DEPARTMENT OF ECONOMICS,
MANAGEMENT AND STATISTICS

UNIVERSITY OF MILAN – BICOCCA

DEMS WORKING PAPER SERIES

Tax Compliance Under Different Institutional Settings in the EU: An Experimental Analysis

Stefania Ottone, Ferruccio Ponzano, Giulia Andrighetto

No. 307 – July 2015

Dipartimento di Economia, Metodi Quantitativi e Strategie di Impresa
Università degli Studi di Milano - Bicocca
http://dems.unimib.it/

Tax Compliance Under Different Institutional Settings in the EU: An Experimental Analysis

Stefania Ottone[*], Ferruccio Ponzano[•], Giulia Andrighetto[°]

Abstract. In this paper we study how people from different European countries would react, in terms of tax compliance, to institutional changes. We choose an experimental setting and we focus on two features of the tax system – efficiency and tax rate. We develop our analysis in three countries characterized by different systems: Italy, Sweden, UK. The main finding is that participants from different countries react with the same intensity to efficiency changes but not to increases in the tax rate. In all countries tax compliance decreases as tax rate increases, but the reaction is stronger in Italy and softer in UK. Policy implications – mostly focused on fiscal harmonization - follow.

Keywords: tax compliance, fiscal harmonization, cross-country comparison, efficiency, tax rate
JEL classification: C9, D31, H26

[*] Corresponding author. Department of Economics, Management and Statistics, University of Milano-Bicocca, 20126 Milano, Italy. Email: stefania.ottone@unimib.it
[•] Department of Law, Politics, Economics and Social Sciences, University of Eastern Piedmont, 15100 Alessandria, Italy, and Department of Social and Political Sciences, European University Institute, I-50014 San Domenico di Fiesole (FI), Italy. Email: ferruccio.ponzano@uniupo.it
[°] Department of Social and Political Sciences, European University Institute, I-50014 San Domenico di Fiesole (FI), Italy. Email: giulia.andrighetto@gmail.com

1. INTRODUCTION

In this study[1] we implement in the lab a small state in order to perform a cross-country analysis on attitudes towards taxation. In particular, we want to detect how people from *different* countries behave when facing the *same* State and how they react under the *same* institutional setting.

Comparisons among different countries are usually based on wide-ranging surveys such as the International Social Survey Program, the World Values Survey and the European Values Survey (see, among others, Torgler and Schaltegger, 2005; Alm and Torgler, 2006; Frey and Torgler, 2007; Torgler, 2006; Torgler and Schneider, 2007; Cummings et at., 2009; Lago-Penas and Lago-Penas, 2010). Generally, the result is twofold. First of all, each country is assigned a measure of tax morale[2], which is positively correlated to the country level of tax compliance. Secondly, scholars investigate a series of socio-demographic and cultural determinates of tax morale concluding that demographics, culture, and institutions shape tax payer behavior. Lago-Penas and Lago-Penas (2010) provide a detailed survey on the most relevant findings concerning the determinants of tax moral Among demographic characteristics, generally, age, religiosity, employment status and social class play a role.

In particular, tax morale tends to be higher among religious and older people, while it decreases among those individuals who are self-employed or belong to the upper class. The effect of education on tax morale is instead erratic: some studies show that they are positively correlated while others affirm that the relation is the other way round.[3] The effect of gender is also controversial. The majority of cases demonstrate that that the level of compliance among women is higher than in men, while this is not so evident in other analyses[4]. Regarding personal characteristics, risk aversion increases tax morale, suggesting that taxpayers tend to be conditional co-operators as far as their level of compliance depends on others' virtuous behaviour. Direct democracy, a high quality of the services provided by the State, and a fair

[1] This experiment has been financed through an ERC fellowship for the project is "Willing to pay? Testing Institutional Theory with Experiments"

[2] Usually, in order to measure tax morale, people are asked to answer the question: 'Please tell me for each of the following statements whether you think it can always be justified, never be justified, or something in between: …Cheating on tax if you have the chance'. The question leads to a ten-scale index with the extreme points labeled as 'never justified' and 'always justified'.

[3] This result is not surprising. In fact, well-educated people are supposed to be aware more than others of the benefits coming from public goods. At the same time, they are more likely to identify situations where the government wastes resources. Furthermore, they are more likely to understand opportunities for evasion. All these factors may drive to different outcomes.

[4] Experiments on tax compliance provide the same controversial evidence. In some papers women are more compliant (see, among others, Hasseldine, 1999, Lewis et al., 2009, Spicer and Becker, 1980), while in others no significant difference emerges (see, among others, Kirchler and Maciejovsky, 2001, Chung and Trivedi, 2003).

treatment of taxpayers increase tax morale. More generally, a high level of trust in legal and political institutions has a positive effect on tax morale. At the same time, a high level of tax burden decreases tax morale and compliance. Torgler (2004) and Torgler and Schneider (2007) identify two further institutional characteristics that are likely to negatively affect an individual's tax morale: corruption[5] and complexity of the tax system.[6] However, at the moment, we have insufficient evidence to understand whether and to what extent they play a role.

Even though cross-country studies based on surveys are robust and provide relevant results, the experimental methodology seems to be the most suitable tool to answer our research question since it allows us to control the institutional enivronment, something that is nearly impossible in a real-life context. In the lab we can directly observe people's behavior towards taxation and their reaction to controlled and targeted economic changes. We are aware of the limits that are intrinsic to a lab where citizens are mainly a pool of university students. However, as Alm et al. (1992a) argue:

> 'experimental results can contribute significantly to policy debates, as long as some conditions are met: the payoffs, and the experimental setting must capture the essential properties of the naturally occurring setting that is the object of investigation. Laboratory methods may offer the only opportunity to investigate the behavioral responses to policy changes' (p.325). Moreover, 'there is also no reason to believe that cognitive processes of students are different from those of "real" people' (Alm, 1998, p.43).[7]

In line with the mixed-methods approach supported by Poteete et al. (2010), according to which the use of different complementary methodologies to a topic provides a better knowledge of the phenomenon, our results represent a useful contribution to better understand how people react to institutional changes.

In this study, we manipulate in the lab institutional characteristics collect data on participant's socio-demographic information and their attitude towards risk and others[8] through a socio-

[5] Corruption implies waste of public resources, inefficiency and unfair treatment of citizens. Consequently, it may reduce people's trust in public institutions. At the moment we have some evidence from Transition Countries (Torgler, 2004) and Africa (OECD Tax and Development report, 2013) where it emerges that corruption in public institutions decreases the level of tax morale.

[6] Torgler and Schneider (2007) report that simplicity is a major issue in tax reforms. This is because 'Complexity may result in unintentional non-compliance if taxpayers have problems filling out the tax form. It can reduce the moral costs of evading taxes and might impose costs to the taxpayers' (p.28).

[7] For a detailed discussion on the relevance of the experimental evidence in studying tax compliance see Torgler, 2002.

