Do people contribute to punish evaders ? *

Cécile Bazart † Aurélie Bonein ‡

March 15, 2015

Prepared for the "ASFEE Annual conference"
June, 15-16 2015, Paris
First draft – Please do not quote without permission

Abstract

It is now well established that people differ with respect to their tax compliance degree. In this laboratory experiment, we investigate whether individuals are willing to contribute, even in a small proportion, to a fund used to increase deterrence of evaders. This voluntary contribution scheme is composed of several experimental treatments that differ as to the degree of freedom left to contributors in their funding and in the levels of contributions in case they are predefined. We first find that a significant proportion of subjects contributes to this fund but less frequently when the cost of the contribution is increased. Secondly, contributors are characterized by a relatively higher compliance level, their reported income is above the average reported income of other group members. Thirdly, as expected the impact on the willingness to contribute is negative, respectively positive, for risk aversion and inequity aversion. Additionnally we have tested the impact of social norms as an additionnal potential explanation to contributions. We conclude that when: (i) the difference in declarations is large, (ii) the declaration made by the others is low, (iii) or when the situation of declarations is seen as inappropriate; the decision to contribute to the fund is seen as a very socially appropriate decision. Finally, looking at reported income after contribution decisions, estimation of difference-in-differences models highlights a strong causal effect: reported incomes are higher in groups in which the audit probability has increased thanks to individuals' contributions.

Keywords: Behavioral economics; Tax compliance; Social norms
JEL Classification:C91; C72; D03; D6

*We acknowlegde financial support from the University of Montpellier 1 under its Bonus Quality Research Program. As usual, all errors remain our own.

†LAMETA UMR 1135 - University Montpellier I, F-34000 Montpellier, France
‡CREM, University Rennes 1.F-35000 Rennes, France

1 Introduction

In the extended literature about tax evasion decisions much has been done to explain individual evasion choices but there is still a lot to do to highlight the social dynamics of evasion. Research shows that taxpayers' decisions are influence by those of their fellow citizens (see **?????**) and that individual decisions are somehow conditional on what the others do, either in terms of frequency or extent of evasion.

There are several ways to explain social interaction in evasion decisions. In some cases the line of explanation lies in unfairness arising because of some taxpayers' decisions to underreport and evade their taxes while others pay honestly. For **?**, taxpayers' intrinsic motivation to pay taxes decreases the more the others evade. In previous work, **?** showed that negative reciprocity existed between taxpayers and explained in part evasion activities. The aim of this experimental work was to dissociate the vertical and horizontal relationships co-existing in any tax evasion decision. Firstly, a vertical relationship opposes taxpayers and the State. In cases where the tax system is unfair taxpayers are expected to evade more. To test this causal effect, identical taxpayers, at least from a fiscal perspective, were charged with different tax rates without justification of any sort. The authors confirmed that those who suffered from this vertical inequity increased their tax evasion activities while those who benefited from it reduced their evasion activities. Secondly, a horizontal relationship links taxpayers to taxpayers. Again considering identical taxpayers any difference in evasion activities introduces a bias, in terms of fairness within the population. Social interaction was introduced displaying information on the average declaration of other group members, i.e., fellow citizens. Subject could then observe if on average the others cheated more or less than they do. The authors confirmed that those who suffered from this unfairness increased their tax evasion activities while those who benefited from it reduced their evasion activities. Social interaction impact on evasion decision is then confirmed and hopefully shown to work in both directions, that is: worsening or improving tax collection. This last result on horizontal dynamics of evasion motivated the present paper. Note that the main result of this study was provided in the last treatment where both vertical and horizontal unfairness co-existed. In such context and in all cases the horizontal effect dominated the vertical one. From there we started considering as crucial the understanding and restoring of this horizontal link. Another explanation of taxpayer's reaction to their fellow citizen's declaration decision lies in the desire to comply with social norms. Considering, as a starting point, that the accepted social norm supports that citizens should pay honestly their taxes, as a consequence examples of evasion activities may weaken this norm. This is in line with the branch of studies focusing on tax morale, conformity to social norms and their perception. **?** showed that taxpayers estimate other's acceptance of tax evasion as being greater than their own and decrease their compliance level to conform to their misperception of social norm. Then the broken window effect, described by psychologists seems to be at

work. Information on the level of evasion and fellow citizens' underreporting decisions becomes then crucial to the overall dynamics of evasion. Of course the population of taxpayers is marked with heterogeneity in moral types: some taxpayers are intrinsically honest, some cheat in a small measure, some are inveterate evaders and are reluctant to any payment of taxes; but the part of honest types may decrease in the total population the more information is given on evaders. Evolutionary models and agent based models describe this alteration in proportions of taxpayers' type within the whole taxpayers' population (see **????**).

In this paper, in line with **?**, we propose a simple mechanism to support honest types and allow them to express their negative horizontal reciprocity against evaders in a different way. We assumed that being able to punish evaders could restore taxpayers' incentive to declare truthfully their income. To do so, we introduce in the experiment an opportunity for taxpayers to contribute to a special fund devoted to tax evasion deterrence. When the sum of contributions to that special fund exceeds some threshold the audit probability is increased. This way, we thought we gave them an alternative to evasion and, at the same time, we may increase the tax authority potential performances relative to tax collection at constant cost. We assume that contributing to the fund may turn out to be the expression of what they believe the social norm should be and gives them an opportunity to defend that norm by punishing evaders. To reach our objective we built a tax experiment where subjects repeat the same declaration decision over 20 periods. At the end of each period, information on the average declaration of the 5 other members of the group is displayed to each subject. The tax game differs from **?** in that at the end of period 10, subjects are informed that they can contribute to a special fund and that amounts collected through that mechanism will be used to increase deterrence against evaders. If contributions to that special fund are high enough then the audit probability is increased for the remaining periods of the game. Experimental treatments vary in the amount of contribution and whether this latter is fixed or variable. We conduct 4 treatments in a between subject design. In the first treatment, Free treatment, subjects are free to pay the amount they want to contribute to this special fund. In the three remaining treatments, we fix the level of contributions to 2, 4 and 6 points which we call Low, Middle and High treatment respectively. We set that the audit probability increases as soon as the threshold of 6 points is reached. It results that a contribution of 2 points requires that half members of a group contribute, a contribution of 4 points requires that 2 out of the 6 members contribute while a contribution of 6 points by one subject is sufficient to increase the audit probability. This was made to analyze the effect of an increase in the cost of contribution on the willingness to contribute to the funding. To be able to provide behavioral explanations of contributions decisions, subjects take part to an experiment dedicated to elicit indivisual risk attitude (**?**) and another for inequity preferences (**?**). Finally, we have adapted **?** social norms elicitation procedure to further analyze decisions of our initial sample of subjects in the tax game.

Our preliminary results show that a non negligible part of subjects contribute to this funding and this frequency decreases as the cost of contribution increases. Second, subjects who contribute are those who report an income higher than the average of income reported by other subjects with which they are paired. Third, risk aversion has a negative impact on the willingness to contribute while inequity aversion has a positive effect. Social norms explain the best contribution decisions: when the difference in declarations is large, the declaration made by other subjects is low or when the situation of declarations is seen as inappropriate, the decision to contribute to this funding is seen as a very socially appropriate decision. Finally, looking at reported income after contribution decisions, estimation of difference in differences models highlights a strong causal effect: reported incomes are higher in groups in which the audit probability has increased thanks to individuals' contributions.

2 Related literature

From the beginning of the 70s and the seminal work of **?** considerable research efforts have been devoted to the analysis of deterrent policies to fight tax evasion. Besides the analysis of the determinants of evasion decision a growing set of research studies focused on tax compliance understanding and enlightened the importance of social interactions.

Firstly, this social component is very important in that it fuels the overall dynamic of tax evasion (see, for instance, **??**. Some early research focused on conformism, tax morale, compliance with social norms and reciprocity concerns (e.g., **?**). **?** showed that under a strong social norm of honesty, deviation by evasion activities resulted in an increasing disutility for the taxpayer. In line **?** underlined a decrease in taxpayers intrinsic motivation to declare truthfully the more the other evade. In a survey about tax morale, **?** stressed that individuals with tax evaders as friends were more likely to evade themselves. Moreover,**?** found no counterbalancing effect of information on high compliance examples on this dynamic. Recently, **?** observed positive reciprocity that implied an increase in reported income when higher declaration levels of others were observed. Differences in individual moral codes, ethics, propensities to conform explain differences in evasion decision (e.g., **?????**) but a generalized result is that information on tax evasion activities could increase evasion. Whatever channel explains the increase in evasion: either pushing honest taxpayers towards evasion or neutralizing guilt feelings of those who already evade these studies in most cases underline a negative social dynamic.

Secondly, this social component can be used to deter evasion. Several studies analyze the effects of punishment on contribution levels, mainly in public goods games. Some deal with monetary punishment some with nonmonetary ones like public disclosure of information or identity of evader to discourage non compliance. The experimental literature offers several examples where individuals sanction, at positive costs, those who deviate from the established social norm (see, for example, **???**).

