Attitude–behavior consistency in tax compliance: A cross-national comparison

Attitude-Behavior Consistency in Taxation: A Cross-National Comparison

Alice Guerra[a†] and Brooke Harrington[a]

[a] Copenhagen Business School, Denmark

March 15, 2018

Abstract

Are individuals' attitudes about paying taxes consistent with their behavior? A direct link between attitudes (tax morale) and behavior (tax compliance) has long been assumed, despite an extensive social scientific literature attesting to the generally weak congruence between the two. This study builds on an emerging body of work questioning the link between tax morale and compliance. It innovates with a cross-national experimental research design whose results indicate that populations with high levels of tax morale exhibit higher evasion rates than those with low levels of tax morale; and individual self-reported tax morale cannot predict actual evasion choices. Methodologically, the paper contributes the first results of a laboratory experiment on taxation in Denmark, comparing them to laboratory findings from Italy; while previous research indicates that these two countries lie at opposite extremes in tax compliance and morale, our findings run contrary to "culturalist" explanations. Although we find that Danes do indeed have higher tax morale than Italians, and are less tolerant of tax evasion by others, our results also show that Danes are more likely to evade tax than Italians, and evade in larger monetary amounts. Finally, the study overcomes methodological flaws in previous studies by measuring risk aversion—a factor known to influence tax evasion, but often confounded with national stereotypes—through observed behavior rather than through self-reporting.

Keywords: tax compliance, tax morale, cross-national comparison
JEL Codes: H2; H3

[†] Corresponding author at: Department of Business and Politics, Copenhagen Business School, Porcelaenshaven 18B, 2000 Frederiksberg, Denmark.
Email addresses: agu.dbp@dbp.dk (A. Guerra); bh.dbp@dbp.dk (B. Harrington).

1. Introduction

Are individuals' attitudes about paying taxes consistent with their behavior? An extensive body of research in social science, dating back to the 1930s, suggests that attitude-behavior consistency is weak (LaPiere, 1934; Kutner et al., 1952; Deutscher, 1966; Liska, 1974). Although individuals are motivated to align their behaviors with their beliefs (Bandura, 2002), empirical studies suggest that such congruence is rarely achieved—particularly when it comes to questions of morality and ethics (White, 2002). Yet taxation research has long been predicated on the opposite claim: that the behavior of paying tax (*tax compliance*) follows directly from attitudes, particularly the intrinsic motivation to pay (*tax morale*) (Riahi-Belkaoui, 2004; Cummings et al., 2009). In fact, tax morale is viewed as so closely tied to behavior as to serve as a proxy for action (Halla, 2012).

However, an emerging stream of taxation research has begun to question this linkage between attitudes and actions (e.g., Casal and Mittone, 2016). For example, a series of studies in Europe, North America and South America have exposed the negative correlation between tax morale and the shadow economy: in those regions, motivation to pay taxes is high, but compliance is low (Torgler, 2005; Alm and Torgler, 2006; Barone and Mocetti, 2009). The present study advances and builds on this literature by comparing attitude-behavior consistency in an experimental setting. The results are novel, showing that high levels of tax morale are associated with high evasion rates and individual self-reported tax morale cannot predict actual evasion choices.

Empirically, the study draws from data collected in two countries that previous research has identified as representing opposite extremes on tax morale and compliance: Denmark and Italy. Previous research has identified the former as having exceptionally high tax morale and one of the highest rates of tax compliance in the world, while Italians have among the lowest rates of morale and compliance in Europe (Alm and Torgler, 2006). In addition, the two countries appear to stand at opposite poles in terms of underlying social norms that might affect taxation-related attitudes and behavior, such as interpersonal trust and honesty—Danes rank very high internationally on both items (Rothstein and Uslaner, 2005), while Italians rank quite low (Floris, 2010).

But while Denmark and Italy appear to represent straightforward cases of tax compliance stemming directly from tax morale, recent research has complicated the attitude-behavior relationship, making this an analytically fruitful case to explore. For example, research on Denmark suggests that Danes' high rates of tax compliance may stem not from high tax morale, but from state-level institutions that make it nearly impossible to evade (Kleven et al., 2011). By the same token, recent work suggests that Italians are *more* tax compliant than some Northern Europeans, despite having lower levels of tax morale (Andrighetto et al., 2016; Zhang et al., 2016).