[8] Generally, in studies based on a survey, the role played by subjects' prosocial attitude is not detected. However, Trivedi et al. (2003) in their experimental analysis proved that prosociality positively affects people's tax compliance. Since fiscal behavior includes a social dimension, we decide to study how people's attitude towards others influences tax compliance in the lab.

demographic questionnaire at the end of the experiment and the participation to the Social Value Orientation (SVO) Survey.[9] Our work focuses on two relevant features of the tax system – efficiency and tax rates. In particular, we examine how people from different countries react to varying tax rates and levels of efficiency. The topic is relevant since a change in the tax system could generate a variation in the level of tax evasion and, consequently, a variation on the tax revenue and the services a state can provide (Clotfelter, 1983). Alm et al., 1992b and Bosco and Mittone, 1997 provide some experimental evidence on these issues. Experimental methods is a common methodological approach in the tax evasion literature.[10] However, to our knowledge, this is the first analysis on a large experimental sample from different countries within Europe. We focus our analysis on three countries: Italy, Sweden and UK. The choice of these three countries is due to the fact that they present few similarities – the presence of a democratic and unitary state being the most relevant - and differences representative of three typical state types in the EU.[11] In particular, these three countries show differences concerning the two institutional characteristics we are focused on. Italy and Sweden show a high tax burden while UK shows a low one. Whereas, Sweden and UK can be considered efficient states, Italy is not.

We measure the tax burden using the data provided by the OECD regarding the total tax revenue as percentage of GDP (see Table 1). We can easily observe that Sweden and Italy present a higher tax burden than UK.

The measure of efficiency is a bit harder to measure. We base our estimation on two different indicators. First of all, we use a measure of corruption – the corruption perception index. Corruption is problematic because it shifts resources from general welfare to self – usually criminal – interests. In a certain sense, it represents a waste of resources and, at the same time, a source of unfair behavior of the State towards its citizens. In a 0 – 100 scale – with the 0 level signifiying maximum corruption and 100 representing the no corruption – perceived corruption measured in 2013 is 89 in Sweden (rank: 3rd), 76 in England (rank: 14th) and 43 in Italy (rank: 69th)[12].

[9] Murphy et al. (2011).

[10] See Alm, 2012, for a survey and a comparison with other methods

[11] Applying the classification proposed by Esping Andersen (1990), Italy belongs to the Corporatist-Statist model, UK to the liberal one while Sweden is representative of the Social Democratic regimes[11]. To sum up the three models, the first one is characterized by services provided by the state for the citizens mainly based on their employment history; the second one is identified by a low level of services provided by the state, especially for poor people, while in the third one the provision of public services is universal. Also in the classification proposed by Titmuss (1974), these three countries belong to different categories of Welfare State.

[12] The index includes 177 countries and territories. Source: Transparency International, 2013. This index is updated every year but the variations are little.

The second proxy is based on the quality of government measure provided by the World Bank Governance Indicators. The 27 members of the EU are ranked on the basis of four indicators[13]. Starting from this classification, the countries can be grouped in three categories ordered on the basis of the quality of government. UK and Sweden belong to the first group (high quality), while Italy belongs to the third one (low quality).

The main result is that participants from different countries react similarly to efficiency changes, while changes in tax rates cause different reactions from people from different countries. Unsurprisingly, tax compliance decreases as tax rate increases in all countries, but Italians react more strongly to tax rate increased than to British participants.

Shedding light on how people from different countries behave under similar tax conditions can provide some hints for policy implications. For instance, the results obtained in this paper can provide a contribution in the debate about fiscal harmonization (see, among others, Tiebout, 1956, Sinn, 1990 and Kirchgassner and Pommerehne, 1996) in European Union. In fact, even if we do not study this topic directly, we find how people coming from different countries facing very different tax systems show similar preferences, in particular regarding the level of efficiency. A possible implication is that, if European Union is able to ensure that fiscal systems of the country members maintain a high level of efficiency, citizens in different countries may be ready to accept changes in the tax rates.

The next section is devoted to a description of the experimental design and procedure. The third section provides the results while the last section provides some policy implications and concludes.

2. EXPERIMENTAL DESIGN AND PROCEDURE

In this laboratory experiment decisions are recorded through the computer. Instructions are read by participants on their computer screen, while a researcher reads them out loudly.[14] The experiment is programmed and conducted with Z-tree (Fishbacher, 2007).[15]

[13] The four indicators are: government effectiveness, control of corruption, rule of law and voice and accountability. For a better description of this measures, see the report "Measuring the Quality of Government and Subnational Variation". In particular, page 22 and 23 are devoted to a brief explanation of the groups and the characteristics and provide a table with the rankings. This table is provided by Kaufmann, Kraay and Mastruzzi (2009)

[14] In each location, instructions are read by a mother-tongue reader.

[15] The experiment has been programmed by Marie-Edith Bissey.

Participants enter the laboratory and take a seat in front of a computer. They are immediately asked to switch off their mobiles and to stop talking to their colleagues. First of all, instructions are presented. Subjects are informed that they will participate in several activities (which the experimenters will gradually describe to them) where they will have to perform some tasks and to make some choices. Participants are informed that, based on their choices, on the choices of the other subjects and on chance they will earn experimental currency units (ECUs), which will be converted into real money at the end of the session.

The first part of the experiment consists of two phases of three rounds each where participants are asked to declare their earnings from a clerical tax for tax purposes according to different fiscal systems. During each round, participants are free to report any amount from 0% to 100% of their gross income and they are only taxed on the income they report. There is, however, a 5% possibility of being audited. At the end of the experiment[16], if a subject is audited and it is discovered that she has under-reported her earnings, she will have to pay a penalty equal to twice the tax she should have paid on the income she did not report. Participants are instructed round by round on the specific characteristics of the fiscal systems and they are not informed till the end of the experimental session of the result of the auditing procedure.

This is a real-effort experiment, in each phase subjects have to earn their experimental gross income through an individual clerical task.[17] In this task, subjects must copy rows of information containing fictitious students from a sheet of paper onto the computer for five minutes (see Figure 1). For each correctly copied row, subjects earn 10 ECUs.[18] Consequently, in each phase, each subjects' experimental gross income is the amount of ECUs earned during the five minutes of clerical task.[19]

Each phase focuses on a particular feature of the tax system – efficiency and tax rate. In Phase 1 we hold tax rate constant at 30% and we vary government efficiency manipulating the provision of a public good. In the No Pot scenario (round 1 - NOPOT), the tax revenue is not redistributed at all. In the Pot situation (round 2 - POT), the tax revenue is equally divided among all the participants, irrespectively of their individual contribution. In the Double pot setting (round 3 - DOUBLEPOT), the tax revenue is doubled and equally divided among all

[16] We reported the outcome of the audits at the end of the experiment since we wanted to be sure that different levels of tax compliance across the rounds were due to the experimental variables and not to the result of the audit procedure.

