally, we analyze the effect of selection on dishonesty and output. It turns out that those subjects who lie *both* in the piece rate and in the tournament are more likely to select the tournament than subjects who are honest under both payment schemes. Interestingly, the behavior of those who *only* lie in the tournament does not differ from that of honest individuals. This finding seems to suggest that the population can be divided into three categories, according to the individual cost of lying. At one end we have honest people: they have a high cost of lying and choose to avoid competition. This is indeed rational since, at the other end, dishonest individuals, characterized by a low cost of lying, are more likely to enter competition. Finally, there is a third group defined by an intermediate cost of lying. They are only dishonest under the pressure of the contest but shy away from competition if given the option. Another way to look at this is to consider how subjects motivate their selection. While those who think the others will be dishonest choose the piece rate, subjects who plan to cheat select the tournament. As a consequence, when participation in the contest is endogenous we have a doubly perverse effect. On the one hand we have the *competition effect* as described above: because of the competitive pressure of the tournament subjects are more likely to lie. On the other hand, individuals with a higher propensity to lie are more likely to enter the tournament, causing an even more pronounced level of dishonesty. We call this the *selection effect*.

In spite of the higher dishonesty induced by competition, we find no overall difference in output between those who enter the tournament and those who select the piece rate scheme. This seemingly surprising result is explained by the behavior of a few honest individuals. When the tournament is exogenous, honest participants put little effort into the task, possibly because they are discouraged by the prospect of others lying. Consequently, in the selection stage, most of them choose the piece rate scheme. A handful of them, however, select the tournament and produce a higher output than when the contest was exogenously assigned. We can only speculate as to why these subjects decide to enter the tournament. They may be competitive individuals who, under a false consensus hypothesis, believe that a greater number of honest people will also choose this payment scheme. Whatever their reason, they behave very differently when they select the tournament than when this is imposed upon them. This makes sense: once an honest individual, who does not consider lying as an option, chooses the competitive scheme, the only rational thing for her to do is to work as hard as possible. Hence, despite the low productivity of dishonest individuals who select the tournament, total output is not different from the piece rate scheme. This is a very interesting and encouraging result. However, its robustness depends on the relative fraction of honest individuals who find competition to be sufficiently attractive.

5. Conclusion

Traditionally, introducing competition is argued to increase productivity. Evaluations about performance are however based on being able to accurately measure output, which is often difficult in the field. This provides an opportunity for individuals to be dishonest about their output. Using a real effort task and a within subject design we study individuals' likelihood to be dishonest when they are paid for their performance via different payment schemes. Further, in most situations participants choose to enter competitive professions rather than being randomly assigned. An important contribution of this paper is that it addresses this selection; that is, whether dishonest people are more likely to self-select into competitive environments.

Our results identify both a *competition* and a *selection* effect with respect to dishonesty. Dishonesty is higher when subjects are paid using a competitive tournament mechanism rather than a piece rate scheme. The performance measure or output is significantly lower in the tournament. This provides strong evidence that tournaments can actually hamper performance, when there is a possibility that individuals are dishonest. The selection stage shows that dishonest subjects have a significantly higher propensity to choose the tournament scheme. This selection effect strengthens the negative effect of tournaments and leads to high levels of dishonesty. Selection however does not influence output as a few honest people choose competition and work hard.

Our findings suggest that competition can in general lead to perverse effects. The incentive structure of contests and tournaments makes winning at all costs compelling for some people. This needs to be kept in mind when we design rules in several areas of society, such as winning in sports and politics, payments schemes in organizations and rewards for academic excellence. Institutional structures and management policies that are less easily manipulated by dishonesty and fraud would also increase interest and improve confidence in these important areas of society.

While generalizing from the controlled environment of the laboratory to the naturally occurring environment requires careful thought, we believe that our results can provide insights into the impact of payment schemes on dishonest behavior and selection of dishonest people into areas where there is more opportunity to be dishonest. Our experiment utilizes a design that can measure the degree and existence of dishonesty at an individual level in situations where there is no monitoring and performance is self-reported. While we acknowledge that this is a stylized design, the controlled environment helps us understand several aspects of dishonest behavior in the field. For example, in a recent paper, using a combination of laboratory and field methods, Hanna and Wang (2013) show that cheating in a laboratory task predicts fraudulent behavior by actual public servants in India. Dishonesty measures hence capture a meaningful propensity towards fraud and corruption.

In future research it would be useful to explore other performance measures such as an objective standard or an improvement on own past performance, and how these measures are influenced by dishonesty. A particularly interesting question would be the influence of these alternative measures on selection.

Table 1: Summary Statistics

Variable	**Description**	**Mean**	**Std Dev**	**Min**	**Max**
Lie2	Amount of lying in second task (first matrix task)	3.41	5.95	-1	20
Lie3	Amount of lying in third task	4.09	5.98	0	20
Lie4	Amount of lying in fourth task	4.66	6.34	-1	20
Matrices2	Number of matrices solved in second task (first matrix task)	3.48	2.26	0	9
Matrices3	Number of matrices solved in third task	4.16	2.43	0	9
Matrices4	Number of matrices solved in fourth task	4.30	2.20	0	9
Report2	Number of reported matrices in second task	6.89	5.66	0	20
Report3	Number of reported matrices in third task	8.25	5.93	0	20
Report4	Number of reported matrices in fourth task	8.97	6.10	0	20
PRFirst	Indicator if undertake Piece Rate task first	0.50	0.50	0	1
AlwaysLie	Indicator if lie in both Tasks 2 and 3	0.36	0.48	0	1
LiePRonly	Indicator if lie *only* in Piece Rate task	0.09	0.29	0	1
LieTonly	Indicator if lie *only* in Tournament task	0.27	0.45	0	1
EverLie	Indicator if lie in either Task 2 or 3 or both	0.72	0.45	0	1
NeverLie	Indicator if lie in neither Task 2 or 3	0.28	0.45	0	1
Lie in PR	Indicator if lie in Piece Rate Task	0.45	0.50	0	1
Lie in Tour	Indicator if lie in Tournament Task	0.63	0.48	0	1
TChoice	Indicator if choose Tournament in Task 4	0.35	0.48	0	1
Invested	Amount invested in Task 1	3.62	1.28	0	5
Male	Indicator if male	0.55	0.50	0	1
EconBus	Indicator if studying economics, business or commerce	0.62	0.49	0	1
LocalBorn	Indicator if born in Australia or New Zealand	0.46	0.50	0	1