Societies use various forms of incentives to facilitate cooperation to the benefit of all. Incentives differ in that they are either institutionalized or voluntary[1] and either rewarded or punished. Since the seminal works of **?**, **?** and **?**, it has been demonstrated that voluntary punishment can be effective to deter free riding in different types of public good game.[2] Interestingly **?**'s experimental work reports that subjects punish even when they cannot alter the current distribution of payoffs but with a positive correlation to the observed inter-individual inequity level. **?** analyze the motivations of punishment and confirm its expressive nature. As a consequence, utility from punishment comes from the personal act of punishing much more than the monetary incentives associated to it. In addition, **?** found that an additional threat of punishment increases contributions of participants even before the punishment decision was taken. **?** in an experiment comparing the effects of voluntary rewards versus punishment scheme insist on the differences in motivations in each case. They state that whether self-interested strategic considerations play a larger role in the case of rewards it is very different when the option was that of voluntarily punishing free riders. In the latter case, intrinsic motivations are more important. Also, and of interest relative to tax matters, **?** test robustness of punishment efficiency when groups are enlarged. They found that despite the rise in coordination problems inherent to larger groups punishment effectiveness is increased. Of course institutionalized punishment has also been shown to be efficient (see **?** for a review).

Building off this, the experimental literature has shown that the ability of groups to establish rules for individual behavior and sanction offenders often a critical element in improving efficiency in such situations even though sanctioning is likely to be neither automatic nor costless to the enforcer. But it has also been shown that social disapproval could leads to the same outcome. The analysis of norm processes in tax compliance,**?**, focused besides individual norms (individual ethics, morality of tax compliance) on the role of social norms in tax reporting decisions. Social norms can be thought of as a jointly-recognized agreement regarding appropriateness or inappropriateness of behaviors (**????**). In this paper, **?** uses a survey with Australian citizens and showed that social norms influenced behaviors of those who identify strongly with the group of people to whom the norms are attributed. Afterwards, **?** detailed that perceived norms causally affect tax compliance partly mediated by their effect on individual norms of tax ethics. But on the other hand, tax compliance also affected the perception of norms. The impact of norms (personal and social) appeared to be quite complex. **?** incorporated tax morale as a social norm for compliance in a model of tax evasion decision. The strength of the norm depending on the share of evaders in the society, individual decision appears to be conditional on what the others do. Thus policies aiming at corroborating beliefs about high compliance level became tools to reduce the evasion level.[3] On the basis of question-

[1] Such incentives are also called formal and informal (**?**), or centralized versus decentralized (**?**).

[2] See, for instance, **????**.

[3] The use of norms as a psychic cost modifying individual choices is also the issue of **?** related to tax avoidance activities.

naire response by Italian and UK students, **?** dealt with the way social norms frame the decision to pay taxes. They found out that differences in education (economics or other social science) and national tax culture both affect tax compliance. Still there is no evidence whether students with different social norms chose to study different disciplines or if studies change the social norms of students. The way social norms influence tax compliance is difficult to assess but social norms exert their influence through the reference point individual use. **?** ran an experiment to test focusing and informational influences of norms in the absence of extrinsic financial incentives. Focusing influence refers to the degree to which individuals' attention is focused on the norm. While informational influence arises from the observation of the actions of others. They found that both effects play a role in pro-social behaviors. For these reasons we complemented our analysis of evasion with social interaction and punishment by an elicitation of social norms relative to declaration decisions and punishment opportunity provided with a special fund using **?**.[4]

3 Experimental design

This study aims at highlighting whether individuals are willing to pay in order to increase the probability to punish evaders. To that purpose, our experiment designs the essential features of the voluntary income reporting and tax assessment system used in many countries. It should be emphasized that, from previous experiments, we know that risk aversion may come into play in such game, so that we also elicited our participants' risk preferences using **?**'s procedure. To analyze to what extent individuals are sensitive to inequity, and if this sensitivity affects both reported income decisions and the willingness to pay to increase the probability of punishment, we also elicited individual estimates of inequity aversion with the procedure of **?**. Finally, we have adapted **?** social norms elicitation procedure to further analyze decisions of our initial sample of subjects in the tax game.[5]

3.1 The original setup

The experimental design captures the main features of the voluntary income reporting system used in many countries. Because our focus is on tax compliance decisions, we did not introduce redistribution through the provision of public goods financed by tax payments. So tax payments are not transferred to the taxpayers in any way.

At the outset of the game, participants are presented with a screen informing them

[4]Earlier examples of social norms elicitation of some help were also **?** and **?**.

[5]Because the main purpose of our experiment concerns (i) individuals' declaration, (ii) conditions that favor individuals' financial participation to the increase in the probability to punish evaders, and (iii) the impact of the cost of punishment, the details of experimental designs regarding inequity aversion, risk preferences and social norms are reported, for exposition purposes, in Appendix A, B and C respectively.

of their individual income and the tax policy parameters (i.e., fine, audit and tax rate). Parameters are set such that full honesty is the theoretical prediction. More precisely, in each period, participants are provided with a constant income of $X = 100$ points.[6] We set the fine rate at $\pi = 350\%$ (i.e., participants have to pay evaded taxes plus a penalty of 250% of unpaid taxes in the case of an audit). Audits are assumed to be random and perfect. Audits occur with a fixed and announced probability equal to $\rho = 1/3$. We used the same random sequence of audit to facilitate data comparisons between experimental sessions. Audits apply only to current reported income, not to previous (or future) periods. Finally, the tax rate is equal to 30%.

At the time they make their decisions, participants have to determine the amount of income they will self-report to the tax authorities. They can choose to report any integer amount from 0 to 100. At the end of each period, they are informed about whether they have been audited, their net payoff and the average income reported by the other members of their group. If the participant has an under-reported tax liability and is audited, then a fine is imposed.

This process is repeated over 20 periods, each representing a tax year. Participants are informed that they will be paid their after-tax earnings, obtained in 5 out of the 20 periods, at the end of the entire experiment. The randomly chosen periods are the same for all participants in the same experimental session. The earned points are converted into Euros at the end of the experiment and the conversion rate used was 100 points=3.80 Euros.

3.2 Contribution to the special fund

From this setup, after 10 periods, subjects were told that they have the opportunity to contribute to the funding of a special fund dedicated to increase the frequency of control of declarations. The amount of contributions varies depending to the experimental treatment. But in all treatments, the audit probability is increased to $\rho = 1/2$ if the sum of individual contributions in the group is at least equal to 6 points. The level of individual contribution is withdrawn for their payoff at period 10. So the level of contribution can be seen as the willingness to pay of the individual to increase the frequency of control and so, to indirectly increase the probability to punish evaders. After making their decision, subjects have to indicate their belief about the sum of contributions in their group and then learn this latter and whether the audit probability increases for the next 10 periods. From period 11, the game restarts with an audit probability equal to $1/3$ or $1/2$ depending to the threshold of 6 points has been reached. Other tax policy parameters and composition of experimental groups remain unchanged.

It is noteworthy that at the beginning of the experiment, subjects are aware that a change may intervene after period 10 but they do not know what it is. They are only in-

[6]The term points refers to the experimental currency unit.

formed that some instructions will appear directly on their computer screen at the end of period 10, just before learning their net payoff for this period. This would help minimize the potential impact of audit at period 10 on the willingness of subjects to contribute to this fund.

As part of our experimental design, we vary the level of contribution in the experimental treatment. In the first treatment (**Free**), subjects are free to pay the amount they want to contribute to this special fund. In the three remaining treatments, we fix the level of contributions to 2, 4 and 6 points which we call **Low**, **Middle** and **High** treatment respectively. Knowing that the audit probability increases as soon as the threshold of 6 points is reached, a contribution of 2 points requires that half members of a group contribute, a contribution of 4 points requires that 2 out of the 6 members contribute while a contribution of 6 points by one subject is sufficient to increase the audit probability. This was made to analyze the effect of an increase in the cost of contribution on the willingness to contribute to the funding.

However, a free contribution lower than 6 points or contribution equals to 2 or 4 points requires that other members of the group also contribute to the funding in order to reach the threshold of 6 points. Some coordination problems may occur and could discourage some individuals to contribute. One can suppose that some people may not contribute because they believe that the threshold of 6 points will be not reached and their contribution can be likened to sunk costs. The study of individual's beliefs about the sum of contributions in their group provides a first indication about this issue but in order to deepen it, we have conducted a second series of experiments in which subjects are reimburse if the sum of contributions does not allow to reach the threshold of 6 points. This second experiment has been run for the Free, Low and Middle treatments. Overall, 138 participated in this second experiment.[7]

While one may believe that only fully honest individuals will contribute to the special fund, if it is not the case, risk aversion may also explain why people do not contribute. Individuals who are risk-averse may be reluctant to contribute to the funding since a successful contribution means an increase in the audit probability and so an increase in the chance of getting caught in case of evasion. We will have the opportunity to confirm this conjecture with the individual estimates for risk preferences used from the replication of the **?**'s procedure.