Building on these suggestive findings, this paper makes three contributions to taxation scholarship. First, it offers an unusual cross-national analysis of the relationship between morale and compliance, in contrast to studies that have focused on single countries (Barone and Mocetti, 2009; Casal and Mittone, 2016). Second, unlike previous work that looked either a positive or negative relationship between

taxation attitudes and behaviors (Alm and Torgler, 2006), this study examines both. Finally, this paper advances a recent stream of research calling into question national stereotypes and "culturalist" explanations for variations in tax compliance (Andrighetto et al., 2016; Zhang et al., 2016).

Methodologically, the paper innovates in two ways. First, it is the first experimental study of tax morale and compliance in Denmark; the study combines the experimental data on compliance with survey-based data on morale. Second, the experimental design included a risk aversion elicitation task, to test the impact of individual risk aversion on task evasion—a potentially confounding variable that has been either neglected (e.g., Alm et al., 2017) or improperly measured (e.g., Andrighetto et al., 2016) in previous research. In the absence of valid risk aversion measures, it has been impossible to argue whether observed differences across subject pools are due to cultural differences or intrinsic motivations. This paper thus overcomes a major methodological flaw in previous research, allowing scholars to untangle competing sources of attitudes toward taxation and tax evasion.

The remainder of this paper is organized as follows. Section 2 describes the methodology, including the experimental design and procedure, the subject pool and the econometric method. Section 3 presents the results. Section 4 discusses the main findings, possible limitations of our analysis, and directions for future research. The Appendix contains tables and figures.

2. Methodology

We used a laboratory experiment to analyze tax compliance behavior, through experimental sessions in Italy and in Denmark. We chose laboratory experiments because, unlike field experiments, this methodology allowed us to measure actual evasion choices and compare them between countries while controlling for formal institutions such as tax rates, audit probabilities and fines. The experimental design and procedure were identical in the two countries. This allowed us to eliminate the effects of the vast institutional differences between Italy and Denmark in terms of the two nations' taxation apparatus and policies, and to focus purely on attitude-behavior consistency among individuals.

2.1. Experimental Design

The experiment involved three main parts. The first part consisted of the tax compliance experiment. The second part was a risk aversion elicitation task, involving lottery choices à la Holt and Laury (2002). The third and final part was a questionnaire eliciting information about individual characteristics and attitudes towards taxation.

Our experimental design follows the basic outlines of recent work by Andrighetto et al. (2016) and Zhang et al. (2016). In our study, as in theirs, subjects earned income through a clerical task and then were asked to decide how much to declare under different audit probabilities; this is the typical approach of most tax compliance experiments (Alm and Jacobson, 2007; Alm, 2010; Doerrenberg and Duncan, 2014; Alm et al., 2017). However, our design innovated in several ways compared to previous work. First, we included a risk elicitation task to test for risk aversion—a known influence on tax evasion decisions which has been improperly measured in past research through self-reporting. In addition, our

design differed from that of Andrighetto et al. (2016) and Zhang et al. (2016) by including an extensive set of explanatory variables, along with questions about the perceived compliance of other experimental participants after each round of the experiment—instead of at the end of the whole experimental procedure.

The Tax Compliance Experiment. Participants in this study earned income through a clerical task, and chose how much income to self-report to the tax authority. The tax authority was simulated by the computer. In each experimental round, participants were informed of the tax rate, the probability of being audited, and the planned redistribution of the tax revenues. If auditing uncovered under-declaration of income, both unpaid taxes and fines were collected.

There were three stages in this part of the experiment, with each stage divided into three rounds (see *Table 1* in the Appendix). Each stage was completely independent from the others, meaning that the choices made in each stage had no effect on the earnings in the other stages. Subjects were paid their after-tax earnings at the end of the experiment.