Table 2: Average Dishonesty Measures by Payment Scheme and Task Order

	N	Size of Lie	Proportion Dishonest	Size of Lie if Dishonest	Percent Maximum Possible Lie
		Pooled	*Pooled*	*Pooled*	*Pooled*
Piece Rate	119	3.29	0.45	7.26	0.21
Tournament	119	4.22	0.63	6.71	0.24
p-value*		0.000	0.001		0.000
		PR First	*PR First*	*PR First*	*PR First*
Piece Rate	59	1.42	0.36	4.05	0.09
Tournament	59	3.05	0.58	5.29	0.18
p-value*		0.000	0.006		0.000
		T First	*T First*	*T First*	*T First*
Piece Rate	60	5.12	0.55	9.30	0.32
Tournament	60	5.37	0.68	7.88	0.30
p-value*		0.087	0.073		0.586

* p-values from a paired Wilcoxon signed rank test (i.e., within-subject). Because some subjects lied in only one task the test could not be applied to the magnitude conditional on lying.

Table 3: Average Dishonesty Measures in First Matrix Task Only

	N	Size of Lie	Proportion Dishonest	Size of Lie if Dishonest	Percent Maximum Possible Lie
		Mean	*Mean*	*Mean*	*Mean*
Piece Rate	59	1.42	0.36	4.05	0.09
Tournament	60	5.37	0.68	7.88	0.30
p-value*		0.000	0.000	0.043	0.000
		Error Adjusted	*Error Adjusted*	*Error Adjusted*	*Error Adjusted*
Piece Rate	59	1.37	0.29	4.76	0.08
Tournament	60	5.25	0.55	9.55	0.29
p-value*		0.000	0.004	0.016	0.000

* p-values from a Mann-Whitney rank sum rank test (i.e., between- subjects).

Table 4: Proportion of Subject Types

	Always Lie	*LiePRonly*	*LieTonly*	*Never Lie*
Raw data	36%	9%	27%	28%
Error adjusted	36%	3%	13%	48%

Table 5: Average Output by Payment Scheme and Order

	Pooled	*PR First*	*T First*	*First Task*
Piece Rate	4.47 (2.34)	4.24 (2.21)	4.70 (2.44)	4.24 (2.21)
Tournament	3.17 (2.23)	3.61 (2.31)	2.73 (2.07)	2.73 (2.07)
p-value*	0.000	0.076	0.000	0.000

The figures in the table are average (standard deviation). * From two-sided non-parametric tests; the first three columns report paired Wilcoxon signed-rank tests and the last column reports the Mann-Whitney rank sum test.

Table 6: Average Output in the First Matrix Task Disaggregated by Subject Type

Raw Measure	*Always Lie*	*LiePRonly*	*LieTonly*	*Never Lie*
Piece Rate	3.38	3.20	5.22	4.30
Tournament	2.74	3.67	3.00	2.00
p-value*	0.208	0.848	0.015	0.003
Error Adjusted	*Always Lie*	*LiePRonly*	*LieTonly*	*Never Lie*
Piece Rate	3.38	4.00	6.20	4.06
Tournament	2.74	5.50	2.67	2.52
p-value*	0.208	0.221	0.009	0.008

* p-values from Mann-Whitney rank sum tests.

Table 7: Probit Regressions of Entry Decisions, Marginal Effects Reported

Dependent variable: TChoice

	(1)	(2)	(3)	(4)
AlwaysLie	0.099*** (0.042)			
EverLie		0.161** (0.080)		
Lie in PR			0.092*** (0.034)	
Lie in Tour				0.139* (0.077)
PRFirst	-0.038 (0.114)	-0.030 (0.103)	-0.036 (0.111)	-0.039 (0.109)
Invested	0.042 (0.031)	0.039 (0.033)	0.044 (0.031)	0.036 (0.0317)
Male	0.256* (0.133)	0.254* (0.133)	0.252* (0.130)	0.260* (0.139)
EconBus	0.128 (0.085)	0.141 (0.086)	0.125 (0.088)	0.143* (0.084)
LocalBorn	-0.043 (0.048)	-0.065 (0.055)	-0.045 (0.049)	-0.057 (0.050)
Observations	118	118	118	118
Log likelihood	-69.102	-68.352	-69.143	-68.541

Standard errors in parentheses; clustered over session. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 8: Error-Corrected Dishonesty Coefficients in Selection Regression

Dishonesty Measure	(1) *AlwaysLie*	(2) *EverLie*	(3) *Lie in PR*	(4) *Lie in Tour*
Estimated marginal effect	0.099***	0.016	0.092**	0.020
	(0.042)	(0.045)	(0.046)	(0.059)

Dependent variable and other regressors are the same as in Table 7. Standard errors in parentheses; clustered over session. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 9: Reasons Stated for Task 4 Selection Choice

			Percent Choosing Tournament[b]		
Stated Reason for Task 4 Choice	**Average percent stating reason[a]**	**Cohen's Kappa (reliability)**	**of those who gave this reason**	**of those NOT giving this reason**	**p-value[c]**
Expect others to be dishonest	9%	0.843	0	38%	0.017
I plan to be dishonest	9%	0.843	80%	30%	0.033
Confidence in ability in matrix task	11%	0.955	92%	28%	0.000
Perception of low ability or lack of confidence in ability in matrix task	15%	0.766	6%	39%	0.008
Preference for lower risk	55%	0.865	2%	72%	0.000
Chance of higher return.	25%	0.843	100%	16%	0.000
Preference for competing	7%	0.594	100%	30%	0.002
Dislike of competing	6%	0.325	0	39%	0.205

[a] Percent is averaged across the two coders where disagreement occurs.
[b] Illustrative values from one coder only.
[c] p-value from a Mann-Whitney test comparing whether the proportion selecting the Tournament differs according to whether or not the stated reason was given. Reports the highest p-value across the two coders.