[17] Before starting the first phase of the experiment, subjects are asked to participate in a practice session in order to be familiar with their activity.

[18] Each ECU is equal to 0.01 euro in Italy, 10 pence in UK and 6 kronor in Sweden.

[19] Subjects' gross incomes vary according to their ability to correctly type rows of text into predefined boxes on their computer screens. However, participants do not know how their earnings compare to others in their session.

the participants. In the second phase we keep redistribution constant as in the Double Pot round, but we change the tax rate switching from 10% (round 4 – TAX10) to 30% again (round 5 – TAX30) and to 50% (round 6 – TAX50). Overall, in the first two phases of the experiment, participants perform the clerical task twice and are asked to self-report their gross income 6 times.

In the third phase of the experiment subjects are coupled and asked to participate in the *Social Value Orientation (SVO) task*. It consists of a series of mini-dictator games whose aim is measuring people's attitude towards the others (see Appendix for details). Choices in the *SVO task* are incentivized in the following way: at the end of this activity, only one person in each pair will have her decisions applied and these decisions will determine the ultimate payoff that both subjects receive at the end of this Survey. The earnings from this task are added to the earnings from the previous two phases of the experiment to determine the final income.

As soon as everyone completes the *SVO* activity, each participant reads on her own screen the final payment and the result of the audit procedure for each round. Finally, before privately receiving their payment, experimental subjects fill-in questionnaire, whose aim is to record people's socio-demographic characteristics, attitudes towards the State and the risk as well as beliefs concerning others' performance and choices. A summary of the experimental phases is reported in Figure 2.

The experiment is anonymous in all stages: contribution from other people and *SVO* partners' identities are unknown even when the experiment is over. Subjects are provided an ID number at the beginning of the experiment and experimenters can link participants's decisions and payments to this ID only.

We run the same experiment in eight locations – three in Italy, three in UK and two in Sweden (see Table 2 for details).

3. DATA ANALYSIS

3.1 The sample

Overall, 998 subjects participated in the experiment - 311 in Italy, 360 in UK and 327 in Sweden. They were either former or current university students recruited through a web-based recruitment system. Table 3 reports some descriptive statistics of the socio-demographic characteristics of our sample.

Before going on with our analysis, we have to check whether we can pool data coming from different locations in the same Country. In other words, we have to check whether people's

choices are homogeneous across all the location in the same Country. Figure 3 suggests that no significant difference exists. A series of t-test on people's compliance comparing experimental subjects' choices in couples of different location of the same country, round by round, confirm the evidence reported in Figure 3 ($p > 0.111$). Consequently, in our analysis we will consider data coming from different locations of the same country as pool data.

3.2 Results

If we analyse whether and to what extent people's attitude toward taxation changes as far as the institution incentives vary, it turns out that:

Experimental subjects react to institution incentives, no matter the country. More specifically, tax compliance increases as efficiency increases and decreases as the tax rate increases. However, although people's reaction to changes in efficiency is homogeneous across countries, subjects from different countries react with a different degree to an increase in the tax rate. In particular, participants who live in Italy or Sweden – countries where the tax burden is usually high – react more strongly to an increase in the tax rate than our British subjects. At the same time, subjects in Sweden – where the efficiency of the public service is high – react less to tax rate increases than Italian subjects.

Figures 4 and 5 display the average level of tax compliance in each of the six rounds, broken down between experimental phases and countries. From these graphs, two relevant points emerge immediately. First of all people reaction to institutional incentives is relatively uniform. Secondly, people's reaction to changes in the tax rate varies across countries.

In order to better understand experimental subjects' attitude towards taxation when the institutional scenario changes, we perform a series of regressions. First of all, we run regressions on each country. The aim is to check whether a change in the institutional scenario affects people's behaviour in all the three countries. The dependent variable is the level of tax compliance and is ranged between 0 and 1, where 0 implies total evasion and 1 total compliance. Consequently, we opt for a random-effect tobit model. Our base specification is:

$$TAX\_COMPLIANCE_{it}=\alpha+\Sigma_l\beta_l INSTITUTION\_INCENIVES_{it}+\Sigma_r\eta_r DEMOG_i+\Sigma_s\gamma_s PERSONAL\_CHARACTERISTICS_i+\mu_i+\varepsilon_{it} \quad \text{(R1)}$$

Regressors are a series of determinants of tax morale we mentioned in the Introduction. Demographic variables (DEMOG), include age and dummy variables for participants who are male, employed and study (or studied) economics.[20] Two dummies for high earners and low earners are included.[21] A further dummy variable is included to identify people who participated in the past in a lab experiment. PERSONAL_CHARACTERISTICS report subjects' attitude towards risk and the relevance played by the others. More specifically, we include a dummy variable for those participants who are labeled as prosocial through the SVO Survey[22] and a dummy variable for people who think that most of other participants report the whole gross income. We include a control for risk attitudes, as measured by a survey item which asks subjects to rank themselves on a 10-point scale, with 1 signifying a person who "normally tries to avoid taking risks" and 10 signifying someone who is "completely willing to take risks".

Since each regression is run on rounds belonging to the same phase in order to focus our analysis on an institutional factor at time, INSTITUTION_INCENTIVES depend on the phase. In Phase 1, they are represented by two dummy variables – Nopot and DoublePot – equal to 1 if the observation comes from the NOPOT or the DOUBLEPOT round respectively. The baseline is the Pot scenario. In Phase 2, the baseline is the TAX30 round and dummy variables are included to study the two remaining settings (TAX10 and TAX50).[23]

From Table 4 and 5, it emerges that as far as the level of efficiency increase, the level of compliance increases as well, no matter the country. The same generalized effect occurs if we

[20] Since we run an experiment with highly educated subjects (University students and graduates), we prefer to avoid any generic control for the level of education and we choose to include a dummy on a specific field of studies - where people are supposed to be taught to behave selfishly.

[21] 'High earner' is a dummy equal to 1 if subject *i* declares in the final questionnaire that her performance in the clerical task is above the average level. 'Low earner' is a dummy equal to 1 if subject *i* declares in the final questionnaire that her performance in the clerical task is below the average level. We decided to use this subjective measure and not the income earned in each phase since the latter does not provide any information concerning subjects' perceived position within the society. We think that this is a more realistic proxy of what in real life is the perception of belonging to a particular social class.

[22] The *SVO Survey* allows computing a continuous measure of subjects' attitude towards others – the *SVO* angle (see Appendix A for more details). Murphy et al. (2013) identify four different types from the value of the *SVO* angle – competitive, individualistic, prosocial and altruistic people. In our analysis we simplify this categorization and we decide to distinguish between two types: self-interested people (the representation of the classical Homo Oeconomicus) and prosocial subjects (those who care, even if minimally, about the others and that people in their everyday conversations call 'altruistic'). From a technical point of view, if the value of the *SVO* angle is lower than 7.82, no positive attitude towards others is registered and people are classified as self-interested. If the value is higher than 7.82, a subject is identified as prosocial.