At the beginning of each stage, participants were asked to perform a real-effort task for three minutes. This task, intended to proxy labor supply and to induce a feeling of ownership of income, was designed by Andrighetto et al. (2016) and Zhang et al. (2016). It consisted of data entry: copying information about fictional students from a sheet of paper onto the computer.[1] For each row of information copied correctly, participants received 10 points. This task had a number of advantages: it was easy to explain and implement, did not allow guessing, did not require any prior knowledge or expertise, and was identical across stages, treatments and subjects.

At the end of each clerical task, subjects were told how many points they had earned as income. As is standard in the experimental tax evasion literature, subjects were then asked to declare any amount between 0 and true gross income under different scenarios. In the instructions, participants did not know the details of the following stage's tax scenario until the end of the previous stage. This option was chosen because we wanted to avoid any influence from subjects' performance and decisions in one stage "leaking" into the subsequent states of the experiment.

Each scenario constituted a round, which was defined by different taxation rules and redistribution mechanisms (see *Table 1* in the Appendix). From the beginning of the experiment, subjects were informed that they were free to report any amount, from 0% to 100%, that only the reported earnings would be taxed, and that there was a risk of being audited at the end of the study.[2] For each round, if subjects were caught under-reporting actual earnings, they had to pay a fine equal twice the tax they should have paid. This is a commonly used penalty structure in the experimental literature on tax evasion. Participants were informed that at the end of the experiment, for each individual and each round, the computer would randomly select a number between 0 and 100. If the number were between 0

[1] If the computer detected a mistake, an error message appeared on the participant's screen and he/she had to correct the mistake before proceeding. On another labor task stage in tax experiment, see Doerrenberg and Duncan (2014).

[2] Before being asked to report their income, participants were given explicit examples under hypothetical decisions to ensure their understanding of the instructions.

and the announced probability of being audited, then the individual would be audited. The automated and randomized nature of this process assured participants that they had a credible opportunity to evade, and that audits were not controlled or manipulated by the experimenters.

Finally, in each round, immediately after reporting income, each participant was asked to estimate the compliance choices of other participants in the room. That is, individuals had to state how many others they believed had declared their true earnings for tax purposes. The possible answers were: none; less than half; about half; more than half; almost everyone; and everyone.

As described in *Table 1*, the different tax scenarios unfolded as follows. In all three rounds of Stage 1, participants faced a 30% tax rate, a 5% probability of being audited, a fine of '2Xunderreported income' and varying redistribution mechanisms. The revenues collected were not redistributed at all in Round 1; in Round 2, revenues were equally redistributed among participants; and in Round 3, revenues were first doubled and then equally redistributed among participants. In all three rounds of Stage 2, participants faced the same probability of being audited, and the same fine and redistribution scheme as in Stage 1; only the tax rate varied, from 5% in Round 4, to 30% in Round 5, and then 50% in Round 6. Finally, in all three rounds of Stage 3, the tax rate, the fine, and the redistribution scheme were identical to those in Stage 1; only the probability of being audited varied, from 5% in Round 7, to 30% in Round 8, to 50% in Round 9.[3]

Audit results were revealed only at the end of the experiment. Importantly, at no point during the experiment did participants have information about the others' declaration of income or their audit rates. This procedure was implemented to avoid influences of reciprocity, conditional cooperation, or reputation and wealth effects, all of which lie beyond the scope of this research.

For the sake of realism (that is, to simulate as closely as possible taxpayers' motives and decision-making), and in line with previous experiments on tax compliance, the experiment was framed in non-neutral terms, using words such as "tax," "income," "audit," and "report." However, it avoided loaded terms such as "cheating" or "lying."[4] Using non-neutral language had two advantages for our design. First, it avoided giving participants the impression that the experiment was a risky gamble instead of a tax compliance decision. Second, it ensured that there was no ambiguity about what honest behavior was: that is, to declare the total amount of earnings. Together with this framing choice, the fact that subjects had to earn real income and pay real fines aligned the experiment with the behavior of interest in the real world.[5]

[3] We considered randomizing the order of the different stages and rounds. However, as pointed out by Andrighetto et al. (2016) and Zhang et al. (2016), such randomization is not necessary for the scope of this research, which is not to evaluate the effects of institutional changes, but rather to analyze how individuals from different countries would respond to the same institutional scenario.