Table 10: Tobit Regressions on the Size of the Lie in Task 4

Dependent Variable: Size of Lie in Task 4 (double censored 0-20)

	(1)	(2)
TChoice	2.218***	2.766***
	(0.477)	(0.648)
PRFirst	-3.436*	-3.526*
	(2.000)	(1.858)
Invested		-0.236
		(0.664)
Male		-0.988
		(1.125)
EconBus		-1.045
		(1.162)
LocalBorn		1.747*
		(1.034)
Constant	3.060*	4.123***
	(1.827)	(1.538)
Observations	118	117
Log Pseudolikelihood	-291.448	-287.504
Pseudo R2	0.008	0.011

Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 11: Tobit Regression on the Output in Task 4

Dependent Variable: Number of Matrices Solved in Task 4 (double censored 0-10)

VARIABLES	(1)
TChoice	-0.556
	(0.366)
Matrices2	0.453***
	(0.066)
*NeverLie*TChoice*	1.354*
	(0.689)
*LieTonly*TChoice*	1.291**
	(0.572)
Invested	0.097
	(0.109)
Male	-0.095
	(0.630)
EconBus	0.307
	(0.328)
LocalBorn	-0.549
	(0.487)
Constant	2.410***
	(0.378)
Observations	118
Log PseudoLikelihood	-243.151
Pseudo R^2	0.066
Joint Test of Significance: *NeverLie*TChoice & LieTonly*TChoice*	
F(2, 110)	4.01
Prob > F	0.0208**
Linear combination of coefficients: *NeverLie*TChoice & LieTonly*TChoice*	
Coefficient	2.645
(Std. Error)	(0.946)
P-value	0.006***

Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Figure 1: Sample Matrix

3.91	0.82	3.75
1.11	1.69	7.94
3.28	2.52	6.25
9.81	6.09	2.46

Figure 2: Distribution of Lying in the First Matrix Task by Payment Scheme

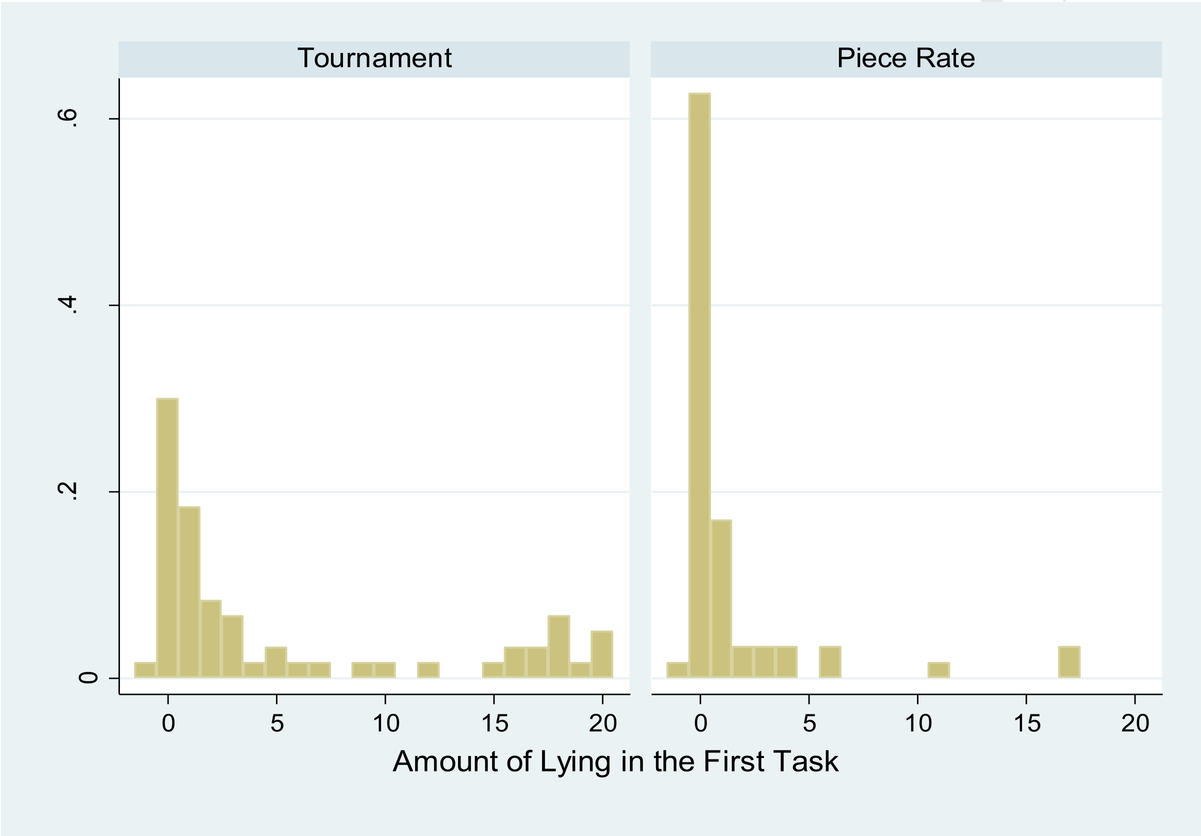
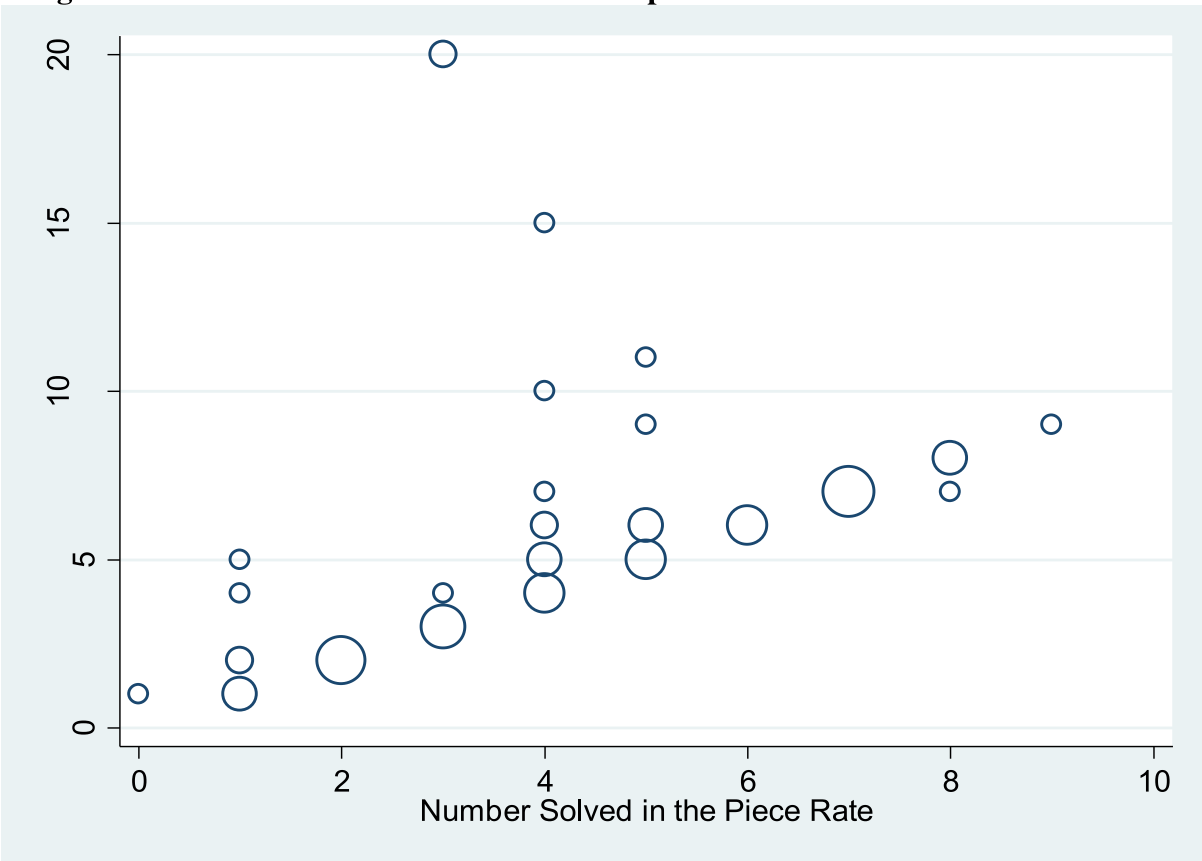