On the contrary, some behavioral characteristics such as aversion towards inequity in declarations or compliance with social norms may explain why people pay to increase

[7]This procedure is usual in provision point public good literature and known as a refund rule. **?** demonstrated that subjects contribute more when their contributions are returned in case of non provision by comparison to their decisions if contributions are lost. Experimental sessions are completed at the end of November. Details for this experiment and results will be given in Appendix in the next version; Overall we had 48 participants for the Free treatment, 48 for the Low treatment and 42 for the Middle treatment.

the probability to punish evaders. We are able to check the impact of inequity aversion on contribution decisions through the individual estimates of inequity aversion obtained when we replicate the procedure of **?**. Suggesting that individual contributions may be the result of a desire to comply with social norms means that there are collectively recognized rules of conduct that prescribe socially acceptable behaviors in a given situation (e.g., **?**). In our experiment, this means that individuals may be willing to contribute to the funding if they perceive the situation inappropriate. Since individuals learn the average declaration of their other group members, this means that if individuals believe that this latter does not comply with social norms, they may be willing to contribute to the funding in order to increase the probability of punishment of individuals who do not comply with social norms. Explaining phenomena by appealing to the influence of social norms can be problematic because of the difficulties of precisely identifying and measuring norms. In particular, often what may or may not constitute a norm is based on intuition or casual empiricism. For a more objective approach toward identifying whether social norms are relevant to our experimental environment, and whether they can explain individual contributions to the funding of the special fund, we adapt the experimental norms elicitation procedure recently introduced by **?**. This procedure uses incentivized coordination games to identify which actions are viewed as most socially appropriate in a given situation. Overall, 61 subjects take part to this experiment, with average earnings equal to 8.77 Euros.

3.3 Participants

Our experiment was conducted at the LABEX-EM, University Rennes 1. The experiment was computerized, using the software Z-TREE (**?**). Participants were recruited from a pool of undergraduate students, using the online Recruitment System for Experimental Economics, ORSEE (**?**). We ran two sessions per experimental treatment of the tax game, with 24 participants per session, so that we have 4 independent observations per session. Overall, 192 subjects participated.

Before the game started, participants were told (i) that there would be three independent experiments,[8] (ii) that money earned in the experiments would depend on their decisions and the decisions of others in their experimental group, and (iii) that they would be paid the earnings they accrue in the three experiments. It was made very clear that information about earnings obtained in each experiment would be given only at the very end of the experimental session. We set this condition to reduce the potential spillover effects of earnings from one experiment to the next.

For the risk and inequity games,[9] instructions were conveyed by a series of computer screens that the participants read at their own pace. Clarification questions were

[8]The experiment dedicated to the elicitation of inequity aversion estimates, the one dedicated to the elicitation of individual estimates of risk aversion and the tax game.

[9]The order of these two games was reversed in half of the experimental sessions

addressed after participants completed reading the instructions and participated in two practice exercises, before the game started. Regarding our main experiment, the tax game experiment, instructions were distributed and read aloud. All participants were required to answer several control questions to ensure that they understood the experimental procedures. In particular, they were required to indicate the payoffs of each player under different outcomes of the game. Answers were privately checked and, if necessary, explained to the participants, and the experiment did not start until all participants had answered all questions correctly. Each session lasted up to one hour and half. Participants earned 17.62 Euros on average (including a show-up fee of 5 Euros).

4 Results

Because the first 10 periods are similar across our experimental treatments, we naturally observe no significant difference in the average declaration across them. Subjects declare on average 65.63 points in the Free treatment, 70.02 points in the Low treatment, 70.73 in the Middle treatment and 70.60 points in the High treatment. This result highlights the robustness of the declaration decisions under the same experimental design. In what follows, we first focus on the behavioral differences between subjects who contribute to the special fund to increase the audit probability and those who do not and the determinants of such decision. In a second step, we analyze the difference in reported income in the last 10 periods between group in which the audit probability has been increased and those in which the audit probability is remained unchanged.

4.1 Participation to the special fund

4.1.1 Motivations to the contribution to the special fund

Our first research question deals with the willingness to pay of individuals in order to increase the audit probability. We observe that a non negligible part of subjects contribute to the special fund: 29.16% in the Free treatment, 27.08% in the Low treatment, 22.92% in the Middle treatment and 14.58% in the High treatment. A noticeable feature is the decrease in these frequencies with the cost of contribution (Chi-square tests: $\chi^2 = 4.4444, p = 0.035$ between the Low and Middle treatments, $\chi^2 = 21.8803, p < 0.001$ between the Middle and High treatments, $\chi^2 = 45.4737, p < 0.001$ between the Low and High treatments). We also note that the frequency of contributions when we impose a price for the contribution is lower, in comparison to the free treatment ($\chi^2 = 66.7574, p < 0.001$). As expected, in case of free contribution, there is a positive and significant relationship between the level of contribution and the average declarations made in the first 10 periods: individuals who contribute 1 point declare on average 63.93 points, those who contribute 2 points declare on average 84.38 points and those who contribute 6 points declare 90 points.

Because subjects are willing to pay to increase the audit probability, two natural questions arise: who contribute and what are the motivations and determinants of contributions? We first note that we have only few subjects who are fully honest during the 10 first periods (none in the Free treatment, 12.5% in the Low treatment, 8.33% in the Middle treatment, and 75% in the High treatment). Among them, 66.66% in the Low treatment, 60% in the Middle treatment and 10% in the High treatment take part to the funding. Conversely, there is only 1 subject who free ride in the first 10 periods (in the High treatment) and he does not contribute. This result suggests that most of individuals vary their declarations between periods, being neither always fully honest or dishonest. We will now examine the patterns of declarations to highlight the determinants of contributions. We observe that subjects who decide to contribute to the funding of the special fund declare significantly more, on average, than those who do not contribute (see Table **??**).

Table 1: Mean declaration of contributors and non contributors (in experimental points)

Treatment	Contribute	Not contribute	Mann-Whitney U test
Free	76.42	61.18	Z=-3.187, p=0.0014
	(33.45)	(40.35)	
2 Points	82.29	65.46	Z=-3.966, p$<$0.001
	(32.47)	(40.20)	
4 Points	81.13	67.65	Z=-4.139, p$<$0.001
	(35.54)	(40.37)	
6 Points	81.78	68.69	Z=-3.181, p=0.0015
	(30.82)	(36.02)	

Mean declaration for periods 1 to 10. Standard deviations in parentheses.

Even if the average of declarations may disguise some discrepancies across experimental groups, we observe from Fig. **??** that, regardless of the experimental treatment, the average declaration of contributors is almost always higher than the one of non contributors.

Figure 1: Average declaration by experimental group depending to the contribution decision

Because contributors declare more, on average, than non contributor, one can believe that subjects who decide to contribute are those for whom their declaration is above the average declaration made by their other group members. Chi-square test confirms that there exists a link between the decision to contribute and the fact that subjects declare more than their other group members in all experimental treatments ($\chi^2 = 13.4599$, $p < 0.001$ in the Free treatment; $\chi^2 = 21.9187$, $p < 0.001$ in the Low treatment; $\chi^2 = 10.5236$, $p = 0.001$ in the Middle treatment; $\chi^2 = 5.6927$, $p = 0.017$ in the High treatment). For instance, 121 declarations of contributors out of the 126 observations (i.e., 96.03%) in the Free treatment are above the average declaration of other group members. The tendency is even more pronounced in the three remaining treatments where 93.16%, 95.96% and 98.41% of declarations are above the average declaration of other group members for

Low, Middle and High treatments respectively.

The gap in declarations generates inequity between individuals and may explain why inequity averse individuals contribute to the special fund. Besides, compliance with social norms may also explain such behavior while risk aversion may explain why individuals do not contribute. To confirm these conjectures, we estimate the determinants of the contribution decisions with a probit analysis. The dependent variable is the decision to contribute or not to the funding of the special fund, and the independent variables are the declaration made by the individual at period 10, the disadvantageous inequity in declaration between the individual and the average of other group members at period 8 and 9 if they exist,[10] the advantageous inequity in declaration between the individual and the average of other group members at period 8 and 9 if they exist,[11] whether the individual has been audited at period 8 and 9, the individual's belief about the contribution of his other group members, the individual parameters of inequity aversion, an index for risk preferences, socio-demographic characteristics of the subjects (ω_i= age, gender, field of studies) and the fixed effects of the experimental treatments, risk-order experience and groups. The results are reported in Table **??**.

[10]The disadvantageous inequity refers to a positive difference between the average declaration of other group members and the one of the subject. If this difference is negative, we set a value of 0.

[11]The advantageous inequity refers to a positive difference between the declaration of the subject and the average declaration of other group members. If this difference is negative, we set a value of 0.