[23] As we mentioned in the Introduction, efficiency and tax rate are the institutional characteristics of our analysis. We intentionally exclude 'subjective' measures like trust in institutions since they cannot be considered independent from the 'objective' peculiarities of a State. Consequently, we choose to check for the effect of these objective economic factors directly. Finally, we do not control for variables like the complexity of the system and auditing. The former is excluded because participants receive detailed instructions and examples in each setting. Thus, we assume that the fiscal system in the lab is clear enough. Then, the same (low) probability of being audited in each round cannot be responsible for subjects' different behavior across the phases.

consider subjects' reaction to an increase in the tax rate. In all 3 countries, higher tax rates imply lower compliance. This is in line with experimental evidence: as Alm (2012, p. 66) affirms: "most (but not all) experimental studies have found that a higher tax rate leads to less compliance" and "The presence of a public good financed by voluntary tax payments has been found to increase subject tax compliance".[24] However, notice that the fact that people from different countries react to the *same* institutional incentives does not mean that the level of compliance in the lab is the *same* in all countries. From Figure 4 and 5, it turns out that experimental subjects in UK comply less than people from the other two countries. This evidence is confirmed through a Tobit regression[25] on the average level of compliance over the 6 rounds (see Table 6 for results), whose specification is:

$$AVERAGE\_COMPLIANCE_i=\alpha+\Sigma_k\omega_k COUNTRY_i+\Sigma_r\eta_r DEMOG_i+\Sigma_s\gamma_s PERSONAL\_CHARACTERISTICS_i+\varepsilon_i \quad \text{(R2)}$$

The step further we want to go is to detect whether the *intensity* of people's reaction to changes in the economic setting is different in different countries. Again, we will go through phase by phase. In the first phase we will perform a series of t-test on the differences between compliance in the different rounds and on people's global reaction along the phase. More specifically, in all countries we will test whether the following differences are different from zero: (POT - NOPOT), (DOUBLEPOT – POT) and (DOUBLEPOT – NOPOT). Moreover, we run a regression where the dependent variable is the difference (DOUBLEPOT – NOPOT) that we will call DIFFPOT. In the second phase we repeat the same procedure as in the first

[24] The fact that people's reaction to efficiency and tax rates is in line with both the literature on tax morale and the previous experimental evidence may represent a proof of the representativeness of our sample as well as of the robustness of our methodology. It is worth underling other findings from our experiment are in line with the existing literature on tax morale and tax compliance. For instance, the correlation between tax compliance and some variables like attitude towards risk, beliefs about others' behavior and economic factors shows the same direction as in literature. At the same time, we think we provide a novel contribution on different issues. In particular: we provide some original results on the role played by gender and prosociality. As mentioned in the introduction, literature on tax morale is erratic with respect to the gender effect. In our analysis, we find that women are more compliant than men in all phases. Moreover, this result is strongly significant in all countries. An analysis of a possible correlation between people's value orientation and tax compliance is another novelty of our study. To our knowledge, literature on the role played by attitudes towards others on tax compliance is nearly missing. The only scholar who checked whether attitudes towards others matter is Trivedi (2003), who found the same result as we did. However, his study was run on a sample of 98 Canadian university students only. As we mentioned in the Introduction, generally, studies on tax morale focus more on taxpayers' beliefs concerning others' contribution as key factors. In our experiment, in all countries, prosocial subjects comply significantly more than individualistic people, no matter the institutional setting.

[25] This result is in line with the existing literature on tax morale where UK score is usually lower than the Italian and the Swedish ones (see for instance Alm and Torgler 2006).

phase. Now, the tested differences are: (TAX30 – TAX10), (TAX50 – TAX30), (TAX50 – TAX10) and, again, we run a regression on the global difference (TAX50 – TAX10) that we call DIFFTAX.

We use a classical OLS model and the specifications are:

$$DIFFPOT_i=\alpha+\Sigma_k\omega_k COUNTRY_i+\Sigma_r\eta_r DEMOG_i+\Sigma_s\gamma_s PERSONAL\_CHARACTERISTICS_i+\varphi ATTC\_INDEX_i+\varepsilon_i \quad (R3)$$

and

$$DIFFTAX_i=\alpha+\Sigma_k\omega_k COUNTRY_i+\Sigma_r\eta_r DEMOG_i+\Sigma_s\gamma_s PERSONAL\_CHARACTERISTICS_i+\varphi ATTC\_INDEX_i+\varepsilon_i \quad (R4)$$

Where:

$\Sigma_k COUNTRY_i$ are dummy variables for countries

$ATTC\_INDEX_i$ (Attitude Towards Tax Cheating) is an index that is aimed at measuring to subject *i*'s sense of duty towards the State and the other citizens (focused on fiscal obligations).

The ATTC_INDEX is computed for each subject on the basis of her responses to three following items of the final questionnaire:

Please indicate how much you agree with each of these statements, where 1 means you completely disagree and 4 means you completely agree:

1) *Paying taxes is a fundamental duty of citizenship*
2) *Not paying taxes is one of the worst crimes a person can commit because it damages the entire community*

and

Please tell me whether you think the following actions can always be justified, never be justified, or something in between. "1" means you think the action can never be justified, and "10" means you think the action can always be justified:

Cheating on taxes if you have a chance

Through this index we want to investigate whether, and if, how subjects' sense of duty affects people's reaction to institutional changes. In particular, we want to understand whether citizens with a higher sense of duty towards the State and the other citizens are more critical on fiscal policies or they are instead more willing to pay taxes without putting their government decisions into discussion. We think this is a relevant since, for instance in Italy, the government has been supporting for years an advertising campaign aimed at increasing

citizens' sense of duty and focused on tax evasion. However, to our knowledge, there is no empirical study on the relation between people's reaction to fiscal reforms and their sense of duty towards the State and the other citizens.

In our analysis subjects' reaction to efficiency is homogeneous across countries (see Table 7 and Table 9). On the other hand, Italy and Sweden show a stronger reaction to an increase in the tax rate, even if Swedish react less than Italians (see Table 8 and Table 9). Our feeling is that our results depend on different issues. First of all, while inefficiency can be defined as an objectively negative feature of a government, a high tax rate is not negative *per se.* In other words, large tax revenues obtained through a high level of taxation may represent a positive element if the State is able to use it in an efficient way. As we mentioned in the introduction, our sample comes from three countries that have quite different welfare states. Italy and Sweden both have large public sectors. Which implies that the State is likely to affect its citizens' lives more than in UK. Even perception of the importance of the public service is different between these countries: in our sample, Italians and Swedes respond that 47.5% and the 51.2% respectively completely agree on the fact that paying taxes is rational because it finances useful and important services. This percentage drops to 28.7% in UK. Moreover, when we ask subjects to place themselves on a 0-10 point scale where 10 represents complete agreement with the statement "Individuals should take more responsibility for providing for themselves" and 0 indicates complete agreement with the statement "The state should take more responsibility to ensure that everyone is provided for", Italians and Swedes score 3.7 and 4.3 respectively, while UK reports a value of 5.6. This implies that Italian and Swedish subjects are more likely to favor a larger role for the government.