Figure 3: Plot of Matrices Solved versus Reported Matrices in First Matrix Task



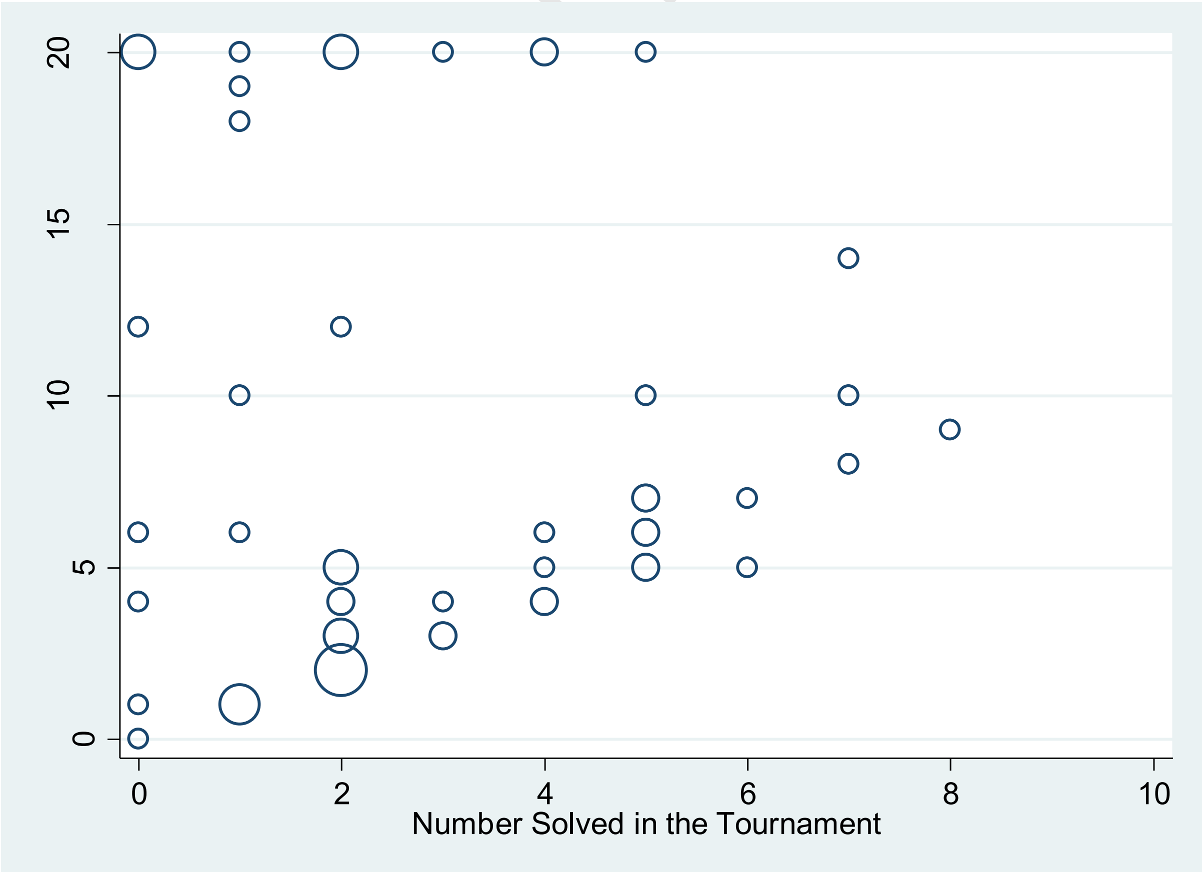


