Cheating Responses to Tax Evasion*

Pre-Analysis Plan

Andrea F.M. Martinangeli† Lisa Windsteiger‡

This draft: February 14, 2020

Abstract

We generalise the findings of the experimental literature investigating conditional behaviours to the setting of tax compliance. In particular, we investigate conditional behavioural responses to tax malpractice in a survey experiment on a representative sample of the Italian population and examine whether cheating rates in a die rolling task differ according to whether the respondent receives information about high or low estimates of the share of unreported income in Italy. Our design allows us to investigate whether asymmetries arise in the incidence of cheating according to whether the estimates of unreported income are higher or lower among the top compared to the bottom income earners.

JEL classification codes: D01, D31, D63, H23, H26
Keywords: Tax evasion, tax avoidance, conditional cooperation, cheating, survey experiment
Declaration of interests: None

*We thank Martin Kocher, Andreas Peichl and the participants of the Experimental Brown Bag Seminar at Ludwig-Maximilian-Universität Munich for their helpful suggestions and comments.
†Max Planck Institute for Tax Law and Public Finance. andrea.martinangeli@tax.mpg.de
‡Max Planck Institute for Tax Law and Public Finance. lisa.windsteiger@tax.mpg.de

1 Introduction

Conditional cooperation has been robustly observed across different societies: individuals cooperate with others (at risk of exploitation) insofar and inasmuch others cooperate in turn (Fischbacher et al., 2001). Such simple yet powerful intuition pertains to a wide range of economic activities, from household management to the funding of publicly provided goods. Frey and Torgler (2007) find in conditional cooperation a relevant force driving tax compliance: tax morale is reduced by the perceived pervasiveness of tax evasion in one's own country, perhaps triggering vicious cycles reminiscent of the cooperation collapse observed in laboratory public good games.

This project aims at further investigating the power of tax compliance (conversely, tax malpractice) in inducing conditional behavioural responses. Specifically, we aim at observing conditional behaviours typically observed in controlled laboratory environments within a broader paradigm, at generalising those findings and at uncovering the real-world bearing and consequences of (potentially asymmetric) conditional behavioural responses in the domain of tax compliance.

The salience of tax malpractice has in fact risen sharply in the past decade following the data leaks about tax evasion and avoidance via international tax shelters (Garside, 2016). It is now imperative to uncover and understand the fall-out of such practices, intensified by the heightened news focus, on social behaviours and particularly, in the light of conditional cooperation, on the propagation of unethical, antisocial and uncooperative behaviours.

Causal inference on conditional behavioural responses to tax compliance in natural settings is however impaired by the lack of credible exogenous variation in the conditioning variable. Within this framework, for instance, while suggestive of the conditional nature of tax compliance on perceived incidence of tax malpractice in society, the results presented by Frey and Torgler (2007) cannot be taken as identifying a causal link. It is in fact impossible to exclude the presence of inverse causality or of spurious correlation between the two phenomena. In this study we therefore adopt a survey-experimental design allowing us to observe direct responses to exogenously induced variation in (information about) estimated tax compliance. We are thus capable of uncovering the causal effect of tax malpractice on our respondents' behaviour.

Further, because unethical practices in tax compliance are most profitable and diffuse among the highest income earners (Alstadsæter et al., 2019), news reporting on the topic has naturally mainly been centred on the upper end of the income distribution. As a result tax malpractice on behalf of the upper echelons of our societies has

received greater resonance and political and popular interest than tax dishonesty on behalf of lower portions of the income distribution. Such impartial focus might well be consequential for the behavioural responses to tax malpractice throughout society, particularly on the propagation of (un)ethical behaviours.