Table 2: Determinants in the decision to contribute to the funding of the special fund

Declaration	-0.002
	(0.004)
Disadvantageous difference in declarations at period 9	0.006
	(0.007)
Disadvantageous difference in declarations at period 8	-0.003
	(0.008)
Advantageous difference in declarations at period 9	0.025***
	(0.009)
Advantageous difference in declarations at period 8	0.034***
	(0.012)
Audit=1 if Yes at period 9	0.052
	(0.341)
Audit=1 if Yes at period 8	-0.408
	(0.342)
Belief	0.070***
	(0.026)
Index for risk preferences	-0.052**
	(0.026)
Alpha[a]	-0.164
	(0.104)
Beta[b]	-0.061
	(0.536)
Constant	-3.220
	(2.086)
Socio-demographic controls	Yes
F-test ($\omega_i = 0$)	3.07
p-value	(0.3811)
Statistics	
N	165
Pseudo R-square	0.3567
Log-likelihood	-61.5568

Notes: ***, **, * denote statistical significance at the 1%, 5% or 10% level, respectively. Robust standard errors in parentheses. Socio-demographic controls include: age, gender (Female = 1) and field of study (Non economics=1). [a]: individual parameter of disadvantageous inequity.[b]: individual parameter of advantageous inequity. See Appendix A et B for individual estimates for inequity aversion and index for risk preferences.

Results show that the current declaration has no impact on the contribution decision. As expected, those who contribute more than the average of declarations made by their other group members contribute as well as those who expect that the sum of contributions made by their other group members is high. We also observe that risk aversion (high index for risk preferences) has a negative impact on the propensity to contribute to the funding. Surprisingly, individuals who express a strong aversion towards disadvantageous inequity are less likely to contribute to the special fund. Overall, we note

that risk aversion and inequity aversion have a little effect on the contribution decisions. It follows that another alternative explanation has to be investigated: compliance with social norms.

To this end, an experiment has been conducted with new participants in which they have to provide rating for the appropriateness of the situation of declarations made by an individual i and another representative individual j. From this situation, they have to rate the appropriateness of the decision to not contribute to the special fund (No Contribution) and to contribute to its funding (Contribution).[12] It was made very clear that even if in the tax game experiment, subjects have to choose between the contributing or non contributing in the funding of the special fund, here individuals have to rate each possible choice. We convert individuals' responses into numerical scores. A rating of "very socially inappropriate" got a score of -1,"somewhat socially inappropriate" a score of $-1/3$,"socially appropriate" a score of $1/3$, "very socially appropriate" a score of 1.[13] Table **??** presents respondents' social appropriateness ratings for the No Contribution and Contribution choice for each situation of declaration. The first two columns report the situation and the corresponding declarations. The next five columns report the mean of the social appropriateness ratings (ranging from complete agreement on "very socially inappropriate" (-1) to complete agreement on "very socially appropriate" (1), and then the full distribution of responses for the No Contribution choice. The last five columns report the same information for the Contribution choice.

From results reported on Table **??**, we first observe that subjects are quite able to anticipate others' ratings because the modal response almost always receives almost half of the respondents. If we turn first to the No Contribution choice, we note that it is socially appropriate for individual i to not contribute when his declaration is lower than the one made by the representative individual or when declarations are close to each other. Another context in which the No Contribution choice is socially appropriate deals with the situation where the individual i is a full evader (situations 21 to 25). In this case, regardless of the declaration made by the representative individual, it is socially appropriate to not contribute to the special fund. If we turn now to the Contribution choice, we observe that the gap in declarations seems to matter in the social appropriateness of such decision. The rating of social appropriateness is even more high that the difference in declaration is high. It is also true when the declaration of the representative individual is low, regardless of the declaration of individual i (except when he is a full evader, i.e., situations 21 to 25). As expected, when individual i is a full evader, most of respondents believe that a contribution of individual i is very inappropriate.

[12] See Appendix C for more details.

[13] We proceed like **?**. This scoring allows to attribute a score of -1 to the least appropriate rating and a score of 1 to the most appropriate rating. Scores of $-1/3$ and $1/3$ for the two intermediate ratings are evenly spaced over the interval -1 to 1.

Table 3: Elicited norms ($N(ak)$) for No-Contribution and Contribution choices

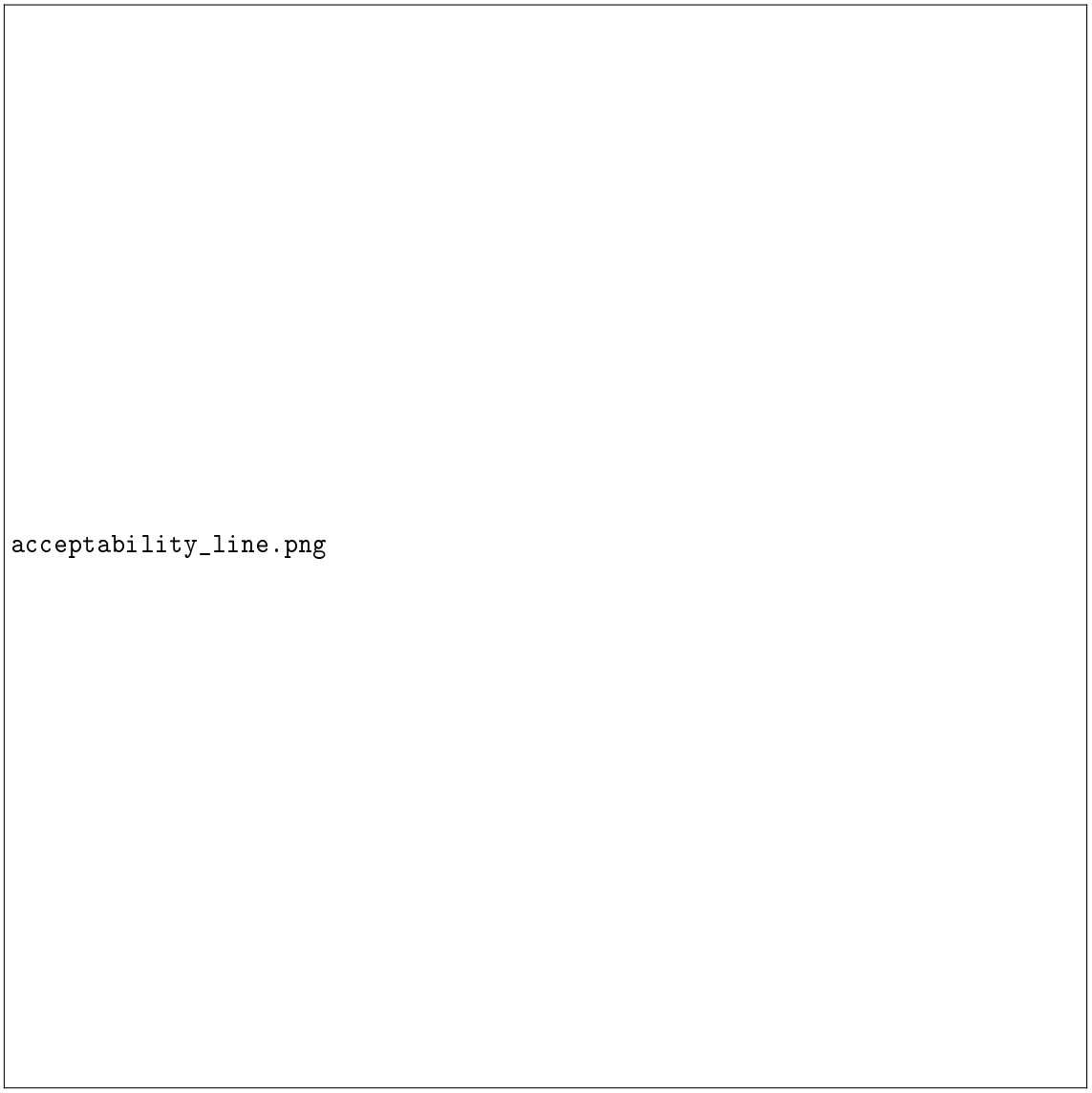
	Declarations	No Contribution					Contribution				
		Mean	- -	-	+	++	Mean	- -	-	+	++
Situation 1:	100, 100	0.7049	1.64	1.64	36.07	60.66	-0.9126	86.89	13.11	–	–
Situation 2:	100, 75	0.3443	4.92	16.39	50.82	27.87	-0.3115	31.15	42.62	18.03	8.20
Situation 3:	100, 50	0.1585	4.92	29.51	52.46	13.11	0.1913	4.92	31.15	44.26	16.67
Situation 4:	100, 25	-0.2131	18.03	54.10	19.67	8.20	0.5082	1.64	11.48	45.90	40.98
Situation 5:	100, 0	-0.4535	44.26	37.70	9.84	8.20	0.8033	4.92	3.28	8.20	83.61
Situation 6:	75, 100	0.3552	1.64	18.03	55.74	24.59	-0.5956	49.18	42.62	6.56	1.64
Situation 7:	75, 75	0.4754	–	11.48	55.74	24.59	-0.4973	42.62	44.26	8.20	4.92
Situation 8:	75, 50	0.1038	1.64	37.70	54.10	6.56	0.1694	3.28	40.98	32.79	22.95
Situation 9:	75, 25	-0.2678	13.11	67.21	16.39	3.28	0.5519	1.64	1.64	59.02	37.70
Situation 10:	75, 0	-0.4754	36.07	54.10	4.92	4.92	0.7596	3.28	1.64	22.95	72.13
Situation 11:	50, 100	0.3661	1.64	13.11	63.93	21.31	-0.4973	40.98	47.54	6.56	4.92
Situation 12:	50, 75	0.2459	4.92	16.39	65.57	13.11	-0.3879	27.87	57.38	9.84	4.92
Situation 13:	50, 50	0.3333	1.64	18.03	59.02	21.31	-0.2240	18.03	55.74	18.03	8.20
Situation 14:	50, 25	-0.2350	8.20	73.77	13.11	4.92	0.4098	4.92	9.84	54.10	31.15
Situation 15:	50, 0	-0.5410	47.54	39.34	9.84	3.28	0.7268	3.28	3.28	24.59	68.85
Situation 16:	25, 100	0.2568	9.84	18.03	45.90	26.23	-0.4426	36.07	47.54	13.11	3.28
Situation 17:	25, 75	0.1913	8.20	19.67	57.38	14.75	-0.3770	26.23	59.02	9.84	4.92
Situation 18:	25, 50	0.1913	4.92	22.95	60.66	11.48	-0.2568	18.03	62.30	9.84	9.84
Situation 19:	25, 25	0.1913	8.20	19.67	57.38	14.75	-0.1913	16.39	57.38	14.75	11.48
Situation 20:	25, 0	-0.4535	42.62	37.70	14.75	4.92	0.5737	6.56	1.64	40.98	50.82
Situation 21:	0, 100	0.1038	13.11	24.59	45.90	16.39	-0.5519	52.46	31.15	13.11	3.28
Situation 22:	0, 75	0.1585	11.48	19.67	52.46	16.39	-0.4645	45.90	36.07	9.84	8.20
Situation 23:	0, 50	0.1257	13.11	18.03	55.74	13.11	-0.4426	49.18	29.51	9.84	11.48
Situation 24:	0, 25	0.0819	14.75	21.31	50.82	13.11	-0.4208	47.54	29.51	11.48	11.48
Situation 25:	0, 0	0.0710	19.67	18.03	44.26	18.03	-0.2678	39.34	29.51	13.11	18.03