Figure 4: Distribution of Output in First Matrix Task by Payment Scheme

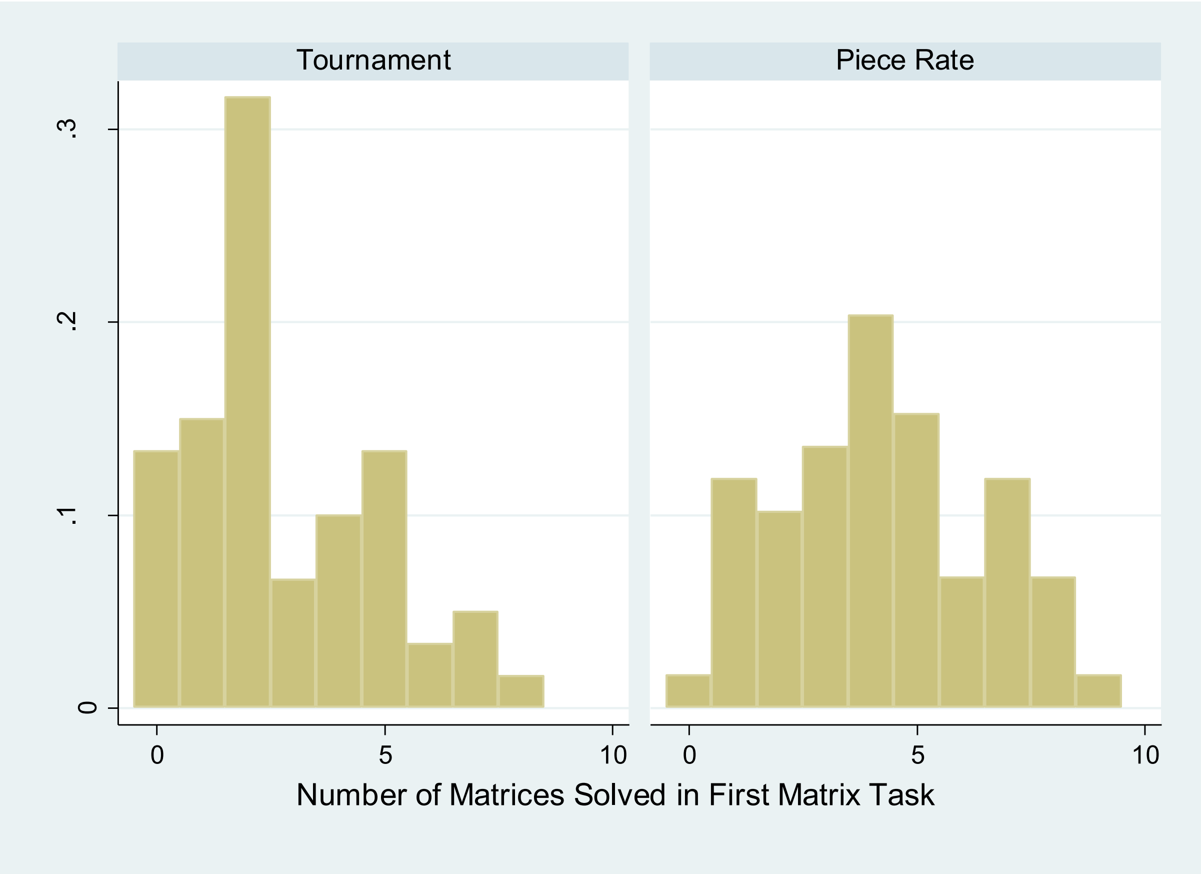
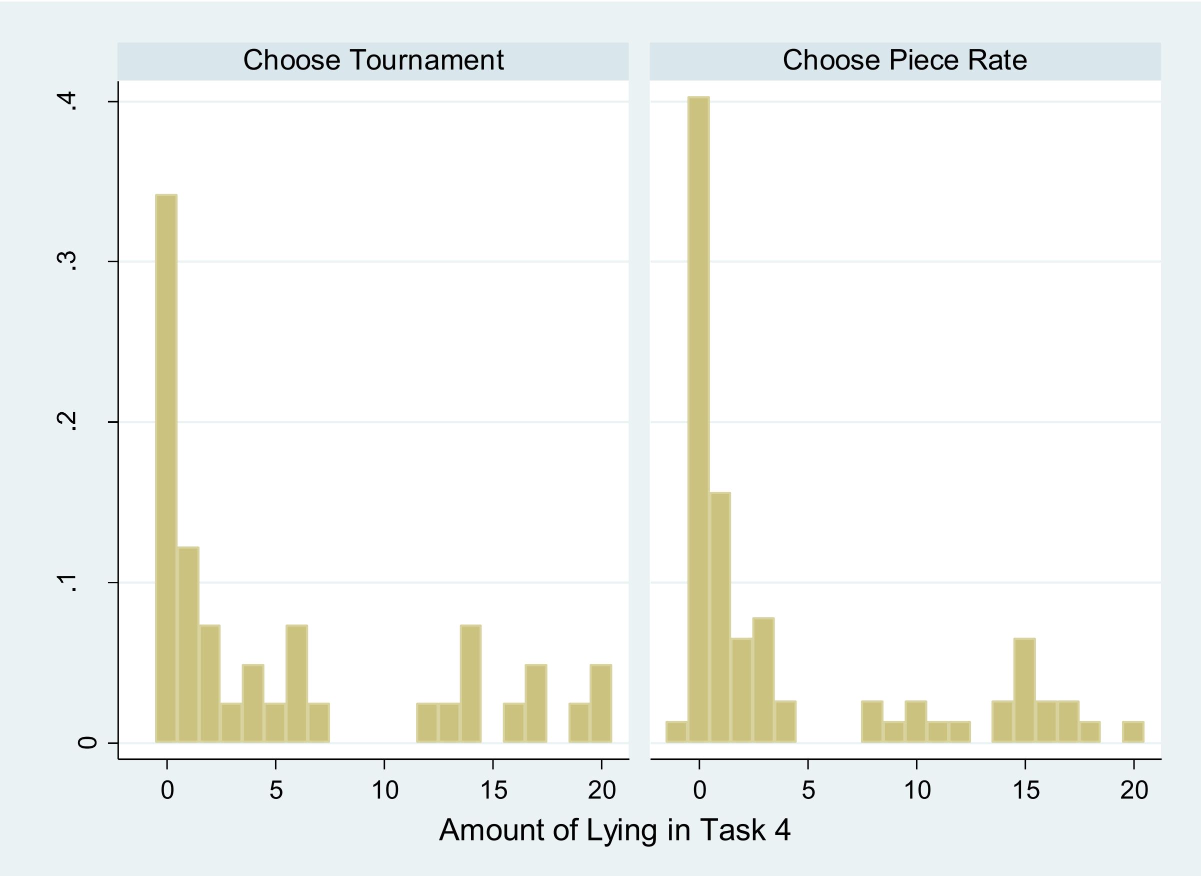