From previous investigations on conditional behaviours (in particular from the large corpus of experimental economic research on cooperation; for instance Fischbacher et al. (2001) and citing literature), it is in fact not clear whether behaviours would be conditioned on those of others differently according to the income of the latter. Martinangeli (2017) finds however asymmetries along the income dimension in both (empirical) expectations and (expectation-driven) conditional behaviours. In his study subjects condition their cooperation on their beliefs over actions of (experimentally) rich but not of poor subjects irrespective of their own endowment size. Moreover, individuals expect greater cooperation from the rich than from the poor. Following those findings, we hypothesize that asymmetries in conditional cooperation might translate into asymmetric responses to tax evasion at the top and at the bottom of the income distribution. If true that behaviours are conditioned more strongly upon actions of rich individuals, the spreading of unethical behaviour might be quicker and more pronounced following increased salience of tax malpractice on behalf of richer rather than of poorer social segments.

Tax malpractice is a sensitive topic. Truthful reporting of behaviours in this domain are hard to obtain and ridden with systematic confounds and biases posing serious threats to the validity of the results obtained. Further complicating such matter, behaviours in the tax compliance domain can be hardly incentivised within the framework of a research design. For this reason, we study a proxy behaviour: cheating towards the experimenter (e.g. Fischbacher and Föllmi-Heusi (2013); Kocher et al. (2017)). Cheating behaviours are widely studied not only because of their intrinsic interest, but also because of their distributional and economic consequences on real world outcomes and their bearing real world economic phenomena, including tax malpractice. We hence adopt the cheating paradigm to shed light on the behavioural consequences of tax dishonesty and to offer a first approach at the generalisation of the findings of the large and growing cheating literature and their application to the real world scenarios they address.

We thus gather survey-experimental evidence from a representative sample of the Italian population offering an insight into whether systematic asymmetric (conditional) shifts in behaviours and in norms of behaviour can be observed as response to variation in information about tax evasion at the top and at the bottom of the income distribution.

2 Experimental strategy

The experimental component of our research design consists of the random provision of information about the estimated incidence of tax malpractice in Italy to subgroups of a representative Italian online survey panel. We aim at systematically varying whether the respondents receive high vs. low estimates of tax malpractice in lower vs. higher portions of the income distribution.

To construct our information intervention, we first collected data from a survey of economists based in Italian research departments (our "experts survey" henceforth). We were able to contact via email 470 out of the top 500 leading economists according to the RePEc ranking list on the 11th of November 2019.[1] We provided them with a link to a Qualtrics survey we designed in which we asked them for their personal estimate of the share of total income that remains undeclared by each of the income quintiles of the Italian income distribution, by the top 10% and 1% income earners.[2] We then classify the responses obtained according to whether they provide relatively high or low estimates of the share of undeclared income for the bottom quintile and the top 10% of income earners. The range of variation in estimated undeclared income is then used to construct the information conditions we provided the respondents to the main survey. Specifically, to construct a *high* (respectively, *low*) estimate of undeclared total income by a given income quantile we take the mean of the estimate provided by the group of experts estimating a share of undeclared total income *above* (*below*) 50% for that same quantile. This strategy allows us to truthfully inform the respondents that "some of" the surveyed economists estimate that the "bottom" and "top" income earners do not declare the computed average shares of their total income. Crucially, as we provide our respondents with information about estimated undeclared total income for both top and bottom income earners, we ensure that both estimates originate from the same group of experts. As will be clear from what follows, our information conditions cover all configurations of high and low estimates for top and bottom income earners.[3]

The average estimated shares differ across the subgroups of experts. As can be seen from Table 1, however, all high estimates are clustered between 61 and 66%. Similarly, all the low estimates are clustered between 23 and 28%. Therefore, in order to provide all respondents with identical information about high and low estimates,

[1] https://ideas.repec.org/top/top.italy.html#authors

[2] The response rate was approximately 22% on the third day after running our experts survey (14th of November 2019), totalling 105 responses over 470 contacted economists. We use these responses to construct our information conditions.

[3] We carefully specify in the survey that the information provided originates from a subset of a number of interviewed experts.

we inform them that the estimated share of undeclared total income is "more than half (around 65%)" or "less than half (around 25%)".