Notes: Frequency for each possible answer by situation. For declarations, the declaration at the left side corresponds to the declaration made by individual i and the declaration at the right side corresponds to the one of the representative individual. Responses are: "very socially inappropriate" (- -), "socially inappropriate" (-), "socially appropriate" (+), "very socially appropriate" (++).

To deepen the the analysis of the relationships between the level of declarations, the difference in declarations and the rating of social appropriateness of the No Contribution and Contribution choices, we represent their average appropriateness ratings for each situation in Fig. **??**. Situations 1 to 5 correspond to situations in which individual i is fully honest (his declaration is equal to 100) and we vary the declaration of the representative individual j from 100 to 0 by interval of 25. Situations 6 to 10 represents situations in which individual i declares 75 and again we vary the declaration of the representative individual j. We proceed in the same way for situations 11 to 15 with a declaration of individual i equal to 50 and situations 16 to 20 with a declaration of individual i equal to 25. Finally, in the last 5 situations, situations 21 to 25, the individual i is a full evader (his declaration is equal to 0). We first observe that, for a given declaration of individual i, the social appropriateness rating decreases with the decrease in the declaration made by the representative individual, except for some focal points corresponding to equal declarations (i.e., situation 7 with declarations equal to 75-75, situation 13 with declarations equal to 50 - 50 and situation 19 with declarations equal to 25 - 25). Second, if we ex-

clude the last five situations where individual i is a full evader, we observe exactly the same trend between the average rating of the social appropriateness of the situation of declarations and that of the No Contribution choice. This observation is strengthened by the computation of Spearman rank correlation coefficient ($\rho = 0.3059$, $p < 0.0001$). As expected, we observe two opposing trends for the average rating of the situation and that of the Contribution choice. Given that a low average appropriateness rating for the situation of declarations means that such situation is very socially inappropriate, we naturally observe in this case that a contribution to the special fund dedicated to increase the audit probability is very socially appropriate. And this is even more true that the difference in declarations is high or the representative individual is a full evader (see for instance situations 5, 10 or 15). Again, the computation of Spearman rank correlation coefficient confirms the opposite relationship between the social appropriateness rating of the situation of declarations and that of the Contribution decisions ($\rho = -0.1578$, $p < 0.0001$).

Figure 2: Average appropriateness rating for the situation of declarations, the No Contribution and Contribution choices



These observations are confirmed with results obtained with ordered logistic regressions. The dependent variable is the social appropriateness rating varying from -1 for "very socially inappropriate" to 1 for "very socially appropriate". For the independent variables, we test 3 specifications. The first specification includes the declaration of individual i and that of the representative individual j. The second specification includes the differences in their declaration. Finally, the third specification includes the social appropriateness rating for the situation. First results are reported in Table **??**.[14]

[14]Preliminary results are reported. Further investigations are in progress.

Table 4: Ordered logistic regression

	Rating for No Contribution			Rating for Contribution		
Declaration i	-0.001			0.015***		
	(0.001)			(0.001)		
Declaration j	0.026***			-0.035***		
	(0.001)			(0.002)		
Difference in declarations		-0.012***			0.024***	
		(0.001)			(0.001)	
Situation rating			0.936***			-0.465***
			(0.072)			(0.067)
Statistics						
N	1525	1525	1525	1525	1525	1525
Pseudo R-square	0.0810	0.0414	0.0451	0.1512	0.1252	0.0116
Prob $> \chi^2$	0.000	0.000	0.000	0.000	0.000	0.000
Log-likelihood	-1801.7434	-1871.3602	-1872.0055	-1774.1999	-1828.3709	-2065.8586

Notes: ***, **, * denote statistical significance at the 1%, 5% or 10% level, respectively.

Regarding the social appropriateness of the No Contribution choice (columns 1 to 3), we observe that it is positively related to the declaration of the representative individual and the appropriateness degree of the situation of declarations. Conversely, it is negatively related to the difference in declarations. As expected, the opposite relationships are found for the social appropriateness of the Contribution choice (columns 4 to 6). We note that the probability that the Contribution choice is very socially appropriate is positively related to the declaration made by individual i as well as the difference in declarations. Conversely, the more a situation of declarations is seen as socially appropriate or the lower is the declaration made by the representative individual, the lower is the probability that the decision to contribute is socially appropriate.

4.2 Consequences to the contributions to the special fund

Thanks to significant individual's contributions to the funding of the special fund, the audit probability has increased in 25% of groups in the Free and Low treatments, 37.5% in the Middle treatment and 62.5% in the High treatment.

Our first investigation refers to the "restart effect". In other words, do average declarations change between the first 10 periods and the last 10 periods? Of course, the answer has to be linked to the effective increase or not in the audit probability. From average declarations reported in Table **??**, we observe that the average declarations do not change before and after period 10 when the audit probability is remain the same, while we observe a significant increase in declarations when the audit probability has increased.

Table 5: Average declarations before and after the contribution decisions

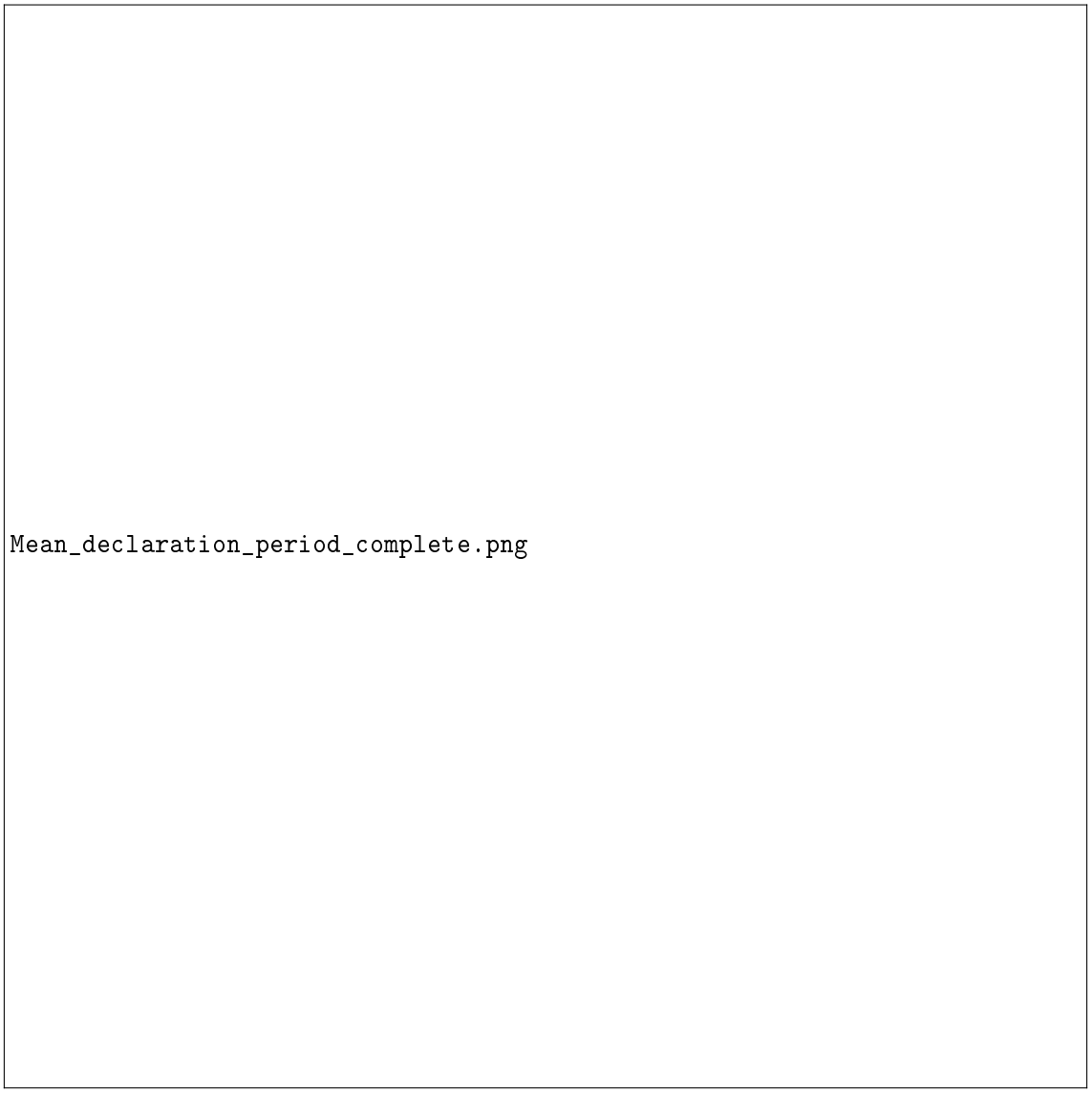
Treatment	Without increase in audit prob.		With increase in audit prob.	
	Period 1 to 10	Period 11 to 20	Period 1 to 10	Period 11 to 20
Free	65.77	67.91	65.21	77.10
Low	69.41	67.16	71.83	80.51
Middle	68.65	62.31	74.21	78.94
High	70.86	67.77	70.46	84.31

