# Twice Losers: How the shadow of cheating affects tax behaviors and norms

**FINAL REPORT**

Michela Boldrini *

July 2019

## 1 Introduction

This report summarizes the activities and the results of the research projects *Twice Losers: How the shadow of cheating affects tax behaviors and norms*, investigator: Michela Boldrini. The project benefited from the generous financial support of the IFREE Small Grants program.
The objective of this project is to investigate if, and how, the origins of inequality in a society can impact both individuals' willingness to pay taxes and perceptions on the social acceptability of evasion in a lab-controlled environment.
More specifically, we focus on whether tax evasion prevalence and tolerance are higher in a society where income inequality may result from cheating rather than from differences in effort or ability. The 'shadow of cheating' introduces an alteration in the degree of fairness of the process which generates incomes, opening up to the suspicion that the top positions in the pre-tax income distribution might have been reached not only through honest means, but also by taking advantage of available cheating opportunities at the expenses of others.
Working on previous literature contributions, which highlighted the importance of some income inequality features in shaping individuals' redistribution preferences and lying behavior (Bortolotti et al. [2017]; Cappelen et al. [2018]; Kajackite [2018]), we design a novel experimental design to contribute to the study of how some features of the income inequality in a society can shape (i) individuals' behavior when they are personally involved in a redistribution action and not only consulted as un-involved third-party actors (ii) individuals' judgements on the social acceptability of redistribution actions.

Our experimental design, based on a two-prong approach, allows to manipulate subjects' perceptions on the presence of the *shadow of cheating* on the income-generation process and to observe - relying on two different non-overlapping sets of participants - both subjects' actual contribution behavior in a context where complete anonymity is guaranteed, and subjects' evaluations on the social acceptability of under-contribution choices.
Our preliminary results could be summarized as follows. First, our manipulation is effective in altering subjects' perception of cheating intensity across treatments. However, we do not observe a significant difference in subjects' aggregate contribution behavior across treatments in Experiment 1. Second, the general pattern of social acceptability ratings elicited in Experiment 2 is the same across treatments. Third, in all scenarios acceptability norms result to be highly sensitive to the size of evasion and this result is stable even after we control for individual beliefs on the actual level of contribution for each level of earnings. Lastly, norms appear to be a relevant driver of subjects' contribution choices, suggesting that subjects derive some form of utility from abiding the existing (and commonly recognized) injunctive social norms.

*Department of Economics, University of Bologna: P.zza Scaravilli 2, 40126 - Bologna (Italy)
E-mail address: michela.boldrini2@unibo.it

## 2 Motivation and Theoretical Framework

Income inequality and tax evasion represent two major issues for our modern societies but the relationship between these two phenomena is under-explored over a series of dimensions.
Previous literature has already pointed out the detrimental distributional effects of the distortions induced by tax evasion (Slemrod and Johns [2010]; Matsaganis and Flevotomou [2010]) but little is known, on the other side, on whether and how inequality - and in particular its origin - can have an impact on tax evasion. The objective of this work is to contribute to the literature studying tax evasion's behavioral determinants, which represent an extremely relevant channel from a policy-perspective especially in contexts where tax evasion is sizeable (as it is the case in Italy for example, where the amount of yearly evaded taxes has been recently estimated to be around 124,5-132,1 billions Euros [1]) and legal enforcement measures lack efficacy.

Literature on tax compliance has only marginally focused on the effects of fairness concerns for inequality over compliance (Spicer and Becker [1980]; Bordignon [1993]). The economic literature on the effects of fairness concerns for inequality on redistributive preferences and lying costs, instead, is abundant and related experimental applications are growing fast (Alesina and Angeletos [2005]; Cappelen et al. [2007]; Durante et al. [2014]; Bortolotti et al. [2017]; Kajackite [2018]). In particular, recent experimental studies (Bortolotti et al. [2017]; Cappelen et al. [2018]) suggested that the suspicion that some individuals in a society got their incomes by taking advantage of cheating opportunities, rather than solely by honest means, can have a substantial impact on individuals' fairness perceptions and redistribution preferences.
The objective of this study is to expand this branch of the experimental literature, by investigating whether the presence of this suspicion – namely, of the *shadow of cheating* – can also have an impact on tax compliance norms and behavior, in a situation which mimics more closely the reality, where the outcome of the redistributive process depends on the actions of the same actors involved in the income production phase and not by external un-involved third-party spectators.

The motivation to focus on this question comes from the observation that in the real world, and in some countries more than others, the suspicion that some individuals in the society got their income at least partially my means of cheating is far from being rare, especially with respect to individuals who hold very top positions in the income distribution.
As reported by the World Values Survey data [2] people tend to attach a high weight to factors such as luck and connections when asked to report whether they believe that in the long-run life success depends on hard work and effort rather than by luck and connections. Overall, 31% of respondents report that luck and connections outweigh effort as long-run life success's driver, but there's a good degree of heterogeneity across countries and in Italy, for example, this share is equal to approximately 46%.
The same scenario is reported by the International Social Survey Programme data [3] where individuals are asked how important they would rate knowing the right people and having political connections to get ahead in life. Overall, 55% and 25.5% of the respondents, respectively, claimed that knowing the right people and having connections is either a very important or essential factor: still there's a good degree of heterogeneity across countries and Italy, again, is positioned above the overall average in both cases with amost 59% and 41% of the respondents recognizing the two channels as at least very important to succeed in life.

[1] http://www.senato.it/application/xmanager/projects/leg17/attachments/documento/...pdf

[2] WVS Wave 5: 2005-2009 - Countries included in the survey: Andorra, Argentina, Australia, Brazil, Bulgaria, Canada, Chile, China, Taiwan, Colombia, Cyprus, Ethiopia, Finland, France, Georgia, Germany, Ghana, Hungary, India, Indonesia, Iran, Italy, Japan, Jordan, South Korea, Malaysia, Mali, Mexico, Moldova, Morocco, Netherlands, New Zealand, Norway, Peru, Poland, Romania, Russia, Rwanda, VietNam, Slovenia, South Africa, Spain, Sweden, Switzerland, Thailand, Trinidad and Tobago, Turkey, Ukraine, Egypt, Great Britain, United States, Burkina Faso, Uruguay, Serbia and Montenegro, and Zambia.

[3] International Social Survey Programme "Social Inequality" data: 2009 - Countries included in the survey: Argentina, Australia, Austria, Bulgaria, Chile, China, Croatia, Cyprus, Czech Republic, Denmark, East-Germany, Estonia, Finland, Flanders, France, Great Britain, Hungary, Iceland, Israel, Italy, Japan, Latvia, Lithuania, New Zealand, Norway, Philippines, Poland, Portugal, Russia, Slovakia, Slovenia, South Africa, South Korea, Spain, Sweden, Switzerland, Taiwan, Turkey, Ukraine, United States, Venezuela, West-Germany.

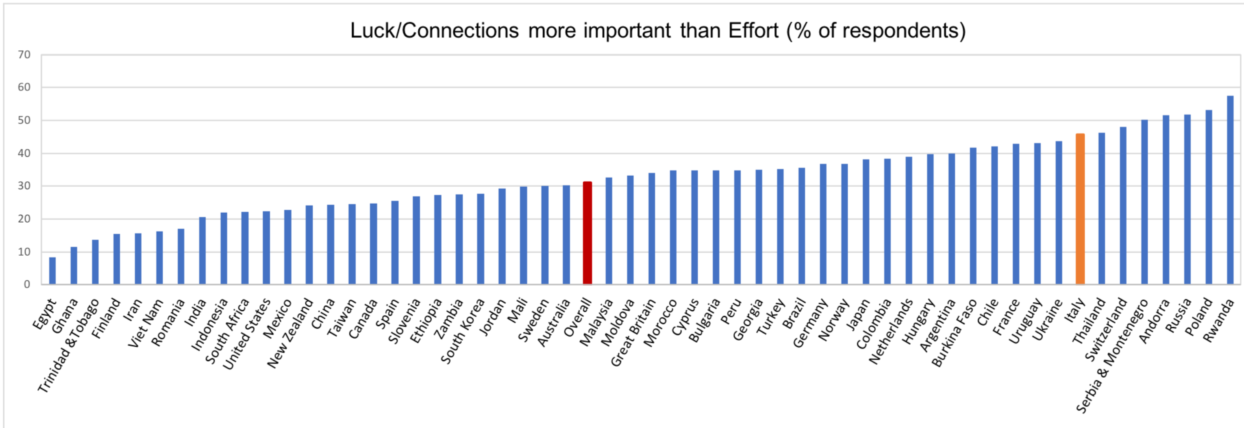



Figure 1: World Values Survey data (2005-2009)