2.1 Information conditions

We randomly assign the respondents to one of 5 information conditions in a 2x2+1 design. To provide a clear baseline, we include a "neutral condition" in which respondents will read a neutral sentence only generically referring to tax malpractice.[4] The other conditions vary whether respondents are informed about:

- a *low* estimated share of undeclared income among both the *top and bottom* income earners (Condition *LL*),
- a *high* share in *top and bottom* income earners (Condition *HH*),
- a *high* and a *low* share in respectively *top and bottom* income earners (Condition *HL*),
- a *low* and a *high* share in respectively *top and bottom* income earners (Condition *LH*)

The data used to generate the information conditions is presented in Table 1 below.

Subgroup of surveyed experts	**Quantile of the income distribution**		**Estimated share of undecl. total income**	**Associated condition**
Subgroup 1:	Top 10%:	$> 50\%$	61.14	Condition HH
	First quintile:	$> 50\%$	62	
Subgroup 2	Top 10%:	$> 50\%$	62.19	Condition HL
	First quintile:	$< 50\%$	25.14	
Subgroup 3	Top 10%:	$< 50\%$	27.28	Condition LH
	First quintile:	$> 50\%$	66	
Subgroup 4	Top 10%:	$< 50\%$	25.63	Condition LL
	First quintile:	$< 50\%$	23.24	

Table 1: Shares of total income undeclared by the first quintile and top decile of income earners in Italy estimated by the surveyed experts.

[4] We purposefully avoid using the terms "evasion" or "avoidance", as we do not wish to restrict our investigation to either of the two, nor do we wish to force our respondents to take a stand on the legal or moral dimension of the two practices. Our wording allows instead respondents to freely interpret our information in either way and without any consequence on our research objectives.

The experimental information is conveyed to the respondents by means of video clips which they visualise in the course of the survey. Stills of the clips are provided in Figures 1 to 5.
Each of the videos begins with a statement concerning how tax malpractice is a theme recurrently discussed in the media. Only the video for the Neutral condition then continues immediately with an invitation for the respondent to proceed with the survey. All the other videos inform our respondents of the estimated incidence of tax malpractice among top and bottom income earners as described above.
In all information conditions the order in which information about top and bottom income earnings is presented is randomised to control for order effects. Moreover, immediately after having viewed the video clips, respondents are asked to restate the information just received, and must do so in order to proceed with the questionnaire. They are in this case given the opportunity to re-play the video. This way, we both ensure that any inattentive respondents will be pushed to go back to the videos, and we obtain information to be used as a manipulation check.

Condition Neutral The information provided in the Neutral condition only generically refers to tax malpractice without making reference to its estimated incidence in the population.

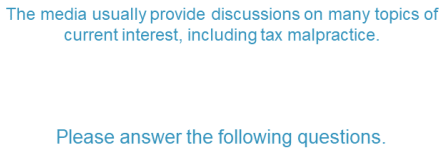
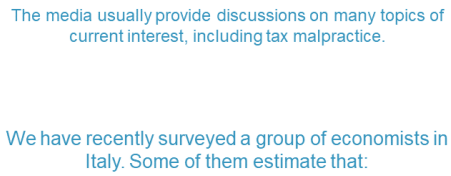
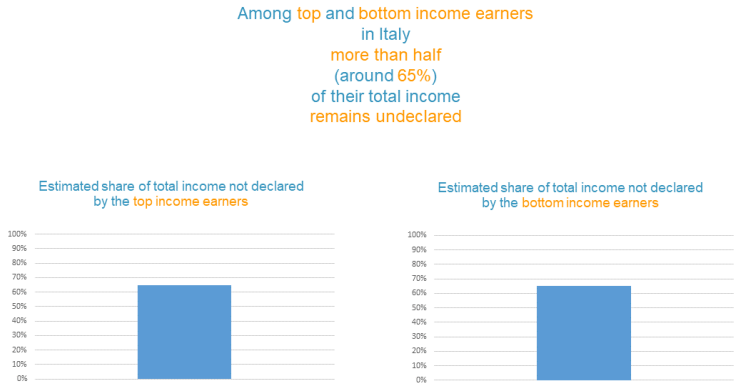


Figure 1: Information provided in the Neutral condition

Condition HH Figure 2 displays the information given to the respondents assigned to Condition HH. This group of respondents receive information about the fact that the estimated proportions of total income that remain undeclared in Italy by both top and bottom income earners are relatively high.