The stronger negative reaction of Italian subjects to an increase in the tax rate may be due to the fact that in everyday life they suffer from high tax rates combined with inefficiency and corruption. This probably makes high tax rates as an undesirable characteristic of the institutional setting – in presence of corruption and inefficiency, the higher the tax rate, the bigger the amount of public money that is wasted by the state. In fact, in the final questionnaire, 67.5% of Italian participants state that people would be more likely to pay taxes if the government were more efficient (vs 34.4% and 30.3% in UK and Sweden respectively) and 54.6% would comply with their fiscal obligations if they had some control over how tax money were spent (vs 30.8% and 25.8% in UK and Sweden respectively); about 14% declares that most of people in Italy evade their fiscal obligations because politicians are

corrupt (the percentage is 4.5% in UK and 2.8% in Sweden) and 25.8% of Italian participants completely agree on the fact that most of people are forced to evade taxes because of high tax rates (this percentage is only 15% and 11.8% in UK and Sweden respectively).

What about the relevance of the ATTC_INDEX? From (R3) and (R4) it turns out that it affects people's intensity of reaction to changes in tax rate, but not in efficiency. More specifically, the higher the index, the lower the intensity of the reaction to an increase in the tax rate. Again, we found homogeneous responses in the efficiency phase, while the tax rate phase makes people's differences emerge. A plausible explanation can be found, again, in the different nature of the institutional changes. As we mentioned before, an increase in the tax rate is not negative *per se*. People who show a higher sense of duty are less responsive to tax rate changes. In any case, notice that this index is higher in Italy than in the other two countries (0.33 vs -0.06 and -0.22 in Sweden and UK respectively). However, Italy is more sensitive to changes in the tax rate. This may imply that the quality of the government and its efficiency matters more. No way to impose a high tax burden on citizens if the tax revenue is wasted through inefficiency and corruption.

More policy implications of our findings are provided in the next and final section of the paper.

3. DISCUSSION AND CONCLUSIONS

Our experimental work is an attempt to provide a contribution in order to understand how people from different countries, facing very different tax systems, react to the *same* changes in efficiency and tax rate. Specifically, we perform an analysis on tax compliance where people from different countries face the same institutional setting and the same fiscal system in the lab. Our evidence suggests that tax compliance increases when efficiency increases, and decreases when the tax rate increases, no matter the country. However, although people's reaction to changes in efficiency is homogeneous across countries, subjects from Italy and Sweden – countries where the tax burden is usually high – react more strongly differences in tax rates than do British subjects. At the same time, subjects in Sweden – where the efficiency of the public service is high – react less than Italian people as the tax rate changes.

Starting from our results, some remarks about policy implications can be made, especially in the context of fiscal harmonization. The take-home message is that efficiency is a key element. In fact, there is no doubt that people from different countries are sensible to efficiency issues. The reaction to inefficiency is really high in each country. In line with this

argument and on the basis of our experimental evidence, we may suggest that increasing efficiency in countries where the level is low might reduce tax evasion. Consequently, the first step to do if we want to harmonize fiscal systems is to enforce governments to implement a similar and very high level of efficiency in each country.[26] In our case, Britons and Swedish governments need to maintain a high level of efficiency while in Italy politicians have to work with it. As efficiency is ensured, shifts in the tax rates seem to be accepted by citizens. In fact, as Swedish and Britons, which are accustomed to efficiency, show a low level of variation even if the jumps in the tax rate are really high, Italians show a strong reaction. Then, we could expect that Italian people behave in the same way if they experience a higher level of efficiency for a long period. Thus, we think that a European Authority dedicated to ensure the efficiency target in the UE members would probably be more useful than restrictive debt and deficit parameters.[27]

Finally, we think that the role played by efficiency in the EU context deserves further inquiry. In particular, in our experiment we do not investigate how differences in efficiency may affect the interaction *between* countries. We think this is a relevant topic. In fact, if efficiency plays a relevant role in the single citizen's decisional process, we may assume that it is a relevant feature for governments too in an international context. For instance, a possible complication against fiscal harmonization may be represented by the resistance of efficient countries who do not want their resources to be transferred to governments who waste money. As far as increasing efficiency in countries where the level is low reduces tax evasion, this is likely to increase trust among governments from different countries.

[26] Notice that we are not affirming that the obvious consequence is the centralization of the fiscal power. At the moment, we do not take a stand on it and we do not provide any evidence on this issue, even if we think it is a topic that deserves further inquiry, maybe through behavioral studies in order to understand how citizens could react.

[27] In the October 2014 *World Economic Outlook* from the International Monetary Fund, focused on the effect of public investment in infrastructure, it emerges that "public investment shocks are found to lead to a significant medium-term reduction in the debt-to-GDP ratio [..] in countries with high public investment efficiency" (p.83).

References

Alm, J. (1998), Tax Compliance and administration, Working Paper, University of Colorado at Boulder.

Alm, J., Jackson, B., McKee, M. (1992a), Deterrence and Beyond: Toward a Kinder, Gentler IRS, in Why People Pay Taxes: Tax Compliance and Enforcement, Joel Slemrod, ed., Ann Arbor, MI: The University of Michigan Press, 311-329.

Alm, J., McClelland, G.H. and Schulze, W.D. (1992b). Why do people pay taxes? Journal of Public Economics, Issue 48, 21 –38.

Alm, J., Torgler B. (2006), Culture differences and tax morale in the United States and in Europe, Journal of Economic Psychology 27 (2), 224-246

Bosco, L. Mittone, L. (1997), Tax evasion and moral constraints: some experimental evidence, Kyklos, 50 (3), 297-324

Chung, J., Trivedi, V. U. (2003) The Effect of Friendly Persuasion and Gender on Tax Compliance Behavior, Journal of Business Ethics 47 (2):133 – 145.

Clotfelter, C. T (1983). Tax Evasion and Tax Rates: An Analysis of Individual Returns, The Review of Economics and Statistics, 65 (3), 363-373.

Cummings, R. G., Martinez-Vazquez, J., McKee, M., Torgler, B. (2009), Tax morale affects tax compliance: Evidence from surveys and an artefactual field experiment, Journal of Economic Behavior & Organization, 70, 447–457

Esping Andersen, G. (1990), The three worlds of welfare capitalism, Cambridge, Polity Press. Experimental Economics, 10, 171-178.

Fischbacher, U., (2007). z-Tree: Zurich toolbox for ready-made economic experiments.

Frey, B. S., Torgler, B. (2007), Tax morale and conditional cooperation, Journal of Comparative Economics 35 (1), 136-159, 2007

Hasseldine, J. (1999). Gender differences in tax compliance. Asia-Pacific Journal of Taxation, 3(2), 73–89.