Notes: In experimental points. For periods 1 to 10, we compute the average of declarations by distinguishing individuals depending to the future increase or not of the audit probability.

As a result, in the groups for which the audit probability has been increased, individuals declare on average more than in other groups (Mann Whitney U test: $Z = -2.476$, $p = 0.0133$ in the Free treatment,$Z = -2.450$, $p = 0.0143$ in the Low treatment, $Z = -2.952$, $p = 0.0032$ in the Middle treatment, $Z = -4.855$, $p < 0.001$ in the High treatment).

By looking into possible differences in declarations across these groups, we can study what are the causes of the observed changes in tax reporting. The starting point refers to the observation of similar average declarations in the first 10 periods across groups in which the audit probability is increased from period 11 and those in which it is not the case. One can assume that average declarations are similar across groups from period 1 to 10 (supported by MW tests). But significant differences arise from period 11 as we can see from Fig. **??**.

Figure 3: Average declarations in the Control and Treatment groups by experimental treatment



To measure the exact size of the increasing audit probability effects shown in Fig. **??**, we run regressions measuring the size of the difference between the Treatment and Control groups. In what follows, we call Control group experimental groups in which the audit probability is remain unchanged and Treatment group experimental groups in which the audit probability has been increased from period 11. In this setting the only difference between the groups is whether the audit probability has been increased from period 11. In order to investigate in what extend the increase in audit probability has caused differences in tax reporting, we use the difference-in-differences approach considering those who are affected by the increase in audit probability as a Treatment group while those who are not affected as a Control group. In particular, we estimate the following specification:

Table 6: Difference-in-Differences Estimates

	Free treat.	Low treat.	Middle treat.	High treat.
Group with increased audit prob (=1 if yes)	-1.609*	0.099	1.139*	0.399
	(0.938)	(0.803)	(0.590)	(0.688)
Periods post contrib. (=1 if yes)	-0.209	-0.855***	-0.901**	-0.846*
	(0.400)	(0.280)	(0.373)	(0.451)
Increased audit prob & periods	1.557*	2.033***	1.746***	1.593***
	(0.816)	(0.458)	(0.620)	(0.566)
Index risk preferences	0.146**	-0.005	-0.059	-0.158
	(0.069)	(0.078)	(0.046)	(0.101)
Index personal risk attitude[c]	-0.434***	-0.617***	-0.368***	-0.341**
	(0.154)	(0.116)	(0.120)	(0.138)
Alpha[a]	-0.266	0.325*	0.275	0.464
	(0.250)	(0.189)	(0.278)	(0.320)
Beta[b]	-2.635	2.050*	0.777	0.991
	(2.026)	(1.130)	(0.972)	(1.486)
Audit in $t-1$ (=1 if Yes)	-3.000***	-4.181***	-3.681***	-1.788***
	(0.554)	(0.585)	(0.493)	(0.387)
Constant	3.344	0.154	4.920	11.584***
	(4.644)	(3.682)	(5.466)	(3.267)
Socio-demographic controls	Yes	Yes	Yes	Yes
F-test ($\omega_i = 0$)	1.52	3.60	0.78	1.61
p-value	(0.2217)	(0.0201)	(0.5126)	(0.2007)
Statistics				
N	912	912	912	912
R-square	0.149	0.265	0.215	0.168

Notes: ***, **, * denote statistical significance at the 1%, 5% or 10% level, respectively. Standard errors in parentheses, clustered at the individual level. [a]: individual parameter of disadvantageous inequity.[b]: individual parameter of advantageous inequity. See Appendix A et B for individual estimates for inequity aversion and index for risk preferences.[c]: Index personal risk attitude measures the individual risk preferences through answers provided in the post experimental questions. High value of index refers to risk lover subjects while low value of index refers to risk averse subjects.

$$Y_{it} = \eta T_{it} + \varphi POST_{it} + \gamma T_{it} \times POST_{it} + \omega X_i + \delta AUDIT_{it-1} + \mu_t + \nu_i + \epsilon_{it} \quad (1)$$

Y_{it} is the value of the dependent variable, the declaration, for individual i in period t.[15] POST and T are indicators of whether the observation is after the treatment and whether the individual is assigned to the Treatment group. The most interesting parameter is γ which stands for the difference-in-differences coefficient. X_i includes socio-demographic controls and individuals estimates for risk and inequity preferences. $AUDIT_{it}$ is an indicator of whether individual i has been audited in period $t-1$. μ_t and ν_i capture the period and individual fixed effects and ϵ_{it} is an error term, which captures all the other things not included in the equation, also those we cannot observe. Results are reported in Table **??**.[16]

We first observe a strong causal effect in all experimental treatments: declarations are higher in groups in which the audit probability has increased after period 10. Individuals in these groups report an income that is 139% larger than the one reported by subjects

[15]To interpret our results in percentage, we convert the dependent variable in logarithm. Due to the null value of some declarations, before transformation, we add a "noise" equal to 10^{-4}.

[16]We only report first results. Other specifications and robustness checks are in progress.

in the Control group after period 10 in the Free treatment. The trend is even more pronounced in the other treatments: individuals in the Treatment groups declare an income that is 153.5% higher than the one reported by individuals in the Control group in the Low treatment. This percentage increases up to 215.33% in the Middle treatment and 174.25% in the High treatment. Second, we find a period effect: all other things being equal, declarations are lower after the contribution decisions, especially in the Low and High treatment. Third, we note that the more individuals are risk lover, the lower are their declarations, regardless of the experimental treatment. Finally, we note that having been audited has a negative effect on reported income. This means that, after having been audited, individuals believe that they will not be audited in the next period and take the risk of reducing the level of the income they report to increase their after-tax income (bomb crater effect). Of course, the after-tax income is increased only if such taxpayers are not audited.

5 Conclusion

In line with public good literature we built an experiment relative to reporting decision including: (i) social information on what other fellow-citizens declare, and (ii) punishment opportunity. The punishment opportunity takes the form of a special fund fed by individual contributions and used to increase audit frequency. In some ways it could be seen as a collective punishment as the taxpayers' population as a whole faces the increased audit probability if the fund threshold has been reached. From this mechanism we expect honest taxpayers only to contribute. In the experiment we test separately voluntary punishment under: a free choice of the individual contribution level (Free treatment) and 3 pre-fixed contribution levels. In a first set of sessions, presented in the preliminary version of the paper, if the threshold is not reached contributors to the fund loose the given amount. But, we have run additional sessions with a refund rule which results are not yet included in the paper. We expect that, in line with the literature on provision point public good games the refund rule will help increasing contributions to the fund. Nevertheless it may well be that social norms explain in most part contributions in which case financial incentives may not dominated the intrinsic and complex ones relative to social norms.[17] To be able to distinguish these effects we ran sessions to elicitate social norms relative to declaration and punishment decisions. Our main results at this stage of the work show that a non negligible part of subjects contribute to the fund: 29.16 % in the Free treatment, 27% in the Low treatment, 23% in the Middle treatment and 14.58% in the High treatment. We also observe a decrease in these frequencies when the cost of contribution increases from 2 to 6 points. Also the mechanism works as there is a significant and positive relationship between the level of honesty in reporting activities over the 10 first periods and the level of contribution. Moreover thanks to significant

[17] Results from these sessions will be included by later.

individual's contribution to the special fund, the audit probability has increased in 25% of groups in the Free and Low treatments, 37.5% in the Middle treatment and 62.5% in the High treatment. As a consequence tax collection has been improved too. Social appropriateness of punishment is also demonstrated. The rating of social appropriateness is even higher that the difference in declaration is high. As expected, when individual i is a full evader, most of respondents believe that a contribution of individual i is very inappropriate. As specified above some data relative to this work have not yet been analyzed and it would certainly be of interest to derive the appropriateness ratio subjects elicited for evasion activities.