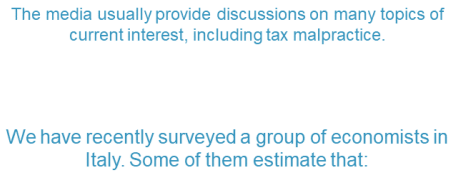
(a)



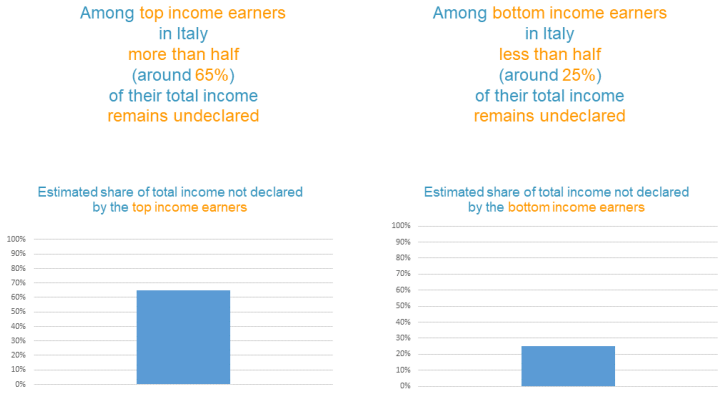
(b)

Figure 2: Information provided in Condition HH

Condition HL Figure 3 displays the information given to the respondents assigned to Condition HL. This group of respondents receive information about the fact that the estimated proportion of total income that remains undeclared in Italy by top income earners is relatively high, and that the estimated proportion of total income that remains undeclared by bottom income earners is relatively low.



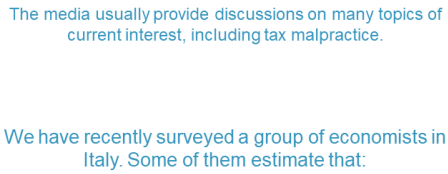
(a)



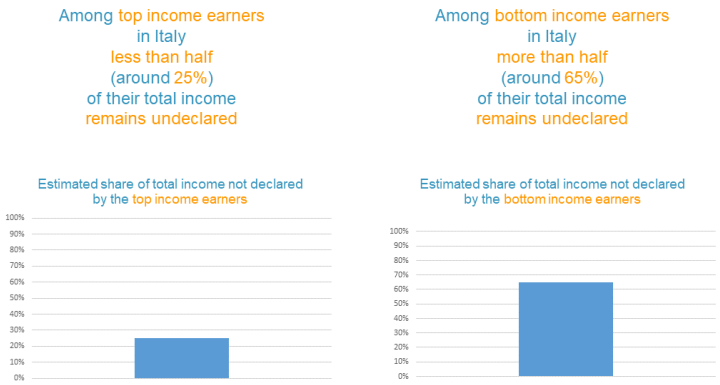
(b)

Figure 3: Information provided in Condition HL

Condition LH Figure 4 displays the information given to the respondents assigned to Condition LH. This group of respondents receive information about the fact that the estimated proportion of total income that remains undeclared in Italy by top income earners is relatively low, and that the estimated proportion of total income that remains undeclared by bottom income earners is relatively high.



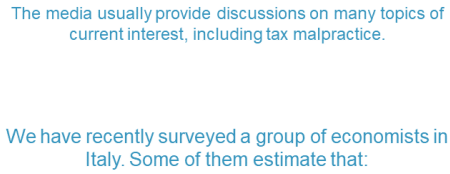
(a)