Kaufmann, D., Kraay, A., Mastruzzi, M. (2009), Governance Matters VIII: aggregate and individual governance indicators 1996-2008, Policy Research Working Paper Series 4978, The World Bank

Kirchgassner, G., Pommerehne, W. W. (1996). Tax harmonization and tax competition in the European Union: Lessons from Switzerland, Journal of Public Economics, 60(3), 351-371.

Kirchler, E., Maciejovsky, B. (2001). Tax compliance within the context of gain and loss situations, expected and current asset position, and profession, Journal of Economic Psychology, 22(2), 173-194.

Lago-Peñas, I., Lago-Peñas, S. (2010). The determinants of tax morale in comparative perspective: Evidence from European countries, European Journal of Political Economy, 26(4), 441-453.

Lewis, A., Carrera, S., Cullis, J., Jones, P. (2009), Individual, cognitive and cultural differences in tax compliance: UK and Italy compared, Journal of Economic Psychology 30, 431–445

Murphy, R. O., Ackermann, K. A., Handgraaf, M. J. (2011), Measuring Social Value Orientation, Judgement and Decision Making, 6 (8), 771-781.

Poteete, A. R., Janssen, M. A., Ostrom, E. (2010), Working Together: Collective Action, the Commons, and Multiple Methods in Practice, Princeton, NJ, Princeton University Press.

Sinn, H.-W. (1990) Tax harmonization and tax competition in Europe, European Economic Review, 34(2-3), 489-504.

Spicer, M., Becker, L. A. (1980). Fiscal inequity and tax evasion: an experimental approach. National Tax Journal 33 (2), 171–175.

Tiebout, C. M. (1956) A Pure Theory of Local Expenditures, The Journal of Political Economy, 64 (5), 416-424

Titmuss R., (1974), Social Policy: An Introduction, Allen & Unwin, London.

Torgler, B. (2002), Speaking to theorists and searching for facts: Tax morale and tax compliance in experiments, Journal of Economic Surveys 16 (5), 657-683

Torgler, B. (2004), Tax morale, trust and corruption: Empirical evidence from transition countries Center for Research in Economics, Management and the Arts (CREMA), Working Papers

Torgler, B. (2006), The importance of faith: Tax morale and religiosity, Journal of Economic Behavior & Organization 61 (1), 81-109

Torgler, B., Schaltegger, C. A. (2005), Tax amnesties and political participation. Public Finance Review 33 (3), 403-431

Torgler, B., Schneider, F. (2007), What Shapes Attitudes Toward Paying Taxes? Evidence from Multicultural European Countries, Social Science Quarterly 88 (2), 443-470

Trivedi, U.V., Shehata, M., Lynn, B. (2003). Impact of Personal and Situational Factors on Taxpayer Compliance: An Experimental analysis. Journal of Business Ethics, 47, 3.

APPENDIX A. THE SOCIAL VALUE ORIENTATION (SVO) SURVEY

Psychologists are used to measure people's attitude towards others through an instrument called Social Value Orientation (SVO). SVO categorization is based on people's choices in a series of Dictator Games (called also Decomposed Games - DG), where the strategic dimension is excluded.

In this work we use the Slider measure since it results to be the more robust measure (see Murphy et al., 2013). It consists of 6 primary and 9 secondary items. In each item, a subject has to choose an outcome allocation between herself and another player out of nine options. From a subject's choices in the primary items (see Figure 1), the *SVO angle* can be computed in the following way:

$$SVO° = arctang\left(\frac{Ao - 50}{As - 50}\right)$$

Where Ao is the mean allocation for the other and As is the mean allocation for self.

The Slider measure provides a continuous measure of people's SVO, starting from a perfectly competitive position at -16.26° and ending with a perfectly altruistic behavior at 61.39°. An Individualist who does not want to damage others chooses (85;85), (100;50), (85;85), (85;15), (100;50), (100;50) in the six Primary items and obtain a score of 7.82. This is why in our work we identify two categories. If the value is lower than 7.82, no positive attitude towards others is registered. If the value is higher than 7.82, a subject is identified as prosocial.

Figure 1A. Primary Items

Item 1

You receive	85	85	85	85	85	85	85	85	85
The other receives	85	76	68	59	50	41	33	24	15

Item 2

You receive	85	87	89	91	93	94	96	98	100
The other receives	15	19	24	28	33	37	41	46	50

Item 3

You receive	50	54	59	63	68	72	76	81	85
The other receives	100	98	96	94	93	91	89	87	85

Item 4

You receive	50	54	59	63	68	72	76	81	85
The other receives	100	89	79	68	58	47	36	26	15

Item 5

You receive	100	94	88	81	75	69	63	56	50
The other receives	50	56	63	69	75	81	88	94	100

Item 6

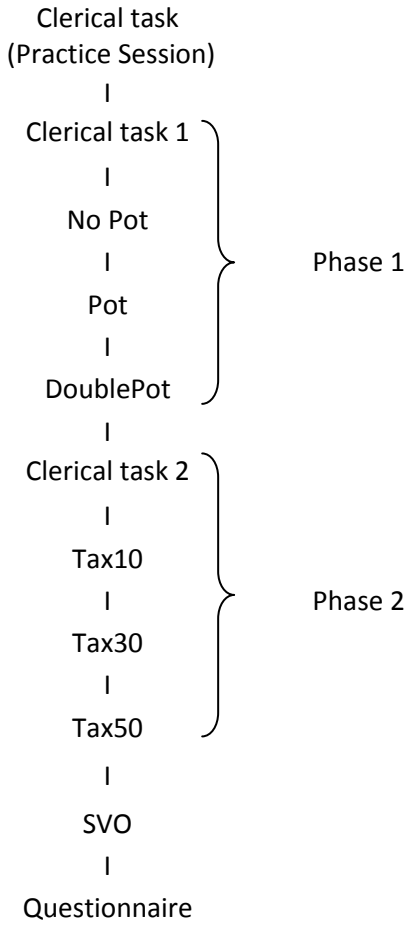
You receive	100	98	96	94	93	91	89	87	85
The other receives	50	54	59	63	68	72	76	81	85

FIGURES AND TABLES

Figure 1. Example of Information to be transcribed in the clerical task

Row	*ID number*	*Last Name*	*First Name*	*Vote*
1	957302	Iklmqy	Dglpqtwy	0
2	023568	Zsefgwt	Ipdbzycx	6
3	469217	Ginvxy	Zuexfkja	8
4	528196	Tmgczu	Zjpvwks	9
5	816593	Bcdhknoy	Klprswx	9
6	125678	Bhmoqsuy	Bcglnvwx	1
7	012468	Amnsux	Cfiloy	0
8	912065	Ysjzmucl	Cgiltvw	7
9	432970	Bdgnsx	Wbzaqv	7
10	234567	Vzbnjp	Zmfbgo	6
11	391586	Fijkln	Zwmvojn	5

Figure 2. Experimental Design and Procedure



Instructions are read round by round, as far as participants are asked either to perform an activity or to make choices.