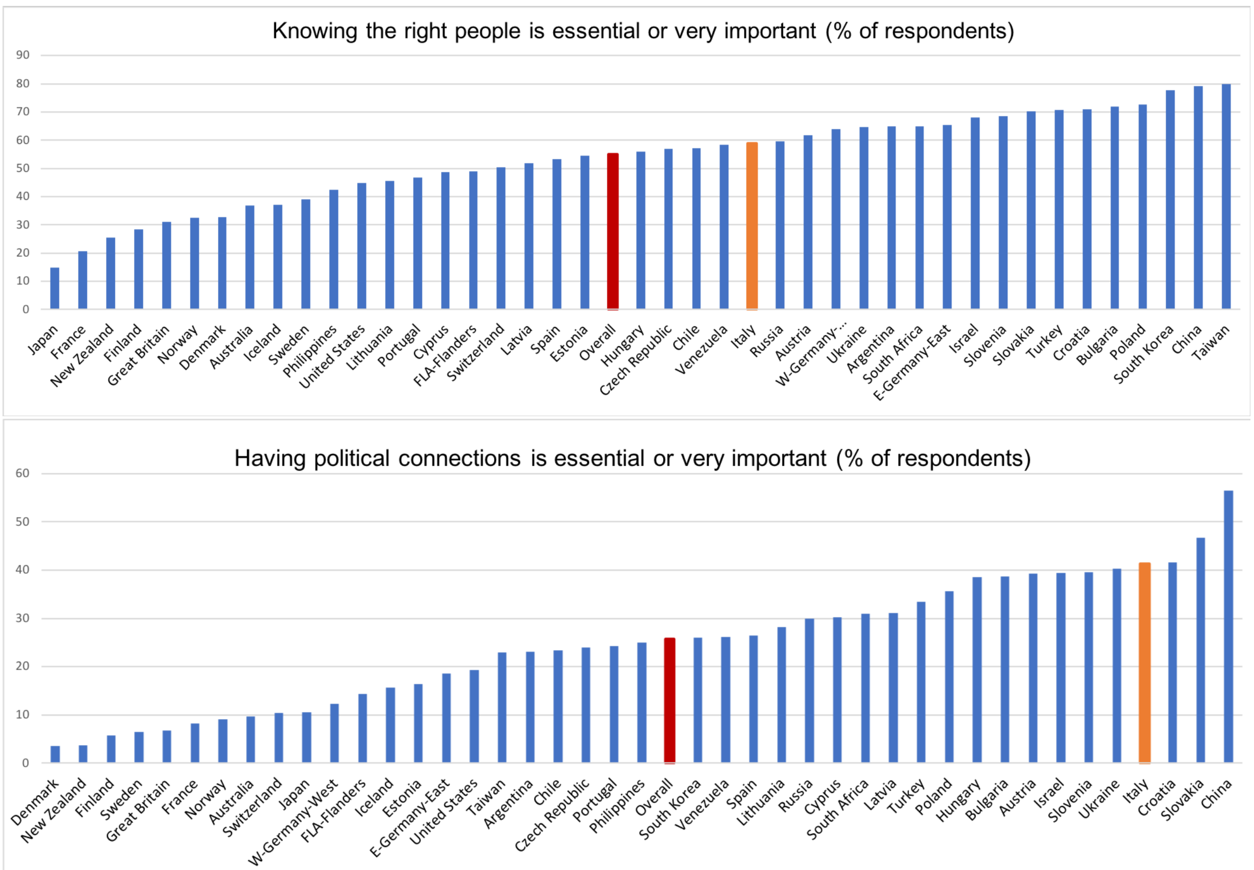



Figure 2: International Social Survey Programme data (2009)

From the World Values Survey data, we can also observe a positive and significant correlation between the perception on how justifiable is tax evasion in a country, and the average weight attached to luck and connections, as opposed to own effort, as the main determinant of life success. This evidence, however, has two main limitations: first, it does not distinguish the role of luck from the role of connections, although these two factor are likely to bring about different considerations in terms of fairness; second, this analysis is able to show only a simple correlation given that it is not possible to draw any casual statement on the impact of the role of luck and connections on tax evasion acceptability perceptions.

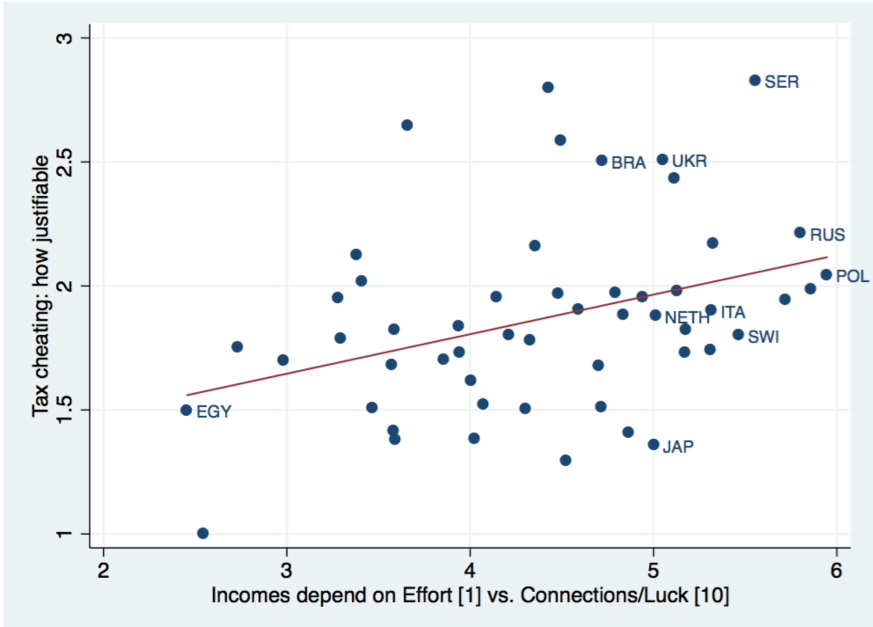



Figure 3: World Values Survey data (2005-2009)

Our experimental approach aims to overcome these limitations to study in isolation what is the impact of cheating, which mimics the opportunity to exploit personal connections and other non-regular or transparent channels to maximize own benefits in the real life, on tax evasion.
We do not aim to directly test a specific theoretical model, but rather to advance experimental knowledge on tax evasion behavioral mechanisms, incorporating new insights coming from the recent experimental evidence on the impact of fairness concerns on redistribution preferences and lying costs cited above. Most of the theoretical and experimental literature on tax compliance, hinging on a mainly crime-economics perspective, focused on the impact of factors such as the probability of detection, the reciprocity of the fiscal exchange, the relevance of efficiency concerns, or the impact of peers' influence on tax compliance behavior (Allingham and Sandmo [1972]; Alm et al. [1992]; Fortin et al. [2007]), but only marginally analyzed the effects of fairness concerns on compliance decisions (Spicer and Becker [1980]; Bordignon [1993]; Fortin et al. [2007]). In these latter works, fairness is defined in terms of horizontal equity and it emerges that individuals who believe to be treated unfairly by the tax system are more likely to evade taxes in order to restore the equity of tax burden they bear. Given that individuals' preferences for redistribution have been shown to be sensitive to how incomes – and related inequalities – are generated, we want to test experimentally whether the presence of the 'shadow of cheating' in the income generating process, altering individuals' preferences for redistribution and thus their perception on the degree of fairness of the tax burden, can also have an impact on their decision to engage in tax evasion and on norms for tax evasion tolerance.

Based on available experimental evidence on how the source of inequality can influence individual redistribution preferences and lying attitudes (Bortolotti et al. [2017]; Kajackite [2018]), and on how individuals react to fairness violations (Houser et al. [2012]; Spicer and Becker [1980]), our main conjecture is that in presence of an opaque income-generation process - where opportunities to cheat and maximize own revenues at the expenses of others are available - individuals would be more likely to engage in under-contribution (H1) and would show on average an higher degree of tolerance towards evasion (H2). We further test whether changes in individuals' contribution behavior can be, at least partially, explained by context-dependent 'injunctive' social norms on evasion tolerance (H3).

# 3 Experimental Design

The objective of the experiment is to re-create in the lab a small-scale economic system where individuals, as it happens in the real-life, first earn their income based on a combination of their own effort and luck, and are later asked to pay taxes contributing with a share of their income to finance publicly shared services. The main manipulation of the experiment refers to the presence, and the role, of *cheating* in the income-generating process, which allows to overcome the limitation of observational data where it would be impossible to distinguish and control whether, and to what extent, individuals ascribe others' earnings to their effort and ability rather than to some forms of cheating (e.g. personal or political connections or small bribes). The experimental design is specifically tailored to introduce cheating in the lab in a controlled way while minimizing subjects' scrutiny and observability concerns and the risks of a strong experimental demand effect.
The experimental design is based on a two-prong approach with two non-overlapping sets of participants: Experiment 1 is used to trace subjects' actual contribution behavior, while Experiment 2 serves to elicit (injunctive) social norms from an un-involved sample of participants.

The experiments were conducted in the Bologna Laboratory for Experiments in Social Sciences (BLESS) between October 2018 and May 2019. It was programmed using z-Tree (Fischbacher [2007]) and subjects were recruited via ORSEE (Greiner [2004]). A total of 270 subjects took part in the two experiments [4]: 180 subjects participated in Experiment 1 (Behaviors) and 90 subjects participated in Experiment 2 (Norms). Each session involved 15 subjects and lasted approximately 75-80 minutes with participants earning on average 13.5 Euros.

| | **EXP 1 - Behaviors** | **EXP 2 - Norms** |
|---|---|---|
| **T1 (*Cheating*)** | 90 obs. | 45 obs. |
| **T2 (*No Cheating*)** | 90 obs. | 45 obs. |

## Experiment 1: Behaviors

The experiment is divided into three phases. The design of the second and the last phase is identical across treatments, while the design of the first phases varies across the two *Cheating* and *No Cheating* treatments. In the $1^{st}$ phase, individuals are grouped in 5-people groups and have the opportunity to earn an income based on their performance in the *slider task* (Gill and Prowse [2012]).
In the slider task, participants see a series of 48 sliders ranging from 0 to 100 on their computer screen and have to adjust each slider to exactly the middle position (50) within the given 120 seconds (see Figure A.1 in the Appendix): subjects are allowed to use only the touch-pad to drag and adjust sliders' position [5] and earn one point for each correctly positioned slider. Subjects' objective is to maximize the number of correctly positioned sliders before the time is over: when the time runs out subjects are ranked based on their performance in the task with respect to their other group-mates, according to a rank-order tournament payment scheme with multiple prizes (Freeman and Gelber [2010], Moldovanu and Sela [2001]).
The earnings scheme is fixed across sessions and does not depend on the actual (absolute) performance of participants: this implies the level of income inequality in the society (group) at the end of the income-generation task is always the same in all sessions. Participants' earnings are expressed in ECUs (Experimental Currency Units), which are then converted in Euros at the exchange rate 1 ECU = 0.5 Euros.