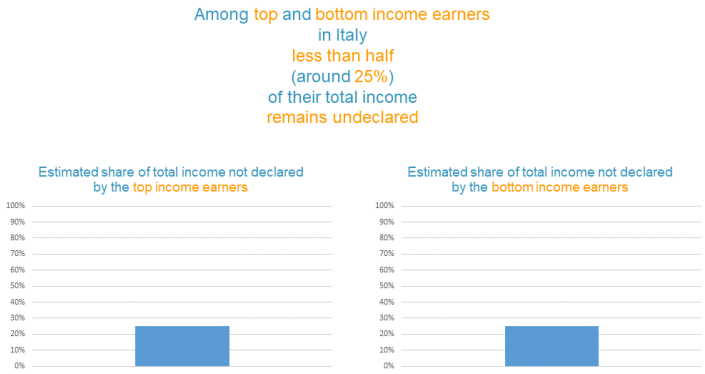
(b)

Figure 4: Information provided in Condition LH

Condition LL Figure 5 displays the information given to the respondents assigned to Condition LL. This group of respondents receive information about the fact that the estimated proportions of total income that remain undeclared in Italy by both top and bottom income earners are relatively low.



(a)



(b)

Figure 5: Information provided in Condition LL

3 Empirical strategy

3.1 Variables

3.1.1 Outcome variables

Our primary outcome of interest is the relationship between cheating behaviours and our information conditions. Following a large body of recent literature, our behavioural outcome variable measures cheating behaviours towards the experimenter (e.g. Fischbacher and Föllmi-Heusi (2013), Kocher et al. (2017)): After receiving one of the above described information treatments, the respondents have to report the outcome of a "lottery" visualised on screen. The video will display the outcome of a six-faced fair die roll. Respondents will be asked to report the visualised outcome and are informed that in case the reported outcome of the die roll is the number "6", they will receive an additional payment of 25 or 50 points (respectively one-third and two thirds of the fixed participation payment), while any other reported outcome will result in no additional payment.[5,6] As the additional payment is conditional on the self-reported outcome of the die roll, respondents have a clear incentive to misreport the outcome. The distribution of reported outcomes can be ex-post contrasted with the implemented distribution of outcomes displayed (i.e. that of a roll of a fair die) such that the incidence of cheating can be measured and compared across conditions. Moreover, as we know which outcome was displayed on video, we can detect cheating at the individual level.[7] We can thus construct an indicator taking value 1 if the respondent has cheated and zero otherwise, given that an opportunity to cheat existed. The following text is displayed to the respondents:

[5]The survey company we hired to distribute our survey will administer the payments according to their standard practices. More details in Section 3.2.

[6]As detailed in Section 3.3, we are currently running a pilot survey to test the size of the incentive.

[7]We stress out in the survey that we collect only anonymous data.

"You will now see one of six videos, randomly selected by the software.

Each of the videos displays the result of a six-faced die roll. Each outcome is displayed in a video (one outcome per video).

Because of this, ***you can visualise any of the outcomes, from one to six, with equal probability****, just as if a real die were rolled.*

Your task will be to report the outcome of the die roll.

Based on the outcome you report, you might get an additional payment. If you report that your outcome was 6, you will receive 25(50) additional points. *Otherwise, you receive no additional points.*

You will have the opportunity to replay the video."

Further, we will elicit norm shifts along the lines of Krupka and Weber (2013) by asking respondents to guess (against additional payments) the modal rate of agreement/disagreement to questions on the appropriateness of questionable behaviours (tax evasion, claim of underserved benefits, free riding on public transport and bribery) elicited in the World Values Survey for Italy, wave 2005. For exact guesses, the respondents receive an extra monetary incentive equal to 15 points (one-fifth of the fixed participation payment).
Finally, we elicit the respondents' unincentivised opinion of commonly debated topics of general interest: whether wealth can be accumulated only at others' expense, the value of hard work for life success, the importance of redistribution, generalized trust, the likelihood of being exploited by others and the appropriateness of current personal and general tax burden in Italy.