Figure 5: Distribution of Lying in Task 4 by Selection Choice



References

Abeler, Johannes, Anke Becker, and Armin Falk. 2014. "Representative Evidence on Lying Costs." *Journal of Public Economics*, doi: 10.1016/j.jpubeco.2014.01.005.

Battigalli, Pierpaolo, Gary Charness, and Martin Dufwenmber. 2013. "Deception: The role of Guilt." *Journal of Economic Behavior and Organization*, 99: 227-32.

Belot, Michele, and Marina Schröder. 2013. "Sloppy work, lies and theft: A novel experimental design to study counterproductive behaviour." *Journal of Economic Behavior and Organization*, 99: 233-38.

Cadsby, C. Bram, Fei Song, and Francis Tapon. 2010. "Are you paying your employees to cheat? An experimental investigation." *The B.E. Journal of Economic Analysis and Policy Contributions*, 10(1): 1-32.

Cappelen, Alexander, Erik Sorensen, and Bertil Tungodden. 2013. "When do we lie?" *Journal of Economic Behavior and Organization*, 99: 258-65.

Carpenter, Jeffrey, Peter H. Matthews, and John Schirm. 2010. "Tournaments and office politics: Evidence from a real effort experiment." *American Economic Review*, 100(1): 504-17.

Cason, Timothy N., William A. Masters, and Roman Sheremeta. 2010. "Entry into winner-take-all and proportional-prize contests: An experimental study." *Journal of Public Economics*, 94: 604-11.

Cason, Timothy N., and Vai-Lam Mui. 2014. "Rich communication, social preferences, and coordinated resistance against divide-and-conquer: A laboratory investigation".

Cohen, Jacob. 1960. "A coefficient of agreement for nominal scales." *Educational and Psychological Measurement* 20: 37-46.

Conrads, Julian, Bernd Irlenbush, Rainer M. Rilke, and Gari Walkovitz. 2013. "Lying and team incentives." *Journal of Economic Psychology*, forthcoming.

Curry, Philip A., and Steeve Mongrain. 2009. "Deterrence in rank-order tournaments." *Review of Law and Economics*, 5(1): 723-40.

Danilov, Anastasia, Torsten Bienmann, Thoirn Kring, Dirk Sliwka. 2013. "The dark side of team incentive: Experimental evidence on advice quality from financial service professionals." *Journal of Economic Behavior and Organization*, 96, 266-72.

Dreber, Anna, and Magnus Johannenson. 2008. "Gender differences in deception." *Economics Letters*, 99: 197-99.

Ehrenberg, Ronald G., and Michael L. Bognanno. 1990. "Do tournaments have incentive effects?" *Journal of Political Economy*, 98(6): 1307-24.

Ellingsen, Tore, and Magnus Johannenson. 2004. "Promises, threats and fairness." *Economic Journa*l, 114(495): 397-420.

Fischbacher, Urs, and Franziska Heusi. 2013. "Lies in disguise – An experimental study on cheating." *Journal of the European Economic Association*, 11(3): 525-47.

Friesen, Lana, and Lata Gangadharan. 2012. "Individual level evidence of dishonesty and the gender effect." *Economics Letters*, 117: 624-26.

Friesen, Lana, and Lata Gangadharan. 2013. "Designing self-reporting regimes to encourage truth telling: An experimental study." *Journal of Economic Behavior and Organization*, 94: 90-102.

Gill, David, Victoria Prowse and Michael Vlassopoulos. 2013. "Cheating in the workplace: An experimental study of the impact of bonuses and productivity." *Journal of Economic Behavior and Organization*, 96, 120-34.

Gino, Francesca, Shahar Ayal, and Dan Ariely. 2013. "Self-serving altruism? The lure of unethical actions that benefit others." *Journal of Economic Behavior and Organization*, 96, 258-92.

Gneezy, Uri, Muriel Niederle and Aldo Rustichini. 2003. "Performance in Competitive Environments: Gender Differences." *Quarterly Journal of Economics*, 118(3): 1049-74.

Gneezy, Uri. 2005. "Deception: The role of consequences." *American Economic Review*, 95(1): 384-94.

Gneezy, Uri, and Jan Potters. 1997. "An Experiment on Risk Taking and Evaluation Periods". *The Quarterly Journal of Economics,* 112: 631-646.

Gneezy, Uri, Kenneth L. Leonard, and John A. List. 2009. "Gender differences in competition: Evidence from a matrilineal and patriarchal society." *Econometrica*, 77(5): 1637-64.

Gneezy, Uri, Bettina Rockenbach, and Marta Serra-Garcia. 2013. Measuring lying aversion. *Journal of Economic Behavior and Organization*, 96, 293-300.

Greiner, Ben. 2004. "The online recruitment system ORSEE – A guide for the organization of experiments in economics." *Working Paper Series in Economics 11*, University of Cologne, Department of Economics.