[4] Additional 64 participants took part in the first four pilot sessions of the two experiments, run in June 2018. Data from these sessions are discarded from the analysis due to a slight difference in experimental procedures.

[5] Throughout the whole duration of the experiment we used a keyboard locker to prevent subjects from using the arrow keys or the mouse wheel

| | |
|---|---|
| $1^{st}$ Best performance | 20 ECUs |
| $2^{nd}$ Best performance | 16 ECUs |
| $3^{rd}$ Best performance | 12 ECUs |
| $4^{th}$ Best performance | 8 ECUs |
| $5^{th}$ Best performance | 4 ECUs |

Subjects face the slider task three times: the first two times represent trial sessions [6], the last session is only one relevant for the determination of subjects' earnings. When the first two trial sessions are over subjects receive instructions on the main execution rule [7]: subjects are instructed to use their non-dominant hand only when executing the task in the last payoff-relevant round. Our treatment manipulation relates to the monitoring over the implementation of the non-dominant hand rule and on the availability of cheating opportunities, which subjects can choose to exploit to boost their performance in order to get to higher rank positions and obtain higher earnings. In this context, exploiting cheating opportunities goes 'at the expenses' of other group members: cheating, in fact, cannot increase the amount of total surplus in the group but can only change the way in which it is distributed among members, favoring cheaters at the expenses of the best-performing honest members.

In the *No Cheating* treatment cheating is impossible because we can can implement perfect monitoring on the respect of the non-dominant hand rule: when subjects arrive to the lab we are able to identify left- from right-hand users by checking which hand they use to sign the lab attendance sheet [8]. We assign each subject textile glove, which has a different color based on whether the subject is classified as a left- (gray glove) or right-hand user (white glove). After subjects learn of the non-dominant hand rule, in the *No Cheating* treatment they are invited to wear the textile glove over their dominant hand and to place the hand wearing the glove on their desk in plain sight: these textile gloves prevent subjects from using the touch-pad with their dominant hand and, given the two different colors, allow us to check whether the subjects are wearing the glove on the right hand during the execution of the payoff-relevant task.

Conversely, no form of monitoring is implemented in the *Cheating* treatment: subjects are simply instructed to use their non-dominant hand during the payoff-relevant task but we are not able to check whether the rule is respected as we do not mark subjects' dominant hands at the beginning of the experiment and therefore we cannot implement any form of monitoring.

When the last session of the slider task is over subjects are informed of their absolute performance, of their relative performance in the group with respect to other group members and of their realized earnings [9].

The $2^{nd}$ phase is the same irrespective of what treatment subjects have been exposed to in the $1^{st}$ phase.

All subjects, irrespective of what is their position in the income distribution, are required to contribute with the same share of their realized earnings - equal to 25% - to a project that is common to all their group mates: subjects know that all contributions made by group members will be summed up, multiplied by a factor of m=2 and then equally divided among participants, as in a standard Public Good Game. This design mimics a proportional income-tax scenario, where tax contributions have moderate redistributive effects. Returns from contributions to the public account are decreasing in the level of individuals' pre-contribution income and, under a full-contribution scenario, we observe the highest efficiency and redistribution gains: global aggregate earnings increase by 25% and inequality decreases, with a 10pp decline in the Gini index, from 0.26 to 0.16. When some individuals decide to under-contribute or not to contribute at all, however, the distribution of gains changes and the effects on inequality can also be negative: if, for example, all individuals but top-earners fully contribute, we will have higher inequality than in the pre-tax scenario, with an ex-post Gini of 0.28, and lower efficiency, with a lower increase in overall aggregate earnings.

[6] The first trial session lasts 200 seconds, the second one 120 seconds just like the final incentivized session

[7] This design allows us to use data on the second trial session lasting 120 seconds as a control for subjects' individual ability in the task

[8] We double check by explicitly asking them which hand they usually use as their dominant-hand in daily tasks

[9] Before providing this information we elicit subjects' guesses on their positions in the group ranking and subjects' perception on the frequency of cheating occurrences in the payoff-relevant slider task to check whether our manipulation is effective in altering subjects' perception on the presence and intensity of cheating across treatments.

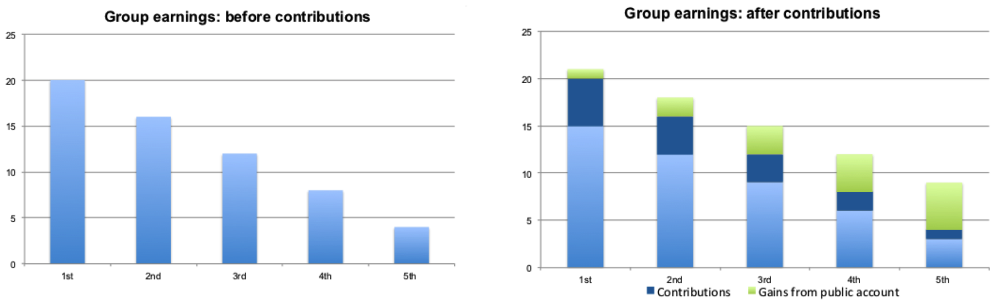



Figure 4a: Income distribution before and after contributions (full contribution scenario)

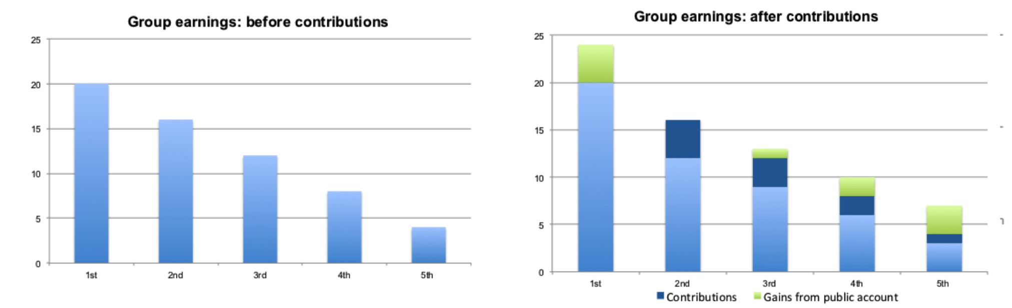



Figure 4b: Income distribution before and after contributions (full contribution by all but top-earners)

To minimize demand effects and guarantee the highest level of anonymity, tax contributions are collected through a paper-based procedure that is largely inspired by the 'double-anonymous' treatment procedures adopted by Barmettler et al. [2012]. This procedure, which ensures subjects' full anonymity towards both the experimenter and the other subjects of their group and sets to zero the detection probability for under-contributers ([10]. By design, before the experiment is over there's no way to detect under-contribution at the individual level thus detection probability for evaders is zero. At the end of the contribution phase subjects are paid in cash privately and the payment procedures ensure that choices taken by subjects throughout the experiment cannot be matched with subjects' real identities.
In the $3^{rd}$ and last phase of the experiment subjects are asked to answer to the 60-items version of the HEXACO personality test (Ashton and Lee [2009]) [11].

## Experiment 2: Norms

The experiment is divided into four phases.
In the $1^{st}$ phase of this experiment, a group subjects who never did or will actually participate in Experiment 1, is asked to perform the slider task twice: first using their dominant hand and later using their non-dominant hand [12]. Subjects are not paid based on their performance and receive a fixed payment of 2 Euros for their participation.
The design of the $2^{nd}$ phase of the experiment heavily relies on the contributions by Krupka and Weber [2013] and Krupka et al. [2017]. In the $2^{nd}$ phase of the experiment, subjects are exposed to the same decision environment faced by participants of Experiment 1: they are exposed to the same instructions that are read aloud to subjects who took part in Experiment 1 (either in the *Cheating* or in the *No Cheating* treatment, subjects are exposed to a single treatment only, mirroring the procedure adopted for Experiment 1) and asked to rate the "degree of social acceptability" of each contribution choice available to subjects who

[10] For details, refer to the instructions reported in the Appendix *XX*

[11] More information available at: http://hexaco.org/

[12] To make the difference between the use of the dominant vs. non-dominant rule more salient we implement the use of textile gloves as in the *No Cheating* treatment of Experiment 1.

took part in Experiment 1 in the contribution-phase, knowing under-contribution was made possible by the implementation of the paper-based double-anonymous procedure. For each level of earnings (20, 16, 12 8 and 4 ECUs) subjects are asked to rate the social acceptability of each and all contribution choices available, using a 6-points scale that ranges from "Very socially unacceptable" to "Very socially acceptable".