A Appendix

A. Experiment dedicated to the elicitation of inequity aversion estimates

A.1. Elicitation

We conduct an experiment aimed at estimating the individual parameters of inequity aversion, following **?**'s model. This model assumes that the utility of player i may be written as:

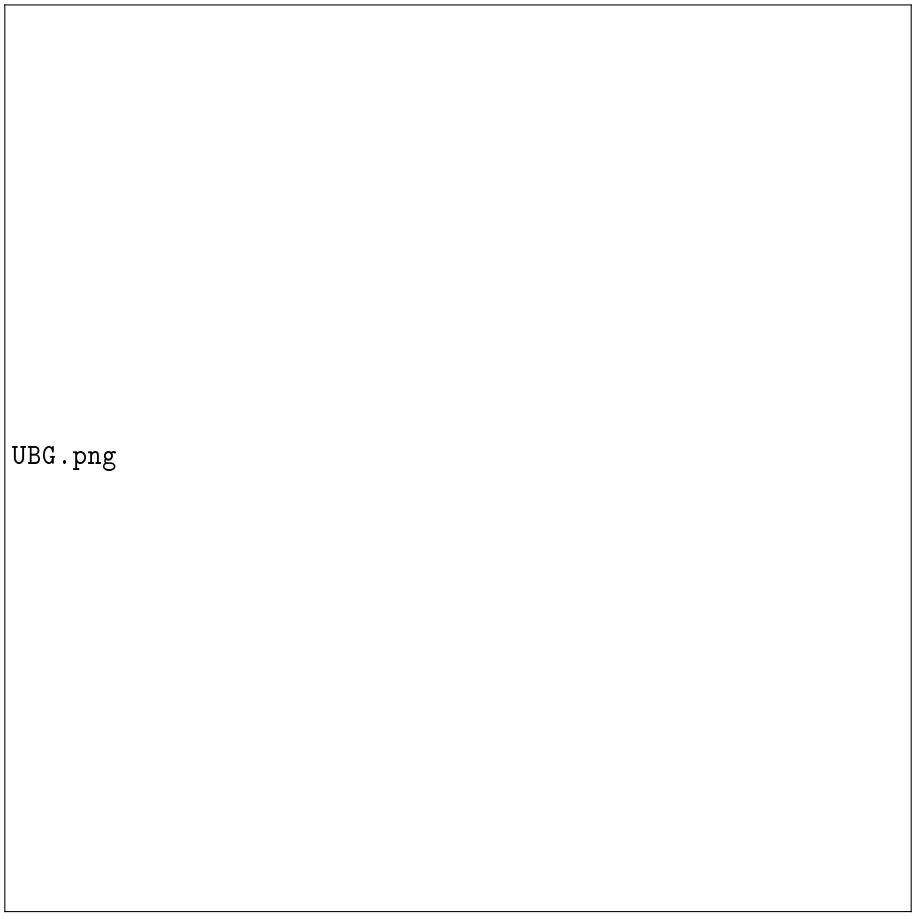
$$U_i = x_i - \alpha_i \max\left(x_j - x_i, 0\right) - \beta_i \max\left(x_i - x_j, 0\right) \tag{2}$$

where x_i is the monetary payoff of player i, x_j is the monetary payoff of player j, α_i is the parameter for disadvantageous inequity of player i and β_i is the parameter for advantageous inequity of player i. It is assumed that $\alpha_i \geq \beta_i$.

We follow the procedure of **?**, whereby subjects make decisions in two different games: an ultimatum game using the strategy method and a modified dictator game. In each game, subjects do not learn their role (for example, proposer or responder in the ultimatum game) until the end of the game.

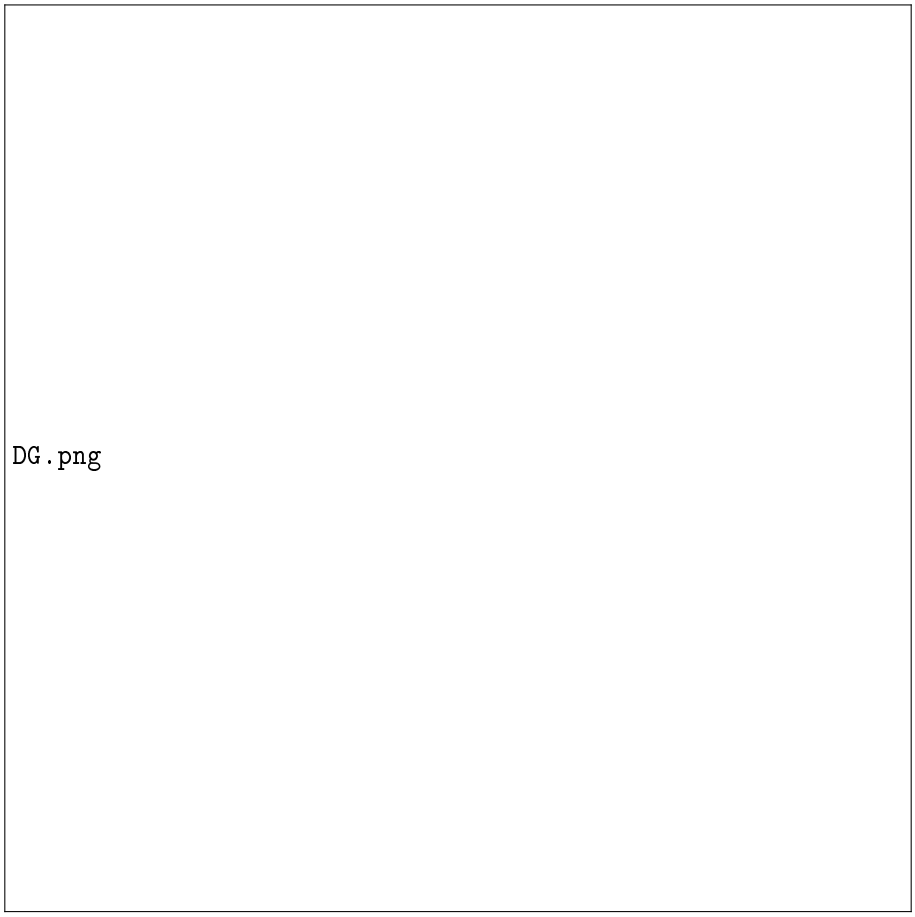
More precisely, the ultimatum game is used to elicit the individual parameter of disadvantageous inequity, α_i. In this game, the proposer must divide 20 points between himself and the responder. Next, the responder must decide whether to accept or reject the proposition. In our experiment, all subjects decide first as a proposer and second as a responder. To avoid any feedback and to elicit the complete strategy of responders, we use the strategy method; that is, responders must decide whether to accept or reject any of the 21 possible distributions (ranging from $(20, 0)$ to $(0, 20)$; see Fig. **??**). The estimation of α_i is obtained through the decisions of the responder i and corresponds to the switch point between rejecting and accepting the distribution.

Figure 4: Table for responder's choices in the ultimatum game



Regarding advantageous inequity, we use the modified version of the dictator game in which subjects must make decisions as a proposer by choosing between two distributions - a non-egalitarian one $(20, 0)$ and an egalitarian one (x_i, x_i), for 21 possibilities (ranging from $(0, 0)$ to $(20, 20)$; see Fig. **??**). The estimate of the advantageous inequity parameter, β_i, corresponds to the switch point from the unfair distribution $(20, 0)$ to the egalitarian one (x_i, x_i).

Figure 5: Table for the modified dictator game



Because we reverse the order between this experiment and that dedicated to risk preferences, we do not change the order of moves between the two games. In all experimental sessions, subjects complete first the dictator game and next the ultimatum game. Finally, subjects know that they will be paired with a different participant in these two games. Only one of the two games is used for the payment, with the conversion rate of 5 points= 1 Euro.

A.2. Results

B. Experiment dedicated to the elicitation of risk estimates

B.1. Elicitation

In this game, participants had to make 20 successive choices between 2 options. The first option is a safe option that varies from row to row, from 0 to 19 points, by increments of 1 point. The second option is a lottery that remains constant: there is a 50 percent chance of winning 0 points and a 50 percent chance of winning 30 points (see Table **??**). The switching point between the two options informs us about the risk attitude of the partic-

ipants. Because the expected value of the lottery is 15, risk-neutral participants should switch at 15. Risk-loving participants should choose the lottery when the offered safe option is higher than 15, while risk-averse participants should prefer the safe option for safe payments lower than 15. In this way, we are able to compute individual risk preferences parameters and subsequently see whether risk preferences and evasion decisions are related. Regarding participants' earnings, they are informed that, after having made their decisions, one of the rows will randomly be selected for payment and that they will be paid according to their choice. The selected row is, again, the same for all participants. In line with **?**, to ensure incentive compatibility, participants are informed that a random device will determine whether they are paid for this experiment. The chance of winning was 1/8 and the conversion rate used was 3.7 points=1 Euro.