Hanna, Rema, and Shin-Yi Wang. 2013. "Dishonesty and selection into public service." *NBER Working Paper 19649*.

Harbring, Christine. 2010. **"On the effect of incentive schemes on trust and trustworthiness."** *Journal of Institutional and Theoretical Economics,* 166(4): 690-714.

Harbring, Christine, and Bernd Irlenbush. 2005. "Incentives in tournaments with endogenous prize selection." *Journal of Institutional and Theoretical Economics*, 161: 636-63.

Harbring, Christine, and Bernd Irlenbush. 2008. "How many winners is good to have? On tournaments with sabotage." *Journal of Economic Behavior and Organization*, 65: 682-702.

Harbring, Christine, and Bernd Irlenbush. 2013. "Sabotage in tournaments – Evidence from the laboratory." *Management Science*, forthcoming.

Innes, Robert, and Arnab Mitra. 2013. "Is dishonesty contagious?" *Economic Inquiry*, 51(1): 722-34.

Konrad, Kai. 2000. "Sabotage in rent-seeking contests." *Journal of Law, Economics and Organization*, 16(1): 155-65.

Krippendorff, Klaus. 2013. *Content Analysis: An Introduction to Its Methodology*, Third Edition, Sage Publications.

Landis, J. Richard and Gary Koch. 1977. "An Application of Hierarchical Kappa-Type Statistics in the Assessment of Majority Agreement among Multiple Observers," *Biometrics,* 33: 363-374.

Lazear, Edward P., and Sherwin Rosen. 1981. "Rank-order tournaments as optimum labor contracts." *Journal of Political Economy*, 89(5): 841-64.

Lundquist, Tobias, Tore Ellingsen, Eric Gribbe, and Magnus Johannenson. 2009. "The aversion to lying." *Journal of Economic Behavior and Organization*, 70: 81-92.

Maas, Victor, and Marcel van Rinsum. 2013. "How control system design influences performance misreporting." *Journal of Accounting Research*, 51(5): 1159-86.

Mazar, Nina, and Dan Ariely. 2006. "Dishonesty in everyday life and its policy implications." *Journal of Public Policy and Marketing*, 25(1): 117-26.

Mazar, Nina, On Amir, and Dan Ariely. 2008. "The dishonesty of honest people: A theory of Self-concept maintenance." *Journal of Marketing Research*, 15: 633-44.

Niederle, Muriel, and Lise Vesterlund. 2007. "Do women shy away from competition? Do men compete too much?" *Quarterly Journal of Economics*, 122: 1067-101.

Niederle, Muriel, Carmit Segal, and Lise Vesterlund. 2013. "How costly is diversity? Affirmative action in light of gender differences in competitiveness." *Management Science*, 59(1): 1-16.

Ploner, Matteo, and Tobias Regner. 2013. "Self-image and moral balancing: An experimental analysis." *Journal of Economic Behavior and Organization*, 96, 374-83.

Rigdon, Mary, and Alex D'Esterre. 2012. "The effects of competition on the nature of cheating behavior." Unpublished manuscript.

Schweitzer, Maurice E., Lisa Ordóñez, and Bambi Douma. 2004. "Goal setting as a motivator of unethical behavior" *Academy of Management Journal*, 47(3): 422-32.

Schwieren, Christiane, and Doris Weichselbaumer. 2010. "Does competition enhance performance or cheating? A laboratory experiment." *Journal of Economic Psychology*, 31: 241-53.

Appendix: Experimental Instructions[27]

Thank you for agreeing to take part in this study. Please read the instructions we give you today carefully. A clear understanding of the instructions will help you make better decisions and increase your earnings. All the money you earn is yours to keep and will be paid to you in private and in cash immediately after the experiment. If you decide to leave early, you will forgo your earnings, except for a $5 participation fee.

In the experiment today you will be asked to complete <u>four</u> different tasks. None of these will take more than 5 minutes. At the end of the experiment you will receive $5 for having completed the experiment. We will also pay you at the end of the experiment for task 1. In addition, we will randomly select one task from tasks 2-4 and pay you for that task. Once you have completed all the tasks we will determine which of the tasks 2-4 count for payment by rolling a standard die (if the number 1 or 5 or 6 is rolled the die will be rolled again). The method we use to determine your earnings varies across tasks. Before each task we will describe in detail how your payment is determined.

Your total earnings from the experiment are: Your $5 for completing the experiment + Payment for task 1 + Payment for the randomly selected task from the tasks 2 to 4.

Please turn your phone off or to silent mode now and place it on the floor, along with any other materials you have brought into the room. If you have any questions, please raise your hand. Please do not consult with other participants in the room.

All decisions that you make today are recorded only by an anonymous subject number and will only be used for research purposes. Your decisions will remain completely anonymous.

Are there any questions before we begin?

[27] These instructions are for the sessions where the Piece Rate task came first. In the sessions with Tournament first, the order of Tasks 2 and 3 was reversed.

Task 1

We are about to begin the 1st task.

In this task, you will be given $5. (No money will be given at this point. All actual payments will be made at the end of the experiment). You have the opportunity to invest a portion of this amount (between $0 and $5).

<u>The investment:</u>
There is an equal chance that the investment will fail or succeed. If the investment fails, you lose the amount you invested. If the investment succeeds, you receive 3 times the amount invested.

<u>How do we determine the outcome of the investment:</u>
After you have chosen how much you wish to invest, you will toss a coin to determine whether your investment succeeds or not. If the coin comes up heads, you win three times the amount you chose to invest. If it comes up tails, you lose the amount invested. You will toss the coin when you come to collect your payment at the end of the experiment.

Here are some examples:

1. You choose to invest nothing. You will get $5 for sure.
2. You choose to invest all of the $5. Then if the coin comes up heads, you get $15. If the coin comes up tails, you get $0.
3. You choose to invest $3. Then if the coin comes up heads, you get $11 ($3 x 3 + $2 = $11). If the coin comes up tails, you get $2.

Please fill in the amount that you would like to invest in the decision form.