| **Earnings = 4 ECUs ⇒ Expected contribution = 1** | | | | | | |
|---|---|---|---|---|---|---|
| | Very Socially Unacceptable | Socially Unacceptable | Somewhat Socially Unacceptable | Somewhat Socially Acceptable | Socially Acceptable | Very Socially Acceptable |
| C = 1 | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| C = 0 | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |

We rely on an incentivized procedure to elicit subjects' perceptions on existing *injunctive* social norms on the acceptability of all contribution choices available for each possible level of realized earnings. As in Krupka et al. [2017], we define injunctive social norms as the jointly recognized beliefs, among members of a population, regarding the acceptability of different behaviors (as distinct from descriptive social norms, which correspond to actions taken by most others).
We provide respondents with incentives to match their ratings to the responses of other subjects in the session, rather than to provide their personal opinions. Subjects are asked to provide what they believe would be the "average answer" provided by other participants in the session and are informed that at the end of this phase one of the contribution choices they rated will be randomly selected and they will be randomly paired to another participant in the room: their payoff depends on how similar their response will be to the answer provided by the matched participant, according to an incentive-compatible quadratic scoring rule [13].

| **If the response..** | **The payoff is equal to:** |
|---|---|
| Exactly matches partner's response | 7 Euros |
| Differs from partner's response by 1 category | 6.5 Euros |
| Differs from partner's response by 2 categories | 5 Euros |
| Differs from partner's response by 3 categories | 2.5 Euros |
| Differs from partner's response by 4 categories | - 1 Euro |
| Differs from partner's response by 5 categories | - 5.5 Euros |

In the $3^{rd}$ phase of the experiment we elicit subjects' beliefs on actual contribution behavior by subjects who actually took part in Experiment 1 in first person. We elicit beliefs on what would be the selected contribution option under all different earnings' levels. The elicitation of beliefs is also incentivized: subjects are informed that their guesses will be compared to the actual choice taken by a randomly-selected participant who took part in one of the previous sessions of Experiment 1. If their guess matches the randomly-selected actual choice they receive a fixed prize of 3 Euros.
In the $4^{rth}$ and last phase of the experiment subjects answer the 60-items version of the HEXACO personality test.

# 4 Preliminary Results

It emerges that our manipulation is effective in altering subjects' perception of the intensity of cheating across treatments. Both in Experiment 1 (Behaviors) and Experiment 2 (Norms) subjects have significantly different perceptions of the concentration of cheating occurrences in the pre-tax income-generating task across treatments (Figure 5): subjects perceive that violations of the non-dominant hand protocol for execution of the income-generating task, which result in a form of cheating at the expenses of others, are more frequent

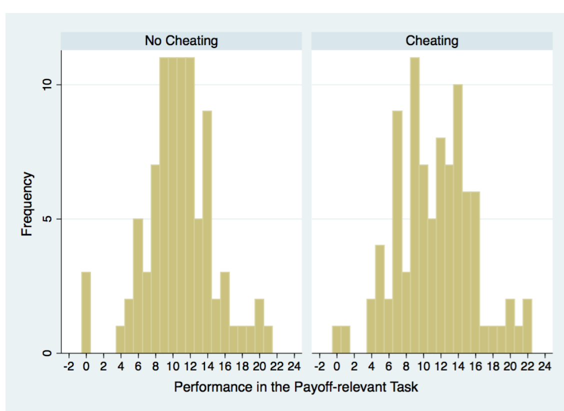
[13] According to the quadratic scoring rule, the payoff corresponds to the maximum score possible $\alpha$ minus the inaccuracy of the forecast, computed as the sum of squared deviations: given subject's $i$ guess p on choice k, the payoff of subject $i$ will depend on how similar her guess will be from the guess reported by her randomly-selected partner j $P_i(p_{i,k}) = \alpha - \beta(p_{j,k} - p_{i,k})$, where $\alpha = 7$ and $\beta = 1/2$.

in ‘Cheating’ treatments than in ‘No-Cheating’ treatments.
The distributions of subjects’ perceptions are significantly different in both Experiments: the Rank-sum z statistic is equal to -9.346 (p-value 0.000) in Experiment 1 and to -3.987 (p-value 0.000) in Experiment 2.


Figure 5: Manipulation check - Experiment 1 & 2

With respect to actual cheating occurrence, our design prevents us from detecting cheating at the individual level. If we look at the distribution of slider task’s scores collected by subjects in the payoff-relevant round we can observe that, on aggregate, subjects exposed to the ’Cheating’ treatment collect slightly higher scores but the two distributions are not statistically different (the Rank-sum z statistic is equal to -0.949 with a p-value=0.3427). The same picture emerges if we control for subjects’ individual ability in the task, looking at the difference in performance between the payoff-relevant round and the last trial round (both lasting 120 seconds). As predictable, the implementation of the non-dominant hand rule in the payoff-relevant round leads subjects to accrue, on average, lower scores with respect to the trial round: the difference in performance with respect to the trial round is lower in the ’Cheating’ treatment, however, the difference between the two distributions is not statistically significant (the Rank-sum z statistic is equal to -1.325 with a p-value=0.1851). This evidence suggests that on average subjects followed the non-dominant hand restriction even in absence of monitoring, although the presence of a high degree of heterogeneity in subjects’ learning in the task could also partially explain this attenuated result.

**Result 1**: *Our novel experimental procedure is effective in altering subjects’ perceptions on the occurrence of cheating across treatments.*



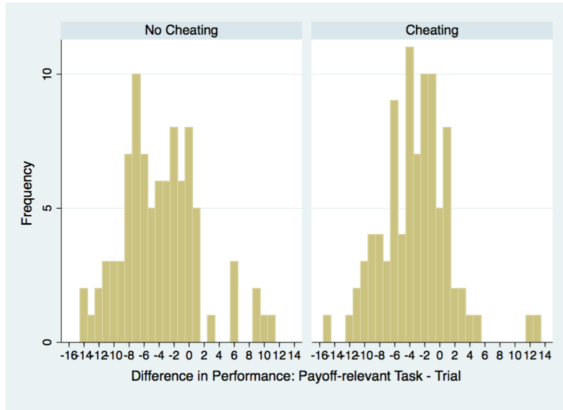



Figure 6: Experiment 1 - Slider Task data

*Experiment 1 - Behaviors*

With respect to tax evasion behaviors, our main conjecture is that in presence of an opaque income-generation process - where opportunities to cheat and maximize own revenues at the expenses of others are available - individuals would be more likely to engage in under-contribution behaviors (H1). We are particularly interested in the behavior of individuals in the upper-middle part of the income distribution (those with pre-contribution earnings equal to 16 and 12 ECUs): these individuals are in the uncomfortable situation of getting smaller gains from redistribution compared to low-income individuals in the full-contribution scenario, and to be severely damaged in case of full or partial undercontribution by top-income individuals. These subjects risk to end up being "twice losers" in the Cheating treatment: being cheated twice by top-income individuals, who could first cheat in the income-generation task to boost their performance at their expenses and undeservedly get to the top position of the distribution, and then cheat again in the contribution phase, contributing less than required to the public account causing a distortion of the relative weight of the contribution burden on the shoulders of upper-middle income individuals.

Our preliminary analysis reveals there is no significant difference in subjects' aggregate contribution behavior across treatments (the Rank-sum z statistic is equal to 0.050, p-value 0.9599). In both treatments we are clearly far from a full contribution scenario, with an average contribution of 59.5% and a non-negligible quota of full under-contribution occurrences (slightly less than 30% of cases). If we look at subjects' contribution behavior conditioning on pre-contribution earnings level, we observe some differences across earnings levels and a slight difference across treatments for top earners. However, a series of non-parametric Mann-Whitney tests reveal that contribution behavior is not statistically different across treatments even after we break down our observations by pre-contribution earnings' levels [14].

[14] Two-sample Wilcoxon rank-sum (Mann-Whitney) tests: (a) Earnings=20, z=-0.860 p-value=0.3898; (b) Earnings=16, z=0.067 p-value=0.9463; (c) Earnings=12,(a) z=0.346 p-value=0.7295; (d) Earnings=8, z=-0.250 p-value=0.8025; (e) Earnings=4, z=0.329 p-value=0.7420