[4] Instructions with tax-specific language are generally used in tax compliance experiments to avoid subjects perceiving the experiment as a risky-choice gamble rather than a tax declaration situation. See, among others, Cummings et al. (2009), Andrighetto et al. (2016), Zhang et al. (2016) and Alm et al. (2017).

[5] For a similar design, see also Bühren and Kundt (2014) and Bühren and Pleßner (2014).

The Risk Elicitation Task. Following the three stages of the main tax compliance experiment, we next carried out a lottery-choice task based on Holt and Laury (2002) to elicit individual risk attitudes.[6] The test involved 10 choices between pairs of two-outcome options (or lotteries), as shown in *Table 2* in the Appendix. In Option A, the possible outcomes were 16 points and 20 points (low variance–low risk). In Option B, the possible outcomes were 1 point and 38.50 points (high variance–high risk). The probability of receiving the higher payoff increased from decision 1 to decision 10, so that expected value initially favored Option A but reversed at decision 5, finally leaving Option B dominant at decision 10. A strongly risk-seeking participant would select Option B throughout. The choice profile of a coherent decision maker is a vector of 10 choices, beginning with Option A and shifting at some point to Option B. A risk neutral participant would switch from A to B at decision 5, where higher switching points indicate greater risk aversion. Decision 10 provided a check on comprehension and attentive responding, since it offered 38.50 points guaranteed (Option B) versus 20 points never guaranteed (Option A). Our measure of risk aversion (operationalized as the variable "risk aversion" in our estimates) was constructed from the number of times an individual picked the safer choice, Option A.

After the risk aversion test, participants were informed about the results in each round of the experiment. That is, they learned of the tax on income they declared, and the benefits they received from the tax-funded common pool. They were also informed whether they had been audited, and if so, what fine they had to pay in case they had underreported their earnings.

The Post-Experimental Questionnaire. In the final stage of the study, subjects were asked to complete a questionnaire that included questions about their gender, nationality, previous participation in experimental research, their household monthly income, and other factors potentially correlated with their income declaration choices. The questions were based in part upon the European Values Study 2008,[7] which includes information about trust in other people, justifications for not paying taxes, religious denomination and church attendance.

2.2. Experimental Procedure

The experimental sessions were conducted during the academic year 2016/2017 at the University of Bologna's Laboratory for Experiments in Social Science (BLESS) in Italy, and at Copenhagen University's Laboratory for Experimental Economics (LEE) in Denmark.[8] The recruitment process, the experimental design and procedure were the same in each site and in each session. To ensure

[6] For other experimental studies that used Holt and Laury's (2002) test on risk aversion, see Kugler et al. (2012). On measuring individual risk attitudes, see Rabin (2000) and Eckel and Grossman (2008).

[7] Available at: https://dbk.gesis.org/dbksearch/file.asp?file=ZA4800_cdb.pdf. Last access: March 15, 2018.

[8] The experiment was approved by the Ethics Committee of the Danish Council for Independent Research. The experiment was conducted in line with the ethics policies of both BLESS and LEE. At the beginning of each experimental session, each participant was asked to fill out and sign an informed consent document outlining the rules under which the experimental lab operated, with information on the voluntary nature of the study that they were participating in and the processing of their personal data. Participants were permitted to opt out at any time with no penalty and were allowed to withdraw their data subsequent to participating. Both BLESS and LEE operate under the methodological paradigm of experimental economics, where participant deception is not allowed.

consistency, the experimental instructions were translated and back-translated between Italian (for the sessions in Bologna) and English (for the sessions in Denmark).

Participants were recruited using ORSEE (Greiner, 2004, 2015), a web-based Online Recruitment System for Economic Experiments specifically designed to optimize recruitment for economic experiments. Since the aim of the study was to examine determinants of tax compliance, including morale and other cultural factors, we recruited only native students—meaning individuals born in Denmark (Italy) to Danish (Italian) parents. Neither the aims nor the cross-cultural scope of the research project were stated in the invitation email to participants or during the experimental sessions.