3.1.2 Covariates

We elicit respondents' prior beliefs about tax evasion rates among the rich and the poor in Italy. This will allow us to classify respondents into different groups, depending on whether their beliefs have been shifted upwards or downwards by the information condition they have received.
In addition, we elicit respondents' prior *and* posterior beliefs about tax malpractice among top and bottom income earners. This will enable us to test how well our information treatments have worked in shifting/focussing people's beliefs (see Martinangeli and Windsteiger (2019)).
Furthermore, we elicit in which province respondents live, whether they and their parents were born in Italy, their ethnicity, political orientation, voting behaviour,

media consumption, education, household income, household size and employment status.
In order to gauge the quality of the gathered responses we include a "trick question" in the middle of the survey, and will exclude respondents who got that question wrong. In addition, we expect that respondents prone to inattention will be pushed to pay attention upon reading a question explicitly mentioning its purpose of "attention checker".[8]

3.2 Sample, sample size and power

We aim at targeting 800 individuals per information condition in the survey, totalling 4000 respondents from a representative sample of the Italian population (representative with respect to gender, age and income). The panel, the distribution of the survey and the payments are administered by the survey company Respondi.[9] As further detailed below, we restrict our investigation to respondents who had an opportunity to cheat on their reporting task (i.e., their random draw would yield no additional payoff to them unless they falsely report a winning outcome). As the winning outcome occurs with a probability p=1/6 we obtain an expected sample size per condition of 5/6*800=666 respondents with an opportunity to cheat. The following power computations refer to pairwise comparisons of cheating rates across information conditions. These computations yield a minimum detectable upward effect size in cheating behaviour (proportion of winning outcomes reported) of delta=0.07 over an assumed baseline proportion b=0.5 at power π=0.8.[10]

3.3 Pilot survey

We are conducting a pilot study (500 respondents) in order to optimally design the extra monetary incentive that respondents get for reporting (truthfully or untruthfully) that they saw the winning video with a die rolling 6. If the incentive is too small (large) compared to the overall payment they get for participating in the study, there might be too little variation in cheating rates because nobody (everybody) finds it worthwhile to cheat (assuming that respondents have an intrinsic cost of cheating).

[8]Notice that we inform the respondents at the very beginning that if screened as inattentive they may not receive compensation.

[9]https://www.respondi.com/EN/

[10]Notice that the above assumed baseline proportion allows us to be as conservative as possible in our power computation, as it is the one associated with the largest variance. Fixing the effect size, the resulting power increases for more extreme values of the baseline proportion (or equivalently, the minimum detectable effect decreases for power fixed at π=0.8).

We test incentives equal to two-thirds (50 points) and one-third (25 points) of the participation compensation.

3.4 Hypotheses

Our primary focus of interest is the relationship between the information conditions provided and the rates of cheating in the reporting task. We will hence formulate hypotheses related to this outcome variable (cheating rate) and investigate the secondary outcome variables in support and generalization of our main findings.
The overarching hypothesis is that exposure to higher estimated tax evasion rates will cause an increase in the cheating rate in the reporting task compared to exposure to lower tax evasion estimates.

Hypothesis 1. *Cheating increases with the reported tax evasion estimates.*

Furthermore, our design allows us to capture asymmetries in the impact of estimated tax evasion rates according to the income bracket for which these rates are reported. Specifically, we hypothesise that the effect on cheating rates will differ according to whether estimated evasion rates increase in high compared to low income brackets. Two alternative predictions can be formulated. First, cheating rates might be higher if high tax evasion rates are reported for high income brackets compared to low income brackets. This hypothesis rests on the fact that higher income brackets have a greater capacity to contribute to public welfare and public good provision, with a lower relative impact on private consumption. This hypothesis is aligned with the findings in Martinangeli (2017). Second, and conversely, cheating rates might be higher if high tax evasion rates are reported for low income brackets compared to high income brackets. Low income brackets, relying more heavily on public support and social welfare systems, might be expected to pay their fair contribution to their financing. This leads to the formulation of two alternative hypotheses.

Hypothesis 2 (Asymmetries)**.**

a. Cheating rates are higher when high estimated tax evasion occurs in high income brackets compared to low income brackets.

b. Cheating rates are higher when high estimated tax evasion occurs in low income brackets compared to high income brackets.

3.5 Heterogeneity

We will study heterogeneity in cheating responses to tax evasion using the background information on our respondents collected in the survey. Specifically, we will test whether there are differences between the following groups: Political orientation (self-reported on a scale from 1 to 10, with values 1-3 indicating left-wing, 4-7 indicating centre, and 8-10 indicating right-wing), gender, income (two groups: below and above median, and alternatively three groups: below poverty line, middle-class and top ten 10%) and education (college vs. no college).