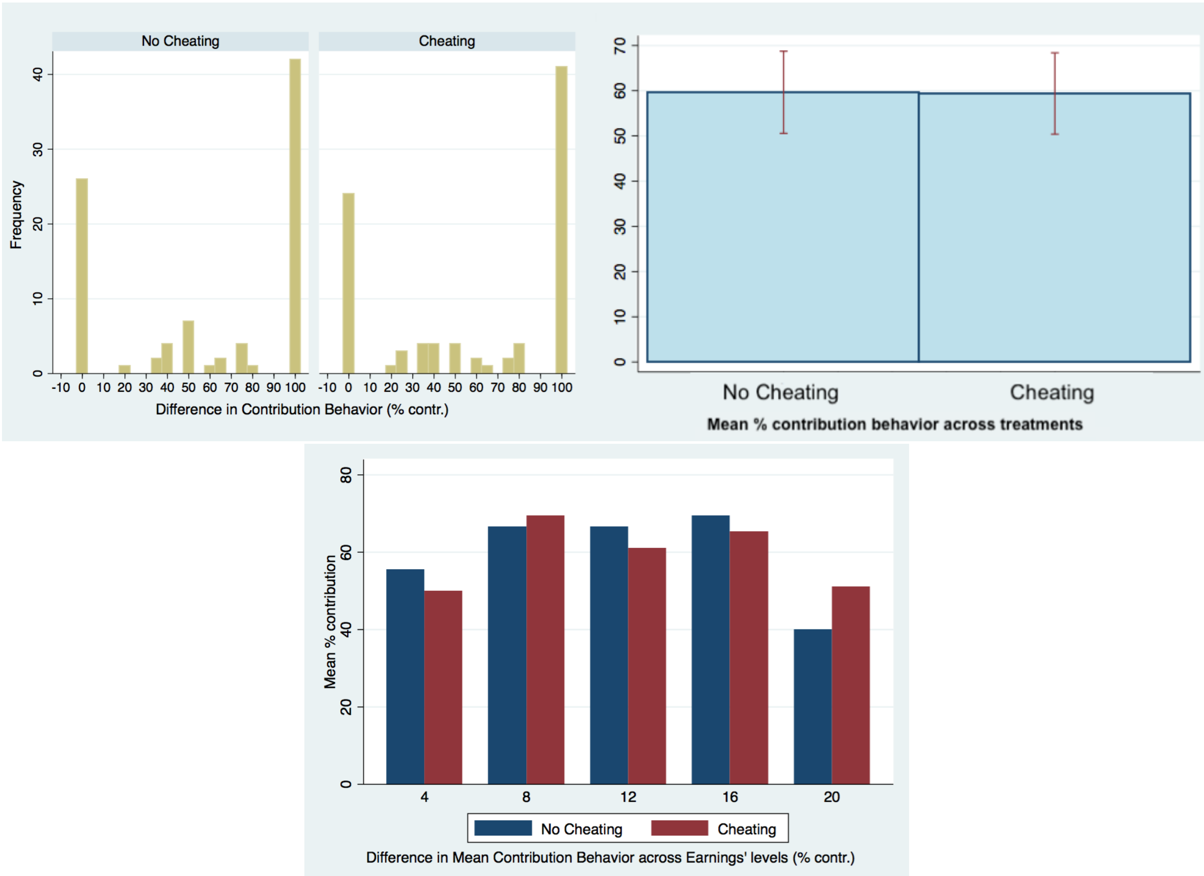


Figure 7: Experiment 1 - Contribution data

Pooling observations across treatments, a non-parametric Kruskal-Wallis equality-of-populations rank test reveals there's a marginal evidence in favor of heterogeneous distributions of contribution choices across pre-contribution earnings' levels ($\chi^2$ with ties=8.371, p-value=0.0789): a parametric analysis shows that this result is mainly driven by the contribution behavior of top-earners as this is the only group whose contribution behavior is, marginally, statistically lower

Since our design for Experiment 1, in order to ensure subjects' complete anonymity in the contribution task, doesn't allow us to observe cheating at the individual level, we are not in the position to disentangle the effects of (the perceptions of) *cheating by others* and of own cheating - which might induce some "moral cleansing" concerns - on contribution behavior. However, we can observe that perceptions of the intensity of cheating by others are decreasing in the level of realized earnings from the slider task, and at the same time, we can reasonably assume that if cheating occurred it occurred more likely at the top position of the distribution rather than at the bottom. Looking at the results on subjects' contribution behavior we can infer that any of these two effects - which would have called for higher(lower)-than-average contribution by top(bottom) earners - seem to be strong. The lower-than-average contribution observed for top earners, instead, is compatible with both: (a) a stronger entitlement effect driven by lower perceptions of cheating intensity, especially in the No-Cheating scenario, and (b) a mere "magnitude of stakes effect", given that top earners by design have the highest direct benefit from evasion in absolute terms.

**Result 2**: *Our treatment manipulation on the intensity of cheating does not lead to difference in subjects' contribution behavior across treatments.*

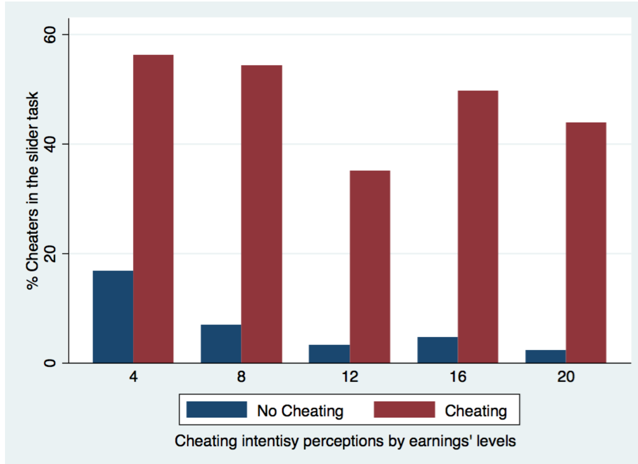



Figure 8: Experiment 1 - Cheating intensity perceptions

*Experiment 2 - Social Norms*

We explore the determinants of norms on evasion acceptability, investigating whether norms are sensitive to (i) the presence of cheating opportunities in the pre-tax income-generation process, (ii) the size of evasion and (iii) individual beliefs on the extent of under-contribution in each contribution context.

The general pattern of social acceptability ratings is the same across treatments.

Norms on tax evasion acceptability appear to be highly sensitive to the size of evasion: average acceptability ratings decrease as the size of evasion increases and the modal evaluations on the two ”extreme” actions (full contribution and full under-contribution) are always polarized at the two respective extremes of the social acceptability spectrum: full-contribution is generally regarded as a ”very socially acceptable” action while full under-contribution is recognized as a ”very socially unacceptable” action. Ratings on the acceptability of ”extreme” contribution actions, however, seem to be sensitive to the level of pre-contribution earnings' levels as well: full-contribution choices tend to be more highly rated when selected by high/top earners, while full under-contribution choices tend to be more highly rated when associated with low/bottom levels of pre-contribution earnings.

These findings are confirmed by our parametric analysis: the size of evasion has a negative and sizeable significant effect on acceptability ratings, while the presence of the shadow of cheating does not lead to any significantly impact. It also emerges that the negative effect of evasion size is larger the higher the income level of the evader, although the interaction effect between income level and evasion size is not dramatic in size. These results are stable even after we control for individual beliefs on the actual level of contribution for each level of earnings, which appear to per se negatively affect acceptability norms: the higher subjects' belief on actual contribution behavior, the lower subjects' tolerance for evasion in those contexts.

**Result 3**: *Norms on tax evasion acceptability are highly sensitive to the size of the evasion and respond to differences in pre-contribution levels of earnings but do not statistically differ across treatments.*

Table 1: Experiment 2 - Social acceptability ratings

| Action | NO-CHEAT (n=45) | | | | | | | CHEAT (n=45) | | | | | | | Rank-sum z test |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Mean | --- | -- | - | + | ++ | +++ | Mean | --- | -- | - | + | ++ | +++ | |
| **Earnings = 4 \| C=1** | 5.26 | 2.22% | 0.00% | 2.22% | 17.78% | 20.00% | 57.78% | 5.26 | 0.00% | 2.22% | 2.22% | 15.56% | 26.67% | 53.33% | 0.233 |
| **Earnings = 4 \| C=0** | 2.86 | 17.78% | 20.00% | 31.11% | 20.00% | 11.11% | 0.00% | 2.73 | 15.56% | 26.67% | 31.11% | 22.22% | 4.44% | 0.00% | 0.486 |
| **Earnings = 8 \| C=2** | 5.42 | 2.22% | 0.00% | 2.22% | 6.67% | 26.67% | 62.22% | 5.51 | 0.00% | 0.00% | 0.00% | 8.89% | 31.11% | 60.00% | 0.944 |
| **Earnings = 8 \| C=1** | 3.82 | 0.00% | 11.11% | 13.33% | 57.78% | 17.78% | 0.00% | 3.77 | 0.00% | 11.11% | 17.78% | 53.33% | 17.78% | 0.00% | 0.286 |
| **Earnings = 8 \| C=0** | 2.11 | 33.33% | 31.11% | 28.89% | 4.44% | 2.22% | 0.00% | 2.17 | 28.89% | 33.33% | 28.89% | 8.89% | 0.00% | 0.00% | -0.410 |
| **Earnings = 12 \| C=3** | 5.44 | 0.00% | 2.22% | 2.22% | 4.44% | 31.11% | 60.00% | 5.55 | 0.00% | 0.00% | 0.00% | 4.44% | 35.56% | 60.00% | -0.187 |
| **Earnings = 12 \| C=2** | 4.22 | 0.00% | 2.22% | 8.89% | 53.33% | 35.56% | 0.00% | 4.31 | 0.00% | 0.00% | 6.67% | 57.78% | 33.33% | 2.22% | -0.350 |
| **Earnings = 12 \| C=1** | 2.8 | 2.22% | 33.33% | 46.67% | 17.78% | 0.00% | 0.01% | 2.84 | 0.00% | 33.33% | 51.11% | 13.33% | 2.22% | 0.00% | -0.136 |
| **Earnings = 12 \| C=0** | 1.73 | 53.33% | 28.89% | 11.11% | 4.44% | 2.22% | 0.00% | 1.71 | 44.44% | 42.22% | 11.11% | 2.22% | 0.00% | 0.00% | -0.419 |
| **Earnings = 16 \| C=4** | 5.4 | 0.00% | 4.44% | 2.22% | 4.44% | 26.67% | 62.22% | 5.51 | 0.00% | 0.00% | 0.00% | 8.89% | 31.11% | 60.00% | 0.051 |
| **Earnings = 16 \| C=3** | 4.31 | 0.00% | 4.44% | 8.89% | 37.78% | 48.89% | 0.00% | 4.48 | 0.00% | 0.00% | 2.22% | 46.67% | 51.11% | 0.00% | -0.704 |
| **Earnings = 16 \| C=2** | 3.26 | 2.22% | 11.11% | 48.89% | 33.33% | 4.44% | 0.00% | 3.46 | 0.00% | 4.44% | 46.67% | 46.67% | 2.22% | 0.00% | -1.252 |
| **Earnings = 16 \| C=1** | 2.33 | 13.33% | 48.89% | 28.89% | 8.89% | 0.00% | 0.00% | 2.42 | 0.00% | 66.67% | 26.67% | 4.44% | 2.22% | 0.00% | -0.387 |
| **Earnings = 16 \| C=0** | 1.46 | 68.89% | 24.44% | 2.22% | 2.22% | | 2.22% | 1.51 | 62.22% | 28.89% | 4.44% | 4.44% | | 0.00% | -0.656 |
| **Earnings = 20 \| C=5** | 5.64 | 0.00% | 0.00% | 6.67% | 0.00% | 15.56% | 77.78% | 5.64 | 0.00% | 0.00% | 0.00% | 6.67% | 22.22% | 71.11% | 0.620 |
| **Earnings = 20 \| C=4** | 4.64 | 0.00% | 2.22% | 4.44% | 22.22% | 68.89% | 2.22% | 4.75 | 0.00% | 0.00% | 2.22% | 22.22% | 73.33% | | -0.554 |
| **Earnings = 20 \| C=3** | 3.73 | 2.22% | 4.44% | 20.00% | 64.44% | 8.89% | 0.00% | 3.97 | 0.00% | 2.22% | 8.89% | 77.78% | 11.11% | 0.00% | -1.639 |
| **Earnings = 20 \| C=2** | 2.66 | 8.89% | 26.67% | 55.56% | 6.67% | 2.22% | 0.00% | 3.02 | 0.00% | 11.11% | 77.78% | 8.89% | 2.22% | 0.00% | -2.413 ** |
| **Earnings = 20 \| C=1** | 1.97 | 22.22% | 64.44% | 8.89% | 2.22% | 2.22% | | 2.22 | 6.67% | 68.89% | 20.00% | 4.44% | 0.00% | 0.00% | -2.106 ** |
| **Earnings = 20 \| C=0** | 1.35 | 80.00% | 13.33% | 2.22% | 2.22% | 0.00% | 2.22% | 1.35 | 75.56% | 15.56% | 6.67% | 2.22% | 0.00% | 0.00% | -0.484 |