Participants were undergraduate and graduate students from various fields (including social sciences, humanities and medicine).[9] Such students are the typical sample group in economics experiments. Though they are unlikely to have paid income tax, this does not undermine the external validity of the results (Alm et al., 2015; Choo et al., 2016).[10]

The experiments were designed using *zTree* (Fischbacher, 2007) and participants performed all the experimental tasks via computer. Computer clients in both labs are partitioned to ensure confidentially and avoid communication between participants. To ensure anonymity and reduce the feeling of scrutiny, participants were informed that their decisions during the experiment and their final payment would be kept confidential and linked to their client ID number, rather than to their names. We assured them that neither the experimenters nor anyone else would be able to link participants to their individual choices. Questions were answered in private and payments were issued in cash at the end of the session to each participant individually, to further ensure confidentiality and anonymity.

On the day of the experiment, participants were given a random clientID number and assigned to the corresponding computer client. Once everyone was seated, the experimenters handed out the clerical task packet and began reading the instructions, which were also shown on each participant's screen. Participants were informed that—based on the others' and their own choices, as well as on chance—they would earn points, to be converted at the end of the experiment into the local currency (Euros in Italy and Danish Kroner in Denmark). The exchange rate was calibrated such that the average payment to participants per hour (including time to read the instructions and payment of participants) would be approximately equal to the average hourly wage for student employment in the local context.[11] In

[9] Both BLESS and LEE have existing databases of active participants. These participants are undergraduate and graduate students. They have expressed interest in participating in behavioral experiments by registering their personal details in ORSEE. This includes their name, gender, degree programme (if applicable), and e-mail address. Both labs have strict policies on the confidentiality of the data provided by our participants.

[10] Recently, Choo et al. (2016) compared tax compliance choices among three distinct populations: undergraduate students, individuals in full-time employment who pay income tax through a third-party reporting system, and individuals who are self-employed and thus self-report their income tax liabilities. They found students as being the least compliant subject pool, but also the most responsive to treatment changes, particularly to ambiguity in the audit probability. In contrast, self-employed taxpayers and taxpayers who paid through third-party reporting were more compliant and mostly non-responsive to differing conditions.

[11] At Copenhagen University, LEE requires external researchers to calibrate the exchange rate from earnings in experimental currency units to kroner such that the average payment to participants per hour is at least 120 DKK per hour. Our experimental sessions lasted 1 hour and a half each, so the average payment to participants was 191.804 DKK.

addition, each participant received a show-up fee for participation (5 Euros in Italy and 100 Danish Kroner—approximately equivalent to 14 Euros–in Denmark).[12] Volunteers who exceed the number of slots available for the session received the show-up fee. Each session lasted about 90 minutes and participants earned an average of 10.641 Euros in Italy and 191.804 DKK (equivalent to 25 Euros) in Denmark.

2.3. Subject Pool

The overall experiment consisted of 16 sessions, of which 6 took place in Italy and 10 in Denmark.[13] Since the number of participants per session might impact individual tax behavior, we controlled in the regression analysis for the number of participants per session (variable "*Pool Size*"). The results indicate that subject pool size had no effect on individual tax behavior.

A total of 180 participants (53.9% male and 46.1% female) took part in the experiment, of which 106 (58.89%) were from Italy (41.51% from the North) and 74 (41.11%) were from Denmark (39.19% from Copenhagen area). Participants were mainly in the 20-24 years old age bracket, and were mainly undergraduates. The undergraduates' majors included Social Sciences (22.78%), Mathematical, Physical and Natural Sciences (11.11%), Engineering (9.44%), Humanities, Philosophy and Arts (8.89%).

The majority of participants (73.33% of all subjects) declared a household monthly net income lower than € 3.000 / kr. 22.000. Moreover, the majority of participants did not regularly attend religious services: only 12% attended religious services at least once per month; another 21.11% attended only on special occasions; and 53.89% never attended religious services at all. Moreover, 61.05% of participants described themselves as atheist or agnostic; another 38.33% described themselves as Roman Catholic, and 21.67% described themselves as Protestant. In our estimates, we followed the standard practice of including demographic controls.