3.6 Specifications and analysis

3.6.1 Hypotheses 1 and 2

Hypotheses 1 and 2 will be tested as follows. Our dependent variable is an indicator taking value 1 if a respondent cheated and zero otherwise, given that a cheating opportunity existed. Denote the event of a winning random draw as D=1 and the complementary outcome of a losing random draw as D=0. We model the probability that respondent i will cheat as a function of the information condition they received and of a number of controls listed in Section 3.1.2 conditional on D_i=0:

$$\begin{aligned} Pr(L_i = 1 | C_{iNN}, C_{iLL}, C_{iLH}, C_{iHL}, C_{iHH}, X_i, D_i = 0) = \\ \Phi(\alpha + \alpha_{LL} C_{iLL} + \alpha_{LH} C_{iLH} + \alpha_{HL} C_{iHL} + \alpha_{HH} C_{iHH} + \beta' X_i + \varepsilon_i) \end{aligned} \tag{1}$$

where L_i is an indicator variable equal to 1 if respondent i has cheated given that the random draw resulted in no additional payoff (scope for cheating exists). The indicators $C_{i..}$ represent our conditions, with the Neutral condition C_{iNN} serving as excluded category, and where C_{iLL} takes value 1 if respondent i was in Condition LL and similarly for the other conditions. X_i represents a vector of individual and regional covariates. We fit the model using the cumulative distribution function Φ of the standard normal distribution.

We can then test for our Hypotheses as follows:

Hypothesis 1

$$H0 : \alpha_{LL} < \alpha_{HH} \qquad H1 : \alpha_{LL} \geq \alpha_{HH}$$

Hypothesis 2a

$$H0 : \alpha_{LH} > \alpha_{HL} \qquad H1 : \alpha_{LH} \leq \alpha_{HL}$$

Hypothesis 2b

$$H0 : \alpha_{LH} < \alpha_{HL} \qquad H1 : \alpha_{LH} \geq \alpha_{HL}$$

Notice that an equivalent way of performing these tests is that of excluding the neutral condition subsample and writing Equation (1) relative to, say, C_{iLL} which would serve as the excluded category, and accordingly rewriting the hypotheses. This is possible because our hypotheses do not involve comparisons of any of our information conditions with the neutral one. We will perform such analysis as a robustness check of our results.

3.6.2 Heterogeneity analysis

Heterogeneity analysis will be performed along the political orientation, income, education and gender dimension.

3.6.3 Multiple Hypotheses Testing

We will correct our p-values for multiplicity according to List et al. (2016).

References

Alstadsæter, A., Johannesen, N., and Zucman, G. (2019). Tax evasion and inequality. *American Economic Review*, 109(6):2073–2103.

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

Fischbacher, U., Gächter, S., and Fehr, E. (2001). Are people conditionally cooperative? evidence from a public goods experiment. *Economics letters*, 71(3):397–404.

Frey, B. S. and Torgler, B. (2007). Tax morale and conditional cooperation. *Journal of comparative economics*, 35(1):136–159.

Garside, J. (2016). A world of hidden wealth: why we are shining a light offshore. *The Guardian*.

Kocher, M. G., Schudy, S., and Spantig, L. (2017). I lie? we lie! why? experimental evidence on a dishonesty shift in groups. *Management Science*, 64(9):3995–4008.

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.

List, J. A., Shaikh, A. M., and Xu, Y. (2016). Multiple Hypothesis Testing in Experimental Economics. *NBER Working paper series*.

Martinangeli, A. F. M. (2017). Do what (you think) the rich will do: inequality, belief formation and group identity in public good games. *Working paper series of the Max Planck Institute for Tax Law and Public Finance No. 2017-6*.

Martinangeli, A. F. M. and Windsteiger, L. (2019). Immigration vs. poverty: Causal impact on demand for redistribution in a survey experiment.