Table 7: The 21 decisions in the risk experiment

Situations	Option A	Choose Right or Left	Option B
Situation 1	0		50% to get 0 and 50% to get 30
Situation 2	1		50% to get 0 and 50% to get 30
Situation 3	2		50% to get 0 and 50% to get 30
Situation 4	3		50% to get 0 and 50% to get 30
Situation 5	4		50% to get 0 and 50% to get 30
Situation 6	5		50% to get 0 and 50% to get 30
Situation 7	6		50% to get 0 and 50% to get 30
Situation 8	7		50% to get 0 and 50% to get 30
Situation 9	8		50% to get 0 and 50% to get 30
Situation 10	9		50% to get 0 and 50% to get 30
Situation 11	10		50% to get 0 and 50% to get 30
Situation 12	11		50% to get 0 and 50% to get 30
Situation 13	12		50% to get 0 and 50% to get 30
Situation 14	13		50% to get 0 and 50% to get 30
Situation 15	14		50% to get 0 and 50% to get 30
Situation 16	15		50% to get 0 and 50% to get 30
Situation 17	16		50% to get 0 and 50% to get 30
Situation 18	17		50% to get 0 and 50% to get 30
Situation 19	18		50% to get 0 and 50% to get 30
Situation 20	19		50% to get 0 and 50% to get 30

B.2. Results

C. Experiment dedicated to the elicitation of norms compliance

We adapt the experimental norms elicitation procedure recently introduced by **?**. As **?**, our focus is on injunctive social norms, i.e., collective perceptions, among members of a population, regarding the appropriateness of different behaviors like tax evasion and contribution to a special fund dedicated to the funding of an increase in the probability of punishment.

To embed this definition of social norms in a simple utility framework which will al-

low us to subsequently estimate the concern that individuals have for norm compliance, relative to money, we follow the formalization of **?**. We assume that an individual cares about both the monetary payoff produced by the selected action, $(\pi(a_k)$, and the degree to which the action is collectively perceived as socially appropriate:

$$u(a_k) = V(\pi(a_k) + \gamma N(a_k) \tag{3}$$

The function $V()$ represents the value the individual places on the monetary payoff, that is assumed increasing in $(\pi(a_k))$. $N(a_k)$ is an empirically measurable collective judgment that assigns to each action a degree of appropriateness or inappropriateness. Therefore, we assume that if the action a_k is viewed as socially appropriate, then $N(a_k) > 0$, while if there is joint recognition that this action is inappropriate or socially proscribed then $N(a_k) < 0$. Thus, $N(a_k)$ identifies the degree to which a specific action, a_k , is collectively perceived as one that should or should not be taken. Finally, the parameter $\gamma \geq 0$ represents the degree to which the individual cares about adhering to social norms. An individual entirely unconcerned with social norms ($\gamma = 0$) will always select the payoff-maximizing action. On the other hand, as γ increases, an individual will derive greater utility from selecting actions that are socially appropriate relative to the utility from those that are not.

The experimental procedure introduced by **?** we follow uses incentivized coordination games to identify which actions are viewed as most socially appropriate in a given situation.

We elicit social norms over possible action choices (contribution or non contribution to the special fund) across different context of declarations, from individuals who do not make choices in these contexts. Because the set of possible combinations for the declaration of individual i and the average of declarations made by his other group members is too large to be implemented, we have selected some representative situations of declarations. We have set 5 possible declarations for the individual i: 0, 25, 50, 75 and 100 experimental points and 5 possible average declaration for the group: 0, 25, 50, 75 and 100 experimental points. Overall, there were 25 possible situations of declarations, as reported in Table **??**.

Table 8: Situations in the norms elicitation experiment

100 - 100	75 - 100	50 - 100	0 - 100
100 - 75	75 - 75	50 - 75	0 - 75
100 - 50	75 - 50	50 - 50	0 - 50
100 - 25	75 - 25	50 - 25	0 - 25
100 - 0	75 - 0	50 - 0	0 - 0

Comparison of judgments made per column highlight the impact of the evolution in the gap of declaration on the appropriateness of the situation and the decision to contribute or not to the funding for a given declaration of individual i. We are also able, for a

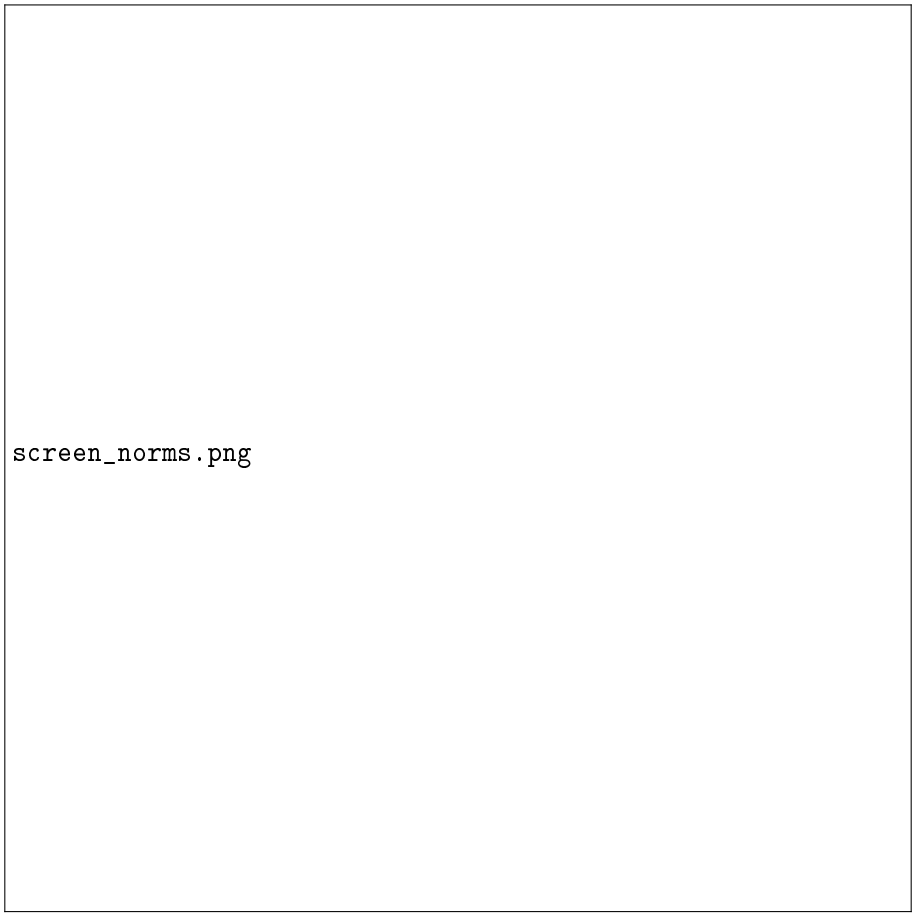
given differential in declarations, to analyze whether the level of declarations themselves impact the degree of appropriateness of the situation and contribution decisions.

For each one of these situations, we present respondents with a description of a situation and the choice environments. To simplify the exposure of the situation, we replace the average declaration of other group members by the declaration of a representative individual. Participants learn the declaration made by an individual i and the one of the representative individual. For each situation, they have to indicate to what extent the situation is socially appropriate or inappropriate and for each possible choices available to individual i (i.e., contribution or non contribution), the extent to which actions are socially appropriate or inappropriate. While we vary the cost of contribution in our experiment, here we use only a fixed cost of 2 points in the norm elicitation experiment.EXPLIQUER POURQUOI. Finally, to help respondents in their decision, the payoff resulting from declaration and contributions decisions are given. In order to not introduce risk concern with the audit probability whose perception made differ across respondents, we provide them the expected payoffs individual i and the representative agent earn:

$$\begin{aligned} u_i &= (1-\rho)(W-\theta X_i) + \rho(W-\theta X_i - (W-X_i)\theta\pi) \\ u_i &= \frac{2}{3}(100-0.3X_i) + \frac{1}{3}(100-0.3X_i-(100-X_i)0.3\times 3.5) \end{aligned} \tag{4}$$

By asking respondents to judge the social appropriateness of each action, we elicit $N(a_k)$, on a four point scale that ranges over "very socially inappropriate", "somewhat socially inappropriate", "somewhat socially appropriate", to "very socially appropriate", like **?**. Fig. **??** below represents the computer screen for one situation

Figure 6: Computer screen for the social norms elicitation experiment



Finally, we provide respondents with incentives not to reveal their own personal preferences but instead to match the responses of others. Thus, respondents play a pure matching coordination game in which their goal is to anticipate the extent to which others will rate an action as socially appropriate or inappropriate, and to respond accordingly. 2 out of the 25 situations were randomly selected; and for each selected situation, one of the rows is again randomly selected. If the choice of the respondent matches with the modal choice of all participants in the room, he earns 3 Euros. In addition, participants received a show-up fee of 6 Euros.