Notes: * p < 0.1; ** p < 0.05; *** p < 0.01; all two-tailed.
Responses are: "Very socially acceptable" (+++), "Socially acceptable" (++), "Somewhat socially acceptable" (+), "Somewhat socially unacceptable" (-), "Socially unacceptable" (--), "Very socially unacceptable" (+++). Responses were converted into numerical scores: "Very socially acceptable" = 6, "Socially acceptable" = 5, "Somewhat socially acceptable" = 4, "Somewhat socially unacceptable = 3, "Socially unacceptable" = 2, "Very socially unacceptable" = 1. Modal responses are shaded.

Table 2: Experiment 2 - Regressions on acceptability ratings

| | [1a] | [1b] | [2a] | [2b] |
|---|---|---|---|---|
| *Cheating* | -0.0640 | -0.0732 | -0.0886 | -0.0983 |
| | (0.225) | (0.229) | (0.227) | (0.231) |
| *Size of evasion (%)* | -2.271*** | -2.271*** | -2.271*** | -2.271*** |
| | (0.281) | (0.281) | (0.281) | (0.282) |
| *Level of earnings* | 0.0103 | 0.0103 | 0.0198*** | 0.0200*** |
| | (0.00764) | (0.00766) | (0.00655) | (0.00662) |
| *Cheating x Size of evasion* | -0.0543 | -0.0543 | -0.0543 | -0.0543 |
| | (0.291) | (0.292) | (0.292) | (0.292) |
| *Cheating x Level of earnings* | 0.0129 | 0.0129 | 0.0208* | 0.0203* |
| | (0.00987) | (0.00988) | (0.0115) | (0.0115) |
| *Size of evasion x Level of earnings* | -0.107*** | -0.107*** | -0.107*** | -0.107*** |
| | (0.00995) | (0.00997) | (0.00996) | (0.00998) |
| *Belief* | | | -0.0637*** | -0.0652** |
| | | | (0.0244) | (0.0262) |
| *Cheating x Belief* | | | -0.0280 | -0.0250 |
| | | | (0.0508) | (0.0501) |
| *Constant* | 5.265*** | 5.276*** | 5.278*** | 5.304*** |
| | (0.183) | (0.205) | (0.185) | (0.208) |
| Observations | 1,800 | 1,800 | 1,800 | 1,800 |
| Number of id | 90 | 90 | 90 | 90 |
| Controls | No | Yes | No | Yes |

The Dependent variable is the norm rating [1–6] for each contribution action available in all Earnings' scenarios. Robust standard errors in parentheses. Controls include indicators on whether individuals show levels of Honesty-Humility, Emotionality, eXtraversion, Agreeableness, Conscientiousness, Openness to Experience (as measured by the HEXACO questionnaire) above sample median.

*Experiment 1 & 2 - Tax evasion behavior and norms*

Can social norms on evasion predict subjects' behavior in different tax contribution contexts? Our empirical results show that neither norms nor behaviors, which are elicited on two different groups of subjects, are deeply affected by the characteristics of the process that generates subjects' incomes, and thus, by the origin of economic inequality within the two groups. It emerges, however, that social acceptability ratings are fairly sensitive to some of the characteristics of evasive actions, such as the size of evasion, and, to a smaller extent, the level of evaders' earnings.

Table 3 shows the distribution of actions selected by subjects who took part in Experiment 1 (Behaviors), conditioning on subjects' pre-contribution earnings' level: the pattern of modal actions across all income levels suggests the presence of a tension between immediate monetary gains from evasion on one side, and norms' adherence on the other side. Immediate monetary gains from evasion are obviously higher the lower the individual contribution to the tax-collection pool, and reach their peak with full under-contribution choices; on the side of norms, instead, we can see that higher contribution choices always associated to higher acceptability ratings.

The effect of norms seems to be particularly relevant for subjects who are not at the top of the income distribution: the modal contribution action for these subjects always corresponds to the action regarded as the most socially acceptable. The same is not true for top-earners, whose modal action, although different across treatments, seems to be driven more intensely by immediate monetary benefits rather than by the desire to comply with social prescriptions on what is considered to be acceptable.

This evidence could be consistent with the hypothesis that acceptability norms play a role in subjects' decision process, with subjects exhibiting a desire for norms' compliance on top of the interest for monetary payoff maximization.

Table 3: Experiment 1 - Contribution choices data

| | **NO-CHEAT (n=60)** | | | | | | | **CHEAT (n=60)** | | | | | | | *Rank-sum p value* |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *Mean* | C=0 | C=1 | C=2 | C=3 | C=4 | C=5 | *Mean* | C=0 | C=1 | C=2 | C=3 | C=4 | C=5 | |
| **Earnings = 4** | *0.55* | 44.44 % | 55.56% | . | . | . | . | *0.50* | 50.00% | 50.00% | . | . | . | . | *0.7420* |
| **Earnings = 8** | *1.33* | 22.22% | 22.22% | 55.56% | . | . | . | *1.39* | 22.22% | 16.67% | 61.11 % | . | . | . | *0.8025* |
| **Earnings = 12** | *2.00* | 22.22% | 11.11% | 11.11% | 55.56% | . | . | *1.83* | 22.22% | 22.22% | 5.56% | 50.00% | . | . | *0.7295* |
| **Earnings = 16** | *2.78* | 16.67% | 0.00% | 16.67% | 22.22% | 44.44% | . | *2.61* | 16.67% | 16.67% | 5.56% | 11.11% | 50.00% | . | *0.9463* |
| **Earnings = 20** | *2.03* | 38.89% | 5.56% | 22.22% | 5.56% | 5.56% | 22.22% | *1.82* | 22.22% | 5.56% | 22.22% | 11.11% | 22.22% | 16.67% | *0.3898* |

Notes: * p < 0.1; ** p < 0.05; *** p < 0.01; all two-tailed.
Modal responses are shaded.

Following Krupka and Weber [2013] and Krupka et al. [2017], we examine whether subjects' choices are guided by a desire for norms' compliance by fitting individual utility functions to choice data collected through Experiment 1, while relying on the ratings on the social acceptability of evasion that have been separately identified through Experiment 2.

We assume that subjects, when facing the contribution decision, follow a logistic choice rule, where the likelihood of selecting any action $a$ out of $n$ alternatives depends on the relative utility attached to that action, compared to the other alternatives available:

$$P(a = a_k) = \frac{exp(U_k)}{\sum_{i=1}^{n} exp(U_i)} \tag{1}$$

We model two different utility specifications: the first specification assumes subjects' utility depends only on their monetary payoff (Selfish model - Eq. 2) [15] and allows us to estimate the weight subjects place on the monetary benefit ($mp$) they get from a particular choice ($\beta$). The second specification (Norms model - Eq. 3) assumes subjects care both about the monetary payoff and the degree to which a particular choice is perceived to be social acceptable, where the social norm $N(a_k)$ is the empirically observed judgement on social acceptability, which should reflect the existing norm in the relevant group; this specification allows us to estimate the degree to which the subjects really care about adhering to a particular norm ($\gamma$).

$$U_i(mp, a_k) = \beta mp(a_i, k) \tag{2}$$

[15] We impose a linear restriction on the effect of the monetary payoff on utility.

$$U_i(mp, a_k) = \beta mp(a_i, k) + \gamma N(a_k) \tag{3}$$

The estimation is obtained through a conditional logit regression (McFadden [1974]). In the model, the dependent variable is a binary indicator that identifies whether a particular choice, out of all those available in each subject's menu, is selected. The independent variables capture the characteristics of the possible contribution choices:

- each choice's immediate monetary payoff, which depends on the pre-contribution income level and the contribution choice selected;
- and each choice's degree of social acceptability, obtained as the average social acceptability rating attached to each alternative (N(ak)).