2.4. Econometric Method

In our analysis, we estimate the determinants of tax evasion by using pooled and panel double-hurdle (DH) models.[14] Both models capture important characteristics of our experimental data. Hence, we report results from both specifications.

---

[12] In Bologna, BLESS requires external researchers to guarantee a minimum payment to all invited volunteers who show up on time, and also to volunteers who exceed the number of slots available for the session. LEE does not require a show-up fee, but researchers have to guarantee that only few participants earn less than 50 DKK in total.

[13] The higher number of sessions in Denmark is due to some difficulties recruiting Danish participants. Indeed, despite inviting more than 800 Danes per session via ORSEE, the number of Danish participants at LEE ranged from 4 to 14, whereas the number of Italian participants at BLESS was consistently higher (from a minimum of 18 to a maximum of 25).

[14] Several experimental studies on tax evasion have used the Tobit model to estimate relationships between censored dependent variables (i.e. declared income) and the relevant covariates (e.g. Andrighetto et al. 2016; Zhang et al., 2016). However, as pointed out by Alm et al. (2017), the standard assumption of the censored Tobit regression model—that is, that variables which influence individual choice to evade taxes have the same effect on *whether* a subject evades and *by how much*—are not appropriate to analyze compliance behavior at the extensive and intensive margins.

This empirical strategy, introduced by Cragg (1971), allows us to estimate two distinct processes: "the first hurdle, which can be interpreted as a probit model, determines whether or not a person participates in evasion and is particularly suited to capture effect occurring mainly at the extensive margin; the second hurdle, which can be interpreted as a Tobit model, determines the level of evasion only for those people who ever choose to evade and is therefore relevant for the effect occurring at the intensive margin" (Alm et al., 2017, p.20).

To estimate the DH models, we used the two-stage procedure developed by Engel and Moffat (2014) where subjects must cross two hurdles. The first hurdle needs to be crossed to be an evader. Once the subject becomes an evader, his or her current circumstances, along with the experimental treatments, affect whether s/he contributes (this is the second hurdle). It follows that the double-hurdle model contains two equations, which can be interpreted as a combined Probit and Tobit estimator.

Formally, following Engel and Moffat (2014) and Alm et al. (2017), the observed evasion rate is given as follows:

(1) $ER_{ij} = d_{ij}\, ER^*_{ij}$

where ER_{ij} is the observed evasion rate of subject *i* in round *j*; d_{ij} and ER^*_{ij} represent the first and the second hurdle, respectively, which are defined in the following.

The first hurdle is represented by a binary variable for evasion as follows:

(2) $d_{ij} = \begin{cases} 1 \text{ if } d^*_{ij} > 0 \\ 0 \text{ otherwise} \end{cases}$

The latent variable d^*_{ij} is given as follows:

(3) $d^*_{ij} = z'_{ij}\,\gamma + \varepsilon_{1,ij}$

where $\varepsilon_{1,ij}$ is subject *i*'s idiosyncratic propensity to pass the hurdle in round *j*, which is assumed to be normally distributed with zero mean and variance normalized to unity.[15]

The second hurdle, which is similar to the Tobit model, is given as follows:

(4) $ER^*_{ij} = \max\{ER^{**}_{ij}, 0\}$

where the latent variable ER^{**}_{ij} is given as:

(5) $ER^{**}_{ij} = \boldsymbol{x}'_{ij}\,\beta_1 + \mathbf{y}_i\,\beta_2 + \alpha_i + \varepsilon_{2,ij}$

The regressor vector $\mathbf{x}_{ij}$ includes three variables for the parameters of the experiment (redistribution, tax rate, audit probability) and the intercept. The regressor vector $\mathbf{y}_i$ includes individual-specific

[15] This is usually required for identification in Probit estimates given that the outcome of the first hurdle is binary.

characteristics (e.g. age, gender, education, and others) and other socio-economic and behavioral factors (e.g. tolerance of tax evasion, perceived behavior of the other participants, degree of risk aversion). *Table 5* in the Appendix describes the list of the variables included in the econometric analysis.