Figure 3. Average level of compliance in each phase across locations within the same country

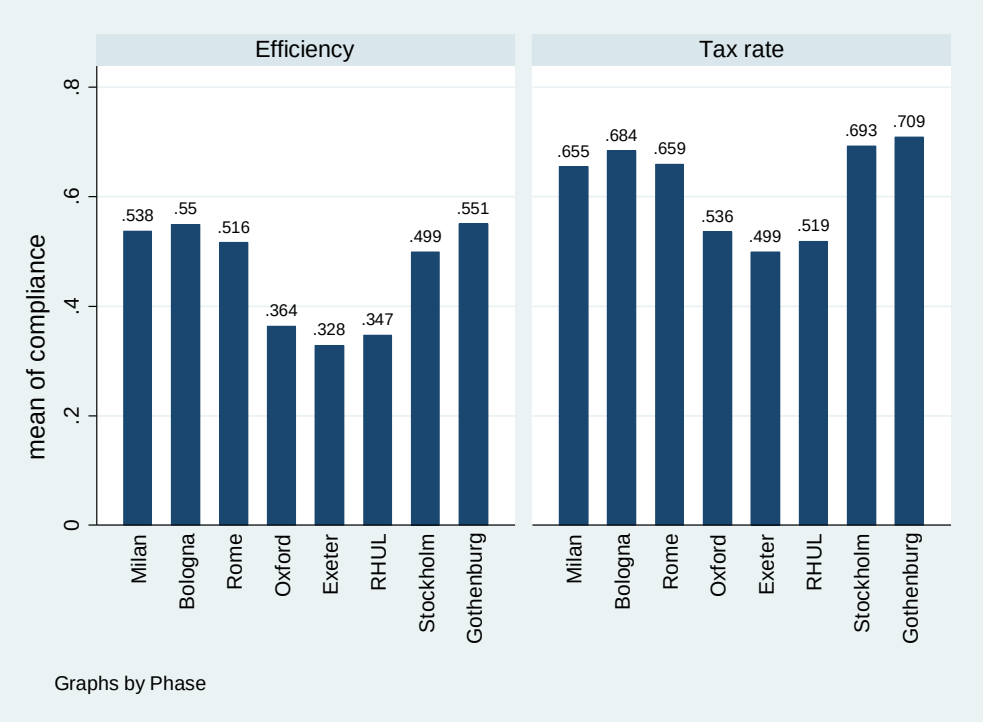


Figure 4. Average level of compliance in each round by country - Phase 1

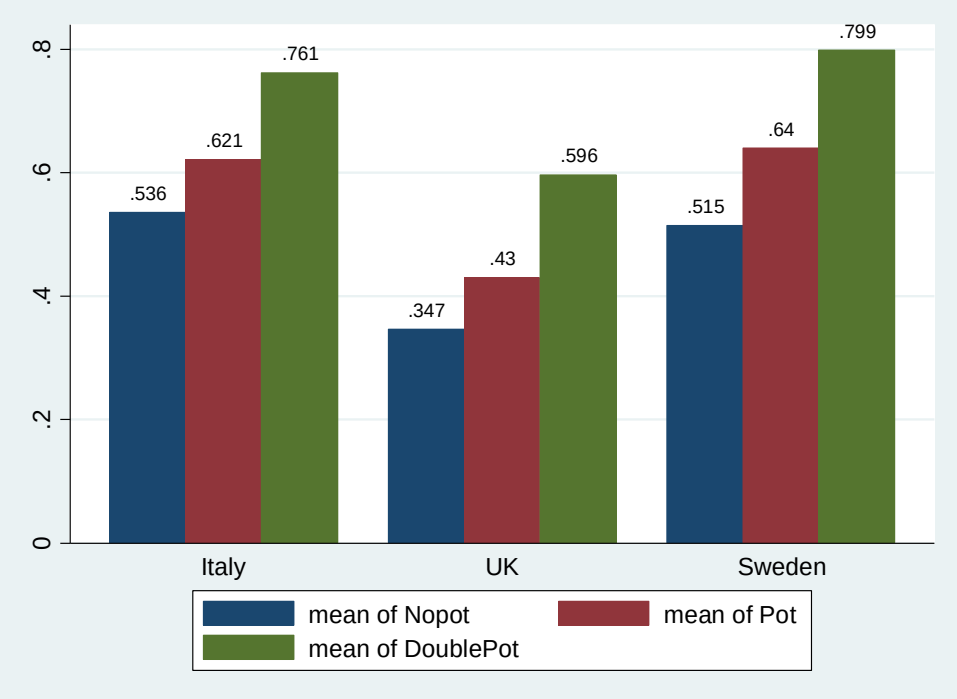


Figure 5. Average level of compliance in each round by country - Phase 2

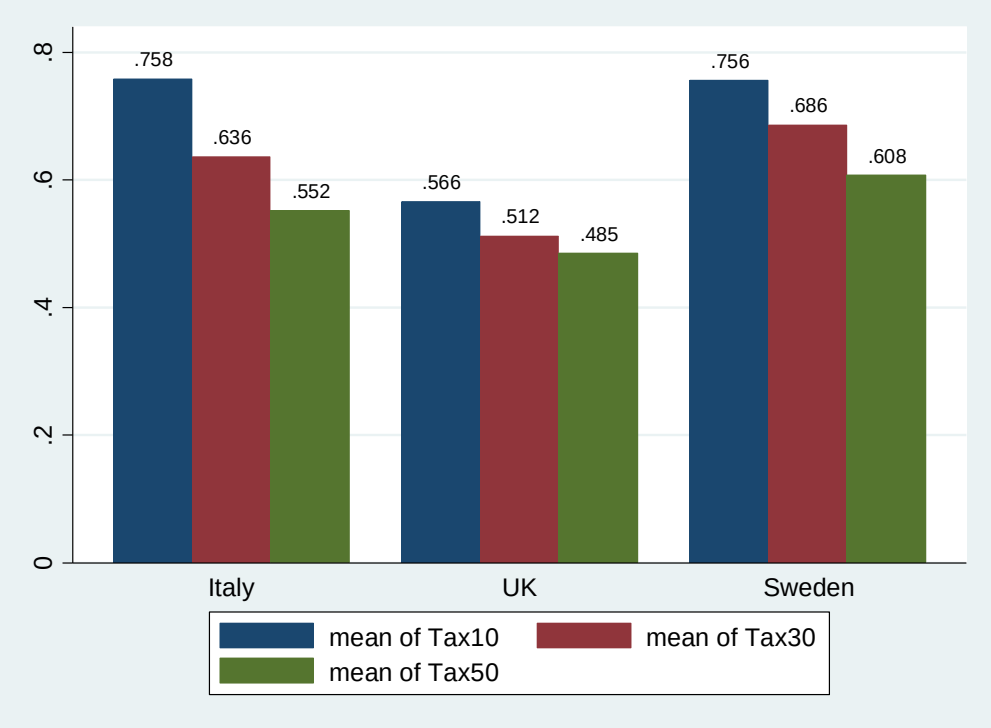


Table 1. Total tax revenue as percentage of GDP

	2007	2009	2010	2011	2012 (provisional)
Italy	43,2	43,4	43,0	43,0	44,4
Sweden	47,4	46,6	45,4	44,2	44,3
United Kingdom	35,7	34,2	34,9	35,7	35,2