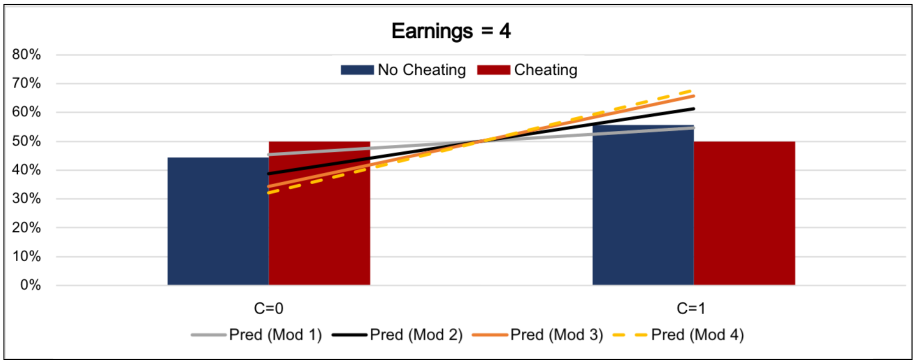
Table 4: Experiment 1 & 2 - Conditional Logit estimation

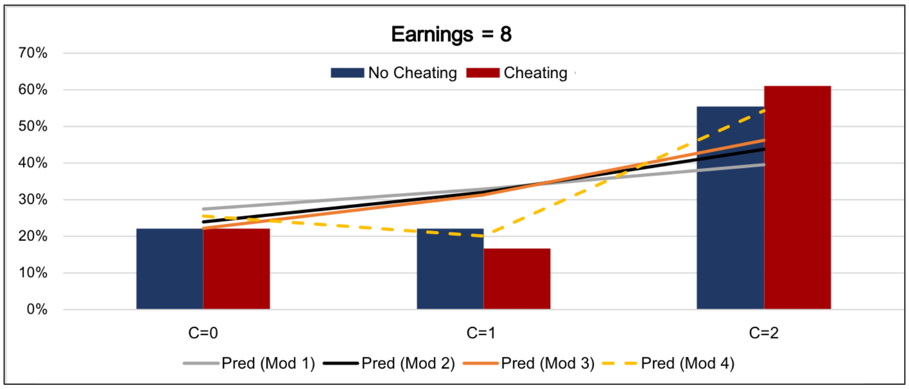
| | [1] | [2] | [3] | [4] |
|---|---|---|---|---|
| *Monetary Benefit* | -0.882** | | 0.881 | |
| | (0.360) | | (1.094) | |
| *Norm Rating* | | 0.257*** | 0.469* | 0.181** |
| | | (0.0874) | (0.274) | (0.0725) |
| *Norm Rating* $^2$ | | | | 0.441*** |
| | | | | (0.106) |
| Observations | 720 | 720 | 720 | 720 |
| Log-likelihood | -232.7 | -231.3 | -230.9 | -222.1 |
| BIC | 471.9 | 469.1 | 474.9 | 457.4 |
| Pseudo-$R^2$ | .0176 | .0236 | .0253 | .0623 |

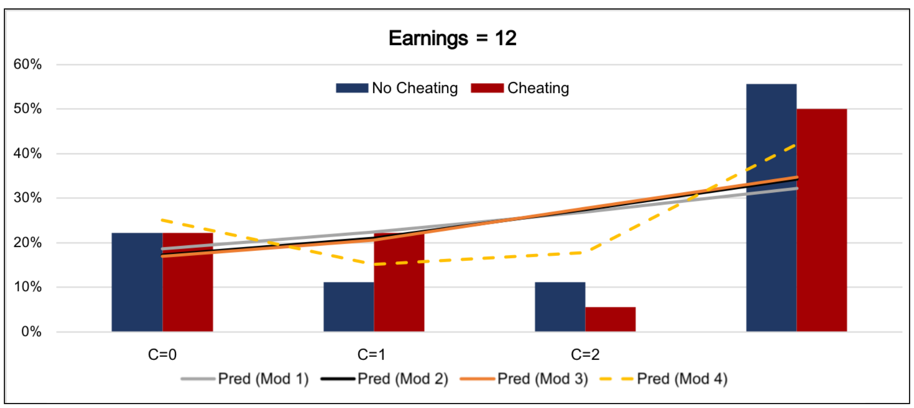
The Dependent variable is the chosen contribution action in the PGG game (Experiment 1). The variable Norm Rating converts subject responses (Experiment 2) to numerical scores: 1="very socially unacceptable", 2="socially unacceptable", 3="somewhat socially unacceptable", 4="somewhat socially acceptable", 5="socially acceptable", and 6="very socially acceptable".
Standard errors are clustered at the subject level and reported in parentheses.

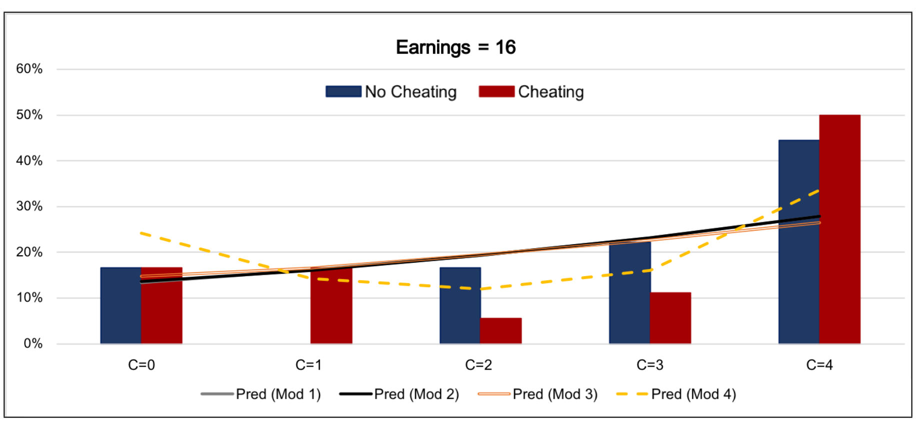
Table 4 reports the estimation results: the coefficient on monetary payoffs in the Selfish model in column (1) is negative and statistically significant, which would imply that the higher the monetary payoffs the lower the probability that subjects will select that specific action. This result is driven by the negative correlation between monetary payoffs and acceptability ratings, which are omitted from the Selfish model. Once we include also social norms as an explanatory variable in the model, the coefficient on monetary payoff turns positive, although not significant, as shown by the Norms model in column (3), while the coefficient on social appropriateness is positive and slightly significant. Norms appear to be the only relevant driver of subjects' choices and moving from the Selfish to the Norms specification results in an improvement in in terms of model's predictive fit. A further improvement is reached when estimating a model where subjects' utility depends only on choices' social appropriateness and subjects' utility is non-linear in the degree of choices' social appropriateness (column 4).
To get a sense of how well the different models can qualitatively account for the choice data collected through Experiment 1, we graphically compare the predictive performance of the models with actual choice frequencies observed. Except for the Earnings=4 scenario, where the Selfish model outperforms the others predicting almost an equal split between the two available options – as it is the case in the data, in all other scenarios the models that account for the role of norms clearly outperform the Selfish model.

**Result 4**: *Social norms on the acceptability of evasion elicited in Experiment 2 are relevant predictors for subjects' actual choices observed in Experiment 1.*


Earnings = 4
No Cheating
Cheating
80%
70%
60%
50%
40%
30%
20%
10%
0%
C=0
C=1
Pred (Mod 1)
Pred (Mod 2)
Pred (Mod 3)
Pred (Mod 4)


Earnings = 8
No Cheating
Cheating
70%
60%
50%
40%
30%
20%
10%
0%
C=0
C=1
C=2
Pred (Mod 1)
Pred (Mod 2)
Pred (Mod 3)
Pred (Mod 4)


Earnings = 12
No Cheating
Cheating
60%
50%
40%
30%
20%
10%
0%
C=0
C=1
C=2
Pred (Mod 1)
Pred (Mod 2)
Pred (Mod 3)
Pred (Mod 4)


Earnings = 16
No Cheating
Cheating
60%
50%
40%
30%
20%
10%
0%
C=0
C=1
C=2
C=3
C=4
Pred (Mod 1)
Pred (Mod 2)
Pred (Mod 3)
Pred (Mod 4)