Do you have any questions? If you are ready, we will proceed.

Form for Recording Decision for Task #1

1. Amount that I wish to invest: ________________________

2. Reason for this decision:

When you are done, fold the sheet in half, raise your hand and we will collect the form from you.

Task 2: Piece Rate Task

<u>Please do not open the envelope. Wait for experimenter instructions!</u>

On your desk is an envelope labelled **<u>TASK 2</u>** in which you will find a sheet with 20 matrixes as the one below:

Example

3.91	0.82	3.75
1.11	1.69	7.94
3.28	2.52	6.25
9.81	6.09	2.46

In each matrix you should look for a unique set of numbers that **sum up exactly to 10**. In some matrixes you may not have a solution.

A matrix is solved correctly by finding two numbers that sum up exactly to 10.

When you find a set, circle the numbers, and mark the corresponding 'Got It' Box below, as in the example below.

Example

3.91	0.82	3.75
1.11	1.69	7.94
3.28	2.52	6.25
9.81	6.09	2.46

Got it ☑

<u>How will you be paid?</u>

For each matrix you report solving correctly, you will receive **$1.00** (if this task is the one randomly selected for payment).

You have **5 minutes** for this task.

Once 5 minutes are up you have to do the following:

Matrix Stage

1. Count the number of correctly solved matrixes and take note of this number.

2. Fold your matrix sheet and place it in the **Orange Matrix Box** located near the entrance to the lab. This box will not be opened until after everyone has left the lab today.

Reporting Stage

1. Write down the number of solved matrixes you wish to report on the collection slip. Only this number will be used to determine your earnings in this task (if selected for payment).

2. Take your collection slip and place it in the **Record Box** for this task located on the front desk. If this task is selected for payment (from tasks 2-4) then we will open the Record Box to determine your earnings; otherwise the Record Box for this task will not be opened until everyone has left the lab today.

Collection Slip for Task 2

ID#

I got _______ Boxes, which translates to $ _______ (=$1.00 per Box)

Task 3: Tournament Task

<u>Please do not open the envelope. Wait for experimenter instructions!</u>
On your desk is an envelope labelled **<u>TASK 3</u>** in which you will find a sheet with 20 matrixes as the one below:

Example

3.91	0.82	3.75
1.11	1.69	7.94
3.28	2.52	6.25
9.81	6.09	2.46

In each matrix you should look for a unique set of numbers that **sum up exactly to 10**. In some matrixes you may not have a solution.

When you find a set, circle the numbers, and mark the corresponding 'Got It' Box below, as in the example below:

Example

(3.91)	0.82	3.75
1.11	1.69	7.94
3.28	2.52	6.25
9.81	(6.09)	2.46

Got it ☑

<u>How will you be paid?</u>

In this task you will be randomly paired with another participant in the experiment, who is in the lab today, and your payment will depend on the number of matrices you report solving relative to that of the person whom you are paired with.

- If you report solving <u>more</u> matrixes than the person you are paired with, you will receive $2 per reported matrix (if this task is the one randomly selected for payment).
- If you report solving <u>less</u> than the person you are paired with, you will receive $0.
- If you report solving exactly the same amount as the person you are paired with, we will toss a coin to determine whether you receive $2 per reported matrix, or receive $0.

You have **5 minutes** for this task.[28]

[28] Instructions regarding the two stages are then repeated here.

Collection Slip for Task 3

ID#

I got _______ Boxes, which translates to $ _______ ($2.00 per Box if I win or $0 if I lose).

Task 4: Payment Method Choice Task

<u>Please do not open the package. Wait for experimenter instructions!</u>
On your desk is an envelope labelled **<u>TASK 4</u>** in which you will find another sheet with 20 matrixes to solve, as in the previous tasks.

In each matrix you should look for a unique set of numbers that **sum up exactly to 10**. In some matrixes you may not have a solution.

When you find a set, circle the numbers, and mark the corresponding 'Got It' Box below, as in the example below:

Example

3.91	0.82	3.75
1.11	1.69	7.94
3.28	2.52	6.25
9.81	6.09	2.46

Got it ☑

<u>Before</u> you solve the matrices you need to choose which payment scheme you want applied to the number of matrixes that you report solving in this task. You can either choose to be paid according to the payment mechanism in the piece rate task or according to the payment mechanism in the tournament task.

If this task is the one randomly selected for payment, then your earnings for this task are determined as follows:

- if you choose the payment mechanism in the piece rate task, then for each matrix you report solving, you will receive **$1.00.**

- if you choose the payment mechanism in the tournament task, you will be randomly paired with another participant who has chosen this payment scheme in the experiment today, and your payment will depend on the number of matrices you report solving relative to that of the person whom you are paired with. If you report solving <u>more</u> matrixes correctly than the person you are paired with, you will receive $**2.00** per correct answer. If you report solving <u>less</u> than the person you are paired with, you will receive $**0**. If you report solving exactly the same amount as the person you are paired with, we will toss a coin to determine whether you receive $2 per correct answer, or you receive $0.

You have **5 minutes** for this task.

Form for Task 4

Q1. Which payment scheme would you like applied (check the appropriate box)

☐ Option 1: The payment mechanism from the piece rate task

☐ Option 2: The payment mechanism from the tournament task

Q2. Please briefly describe how you reached your decision of Option 1 or Option 2.

Q3. If you chose Option 1, did your decision depend on the payment rate under Option 2? If so, what payment rate would have convinced you to choose Option 2?

<u>Collection Slip for Task 4</u>

ID#

I got _______ Boxes

Payment

We will now determine which of the tasks 2-4 to pay for. We will roll a standard die (if the number 1 or 5 or 6 is rolled the die will be rolled again); you will be paid for the task that corresponds to the number that comes up.

If payment is for either of the paired tasks (i.e. the tournament task - task 3 or the payment method choice task – task 4), then we will further determine the number of matrixes solved by your randomly paired partner and pay you accordingly. To do this we will ask you to pick one card from the appropriate box on the front desk. The participant ID number on that card will be matched with you.