Table 2. Number of experimental subjects by location

Location	Number of subjects
Milan (Italy)	116
Bologna (Italy)	106
Rome (Italy)	89
Italy	***311***
Oxford (UK)	126
Exeter (UK)	120
UHRL (UK)	114
UK	***360***
Stockholm (Sweden)	227
Gothenburg	99
Sweden	***327***

Table 3. Participants characteristics

	Italy	UK	Sweden		Pooled sample
Age (mean)	23.9	24.3	27.8	*Δ(UK_Sweden)*** *Δ(Italy_Sweden)***	25.3 s.d. 8.6
Male (%)	52.9	56.2	56.3		55.2
Employed (%)	22	31	43.4	*Δ(Italy_UK)*** *Δ(UK_Sweden)*** *Δ(Italy_Sweden)***	32.3
Major Econ (%)	42.2	24.2	13.4	*Δ(Italy_UK)*** *Δ(UK_Sweden)*** *Δ(Italy_Sweden)***	26.4
Past Participation (%)	81.2	83.7	72.8	*Δ(UK_Sweden)*** *Δ(Italy_Sweden)**	79.3
Prosocial (%)	53.7	56.1	87.1	*Δ(UK_Sweden)*** *Δ(Italy_Sweden)***	65.3
Risk (mean)	5.17	6.28	5.75	*Δ(Italy_UK)*** *Δ(UK_Sweden)*** *Δ(Italy_Sweden)***	5.77 s.d. 2.27
Other Total Compliance (%)	6.1	11	21.7	*Δ(Italy_UK)** *Δ(UK_Sweden)*** *Δ(Italy_Sweden)***	13.04

For Age and Risk, we performed a series of t-tests. In all other cases, differences are detected through a series of chi2 tests

*** 1% significance ** 5% significance

Table 4. Random – effect Tobit Regression: level of compliance in the first phase

	Dependent variable: percentage of income declared in NOPOT, POT and DOUBLEPOT rounds Baseline: POT		
	Italy	UK	Sweden
No Pot	-0.314***	-0.281***	-0.607***
Double Pot	0.516***	0.618***	0.932***
Age	0.035*	0.001	0.022*
Male	-0.451***	-0.564***	-1.021***
Employed	-0.263*	-0.033	-0.097
Econ Major	-0.255**	-0.228	-0.345
Past Participation	-0.254	-0.54***	-0.584***
Prosocial	0.404***	0.732***	1.153***
Risk	-0.105***	-0.139***	-0.136***
Others Report: Total compliance	0.918***	0.849***	0.756***
High earner	0.008	-0.173	-0.087
Low earner	-0.318*	0.059	0.385
Constant	0.967*	1.393***	1.178**
Left-censored obs	205	423	253
Uncensored obs	247	273	168
Right-censored obs	430	339	542

*** 1% significance ** 5% significance * 10% significance

Table 5. Random – effect Tobit Regression: level of compliance in the second phase

Dependent variable: percentage of income declared in TAX10, TAX30 and TAX50 rounds
Baseline: TAX30

	Italy	UK	Sweden
Tax10	0.479***	0.227***	0.531***
Tax50	-0.226***	-0.106	-0.363***
Age	0.039*	0.008	0.031**
Male	-0.428***	-0.525***	-0.874***
Employed	0.047	0.127	-0.136
Econ Major	-0.188	-0.385**	-0.386
Past Participation	-0.14	-0.344*	-0.266
Prosocial	0.494***	0.769***	1.069***
Risk	-0.06**	-0.105***	-0.086*
Others Report:			
Total compliance	1.345***	0.762***	1.008***
High earner	0.251	-0.255	0.038
Low earner	-0.151	-0.076	0.267
Constant	0.337	1.19***	0.613
Left-censored obs	194	351	221
Uncensored obs	248	269	170
Right-censored obs	440	415	569

*** 1% significance ** 5% significance * 10% significance

Table 6. Tobit Regression: average level of compliance in both phases

Dependent variable: average level of compliance over the 6 rounds

	(1)	(2)
Sweden	0.073***	
Italy	0.144***	0.072***
UK		-0.073***
Constant	0.649***	0.722***
N. obs	959	959
Controls	YES	YES

*** 1% significance ** 5% significance * 10% significance

Table 7. Differences across Countries in reaction to different levels of Efficiency

	ITALY	UK	SWEDEN	DIff ITALY-UK	Diff ITALY-SWEDEN	Diff UK-SWEDEN
(POT - NOPOT)	0.086	0.083	0.126	0.003	-0.04	-0.042
(DOUBLEPOT- POT)	0.14	0.166	0.158	-0.026	-0.018	0.007
(DOUBLEPOT-NOPOT)	0.226	0.249	0.284	-0.023	-0.058*	-0.035

We performed a series of t-tests. The null hypotheses are Italy=UK, Italy=Sweden and UK=Sweden respectively
*** 1% significance ** 5% significance * 10% significance

Table 8. Differences across Countries in reaction to different levels of Tax rate

	ITALY	UK	SWEDEN	Diff ITALY-UK	Diff ITALY-SWEDEN	Diff UK-SWEDEN
(TAX30 – TAX10)	-0.122	-0.054	-0.07	-0.068**	-0.052*	0.016
(TAX50 – TAX30)	-0.084	-0.027	-0.078	-0.057**	-0.006	0.051*
(TAX50 – TAX10)	-0.207	-0.081	-0.148	-0.126***	-0.058*	0.067**

We performed a series of t-tests. The null hypotheses are Italy=UK, Italy=Sweden and UK=Sweden respectively
*** 1% significance ** 5% significance * 10% significance

Table 9. OLS Regressions: Differences across Countries in intensity of reaction to different levels of Efficiency and Tax rate

	Dependent variable			
	DIFFPOT		DIFFTAX	
	(1)	(2)	(3)	(4)
Sweden	0.032		-0.096**	
Italy	-0.018	-0.05	-0.18***	-0.084*
UK		-0.032		0.096**
ATTC_Index	0.019	0.018	0.051**	0.051**
Constant	0.151**	0.182**	-0.118	-0.213***
Controls	YES	YES	YES	YES
N. obs[a]	860	860	859	859

*** 1% significance ** 5% significance * 10% significance

a Milan is no more in the sample since in the original questionnaire they did not have to answer all the questions we are computing the ATTC_Index on.