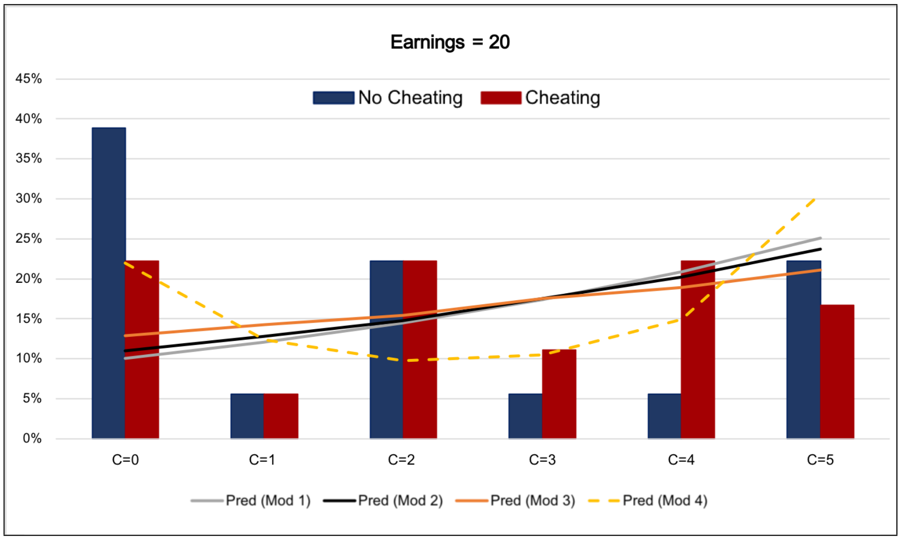


Figure 9: Distributions of predicted and observed choices in Experiment 1

# 5 Expenses Report

This project has been generously funded by one of the IFREE Small Research Grant for a total of $5,200 (approximately 4400 Euros). As agreed, the IFREE grant has been used exclusively for activities connected to the payment of monetary incentives for experimental subjects and to cover materials'expenses.
Specifically:

- 40 Euros have been used to cover materials' expenses;
- 715 Euros have been paid out as monetary incentives for subjects who took part in the Pilot sessions;
- 3600 Euros have been paid out as monetary incentives for subjects who took part in the main sessions of Experiment 1 and 2;

# References

Alesina, A. and Angeletos, G.-M. (2005). Fairness and redistribution. *American Economic Review*, 95(4):960–980.

Allingham, M. G. and Sandmo, A. (1972). Income tax evasion: a theoretical analysis. *Journal of Public Economics*, 1(3-4):323–338.

Alm, J., Jackson, B., and McKee, M. (1992). Estimating the determinants of taxpayer compliance with experimental data. *National Tax Journal*, pages 107–114.

Ashton, M. C. and Lee, K. (2009). The hexaco-60: A short measure of the major dimensions of personality. *Journal of Personality Assessment*, 91:340–345.

Barmettler, F., Fehr, E., and Zehnder, C. (2012). Big experimenter is watching you! anonymity and prosocial behavior in the laboratory. *Games and Economic Behavior*, 75(1):17–34.

Bordignon, M. (1993). A fairness approach to income tax evasion. *Journal of Public Economics*, 52(3):345–362.

Bortolotti, S., Soraperra, I., Sutter, M., and Zoller, C. (2017). Too lucky to be true: Fairness views under the shadow of cheating. *IZA DP No. 10877*.

Cappelen, A. W., Cappelen, C., and Tungodden, B. (2018). Second-best fairness under limited information: The trade-off between false positives and false negatives. *NHH Dept. of Economics Discussion Paper N.18*.

Cappelen, A. W., Hole, A. D., Sorensen, E. O., and Tungodden, B. (2007). The pluralism of fairness ideals: An experimental approach. *American Economic Review N.18*, 97(3):818–827.

Durante, R., Putterman, L., and Weele, J. (2014). Preferences for redistribution and perception of fairness: An experimental study. *Journal of the European Economic Association*, 12(4):1059–1086.

Fischbacher, U. (2007). z-tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, 10(2):171–178.

Fortin, B., Lacroix, G., and Villeval, M. C. (2007). Tax evasion and social interactions. *Journal of Public Economics*, 91(11):2089–2112.

Freeman, R. B. and Gelber, A. M. (2010). Prize structure and information in tournaments: Experimental evidence. *American Economic Journal: Applied Economics*, 2(1):149–164.

Gill, D. and Prowse, V. (2012). A structural analysis of disappointment aversion in a real effort competition. *American economic review*, 102(1):469–503.

Greiner, B. (2004). The online recruitment system orsee 2.0-a guide for the organization of experiments in economics. *University of Cologne, Working paper series in economics*, 10(23):63–104.

Houser, D., Vetter, S., and Winter, J. (2012). Fairness and cheating. *European Economic Review*, 56(8):1645–1655.

Kajackite, A. (2018). Lying about luck versus lying about performance. *Journal of Economic Behavior & Organization*, 153:194–99.

Krupka, E. L., Leider, S., and Jiang, M. (2017). A meeting of the minds: Informal agreements and social norms. *Management Science*, 63(6):1657–2048.

Krupka, E. L. and Weber, R. A. (2013). Identifying social norms using coordination games: Why does dictator game sharing vary? *Journal of the European Economic Association*, 11(3):495–524.

Matsaganis, M. and Flevotomou, M. (2010). Distributional implications of tax evasion in greece. *LSE GreeSE Paper*, 31.

McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. *Zarembka P, ed. Frontiers in Econometrics (Academic Press, New York)*, page 105–142.

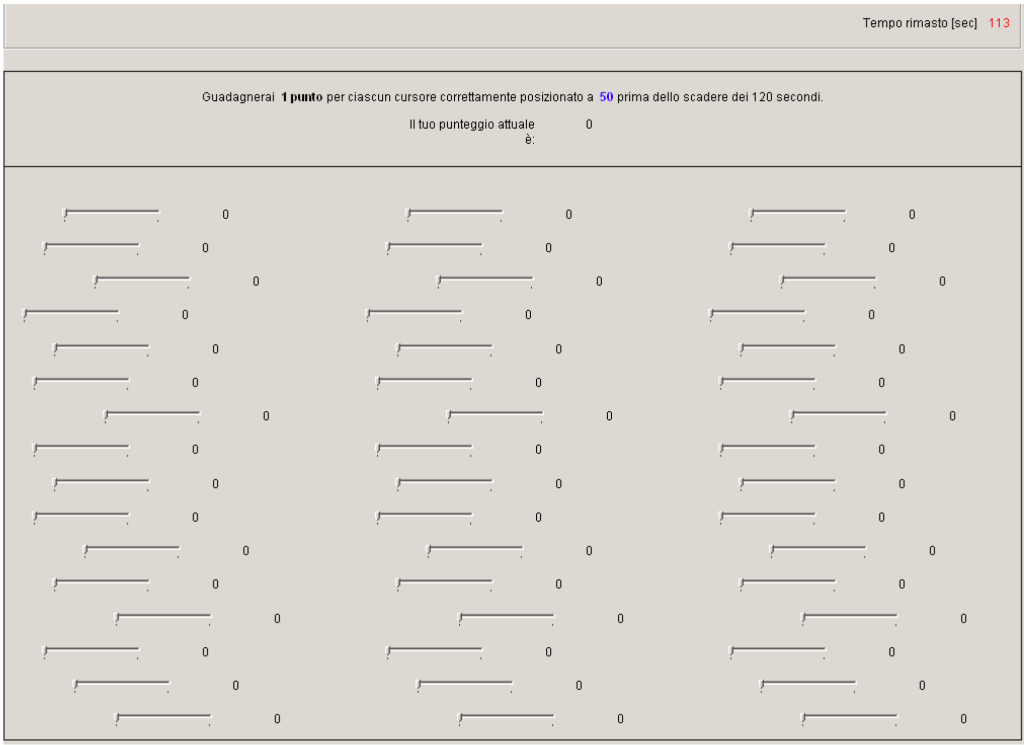
Moldovanu, B. and Sela, A. (2001). The optimal allocation of prizes in contests. *American Economic Review*, 91(3):542–558.

Slemrod, J. and Johns, A. (2010). The distribution of income tax noncompliance. *National Tax Journal*, 63(3):397.

Spicer, M. W. and Becker, L. A. (1980). Fiscal inequity and tax evasion: An experimental approach. *National Tax Journal*, pages 171–175.

# 6 Appendix

### Figure A1




### Table A1

| | [1] | [2] | [3] |
|---|---|---|---|
| *Cheating* | -5.556 | -0.278 | |
| | (14.36) | (6.368) | |
| *Earnings=4* | -11.11 | -11.11 | -11.11 |
| | (14.36) | (10.07) | (10.04) |
| *Earnings=8* | 4.24e-07 | 4.167 | 4.167 |
| | (14.36) | (10.07) | (10.04) |
| *Earnings=16* | 2.778 | 3.472 | 3.472 |
| | (14.36) | (10.07) | (10.04) |
| *Earnings=20* | -26.67* | -18.33* | -18.33* |
| | (14.36) | (10.07) | (10.04) |
| *Earnings=4 x Cheating* | -0 | | |
| | (20.31) | | |
| *Earnings=8 x Cheating* | 8.333 | | |
| | (20.31) | | |
| *Earnings=16 x Cheating* | 1.389 | | |
| | (20.31) | | |
| *Earnings=20 x Cheating* | 16.67 | | |
| | (20.31) | | |
| *Constant* | 66.67*** | 64.03*** | 63.89*** |
| *(Baseline: Earnings=12)* | (10.16) | (7.799) | (7.099) |
| Observations | 180 | 180 | 180 |
| R-squared | 0.048 | 0.043 | 0.043 |

OLS Regression: the dependent variable is individual contribution (expressed as a percentage of due contribution, equal to 25% of realized earnings)