The pooled DH model includes only the contemporaneous error term $\varepsilon_{2,ij}$ (which is assumed to be normally distributed, with zero mean and variance σ_{ε}^{2}), while the subject-specific random effects α_i are assumed to be zero. In the pooled HD model the two error terms $\varepsilon_{1,ij}$ and $\varepsilon_{2,ij}$ are assumed to be independently distributed.

Conversely, the panel DH model includes the subject-specific random effects α_i (which is assumed to be normally distributed, with zero mean and variance σ_{α}^{2}), which measures subject i's idiosyncratic propensity to evade, conditional on passing the first hurdle. In this way, the panel HD model controls for potential dependence for the repeated decisions made by participants by capturing the correlation between the two hurdles (i.e., the correlation between the random effects α_i and the error term $\varepsilon_{1,ij}$, i.e., $\rho = corr(\alpha_i, \varepsilon_{1,ij})$).[16]

As pointed out by Alm et al. (2017), there is one crucial feature that distinguishes the panel DH model from the pooled DH model. Contrarily to the pooled DH model, in the panel DH model a subject is classified as fully compliant (a zero-type subject) only if he/she fully declares his/her true earnings in *all* the rounds of the experiment. Otherwise, the subject is identified as an evader. This means that a subject who fully declares his/her earnings in some, but not all, rounds is classified as an evader. The same holds for a subject who evaded in some or all the rounds. Hence, given the possible presence of zero-type subjects in the experiment, the panel DH model has the potential to significantly improve the estimates.

3. Results

As a general overview, our data reveal that across all subjects, stages, periods, and countries, the mean reporting compliance rate averages 61.16%. This overall level of compliance, which far exceeds the levels predicted by expected utility theory (Allingham and Sandmo, 1972; Yitzhaki, 1974), is in line with the experimental literature on tax compliance and public goods (Bosco and Mittone, 1997; Cummings et al., 2009; Alm, 2012; Andrighetto et al., 2016; Casal and Mittone, 2016; Zhang et al., 2016; Alm et al., 2017). When considering the full evasion rate, we found that subjects fully evade taxes 27.34% of the time.

In the following, we present more detailed descriptive statistics (Section 3.1), estimates from regression analyses (Section 3.2), and results from the post-experimental questionnaire. Some have raised questions about the validity of such questionnaires, and the potential of responses to be influenced by participants' behavior in the preceding experiment: specifically, the prediction is that individuals' questionnaire responses would be biased toward alignment with and moral justification of their evasion behavior in the

[16] For an extensive discussion on pooled and panel DH models, see Alm et al. (2017).

experimental rounds (Halla, 2012). We find no evidence of this. In fact, we find the opposite: a mis-alignment of tax morale and tax compliance in *both* directions.

That is, those with the highest tax morale (the Danes) comply the least (evading the most), while those with the lowest tax morale (the Italians) comply the most (evading the least). Furthermore, the survey revealed that the Italians, despite evading less, are *more* tolerant of others' evasion; Danes showed the contrary pattern. In other words, our findings are the reverse of what we would expect to find if participants regarded the post-experimental questionnaire as an opportunity to justify themselves in their own eyes. Thus, there is a basis for confidence that the questionnaire data are valid and not biased by being collected after the experiment.

3.1. Descriptive Statistics

Evasion rates varied widely depending upon the specific tax scenarios presented in each round. Overall, subjects responded in a predictable manner to changes in the classical economic parameters (i.e. redistribution, tax rate, probability of being audited). *Table 3* in the Appendix reports the summary statistics of evasion rates — defined as the ratio between undeclared income and total earned income— in each of the nine rounds, for each country. The average evasion rate is negatively associated with the redistribution policy (rounds 1-3) and with the probability of being audited (rounds 7-9). Subjects responded to higher tax rates by evading more (rounds 4-6).

Since all subjects faced all three treatments, we report results from a Wilcoxon test for these comparisons (see *Table 3*). Pooling across countries and sessions, we observe that an increase in the probability of being audited from low (5%) to high (50%) reduces the evasion rate from 50.73% to 8.78% (p=0.000). A similar pattern follows when redistribution increases from low (no redistribution) to high (redistribution x2): the evasion rate decreases from 58.85% to 25.21% (p=0.000). An increase in the tax rate from low (5%) to high (50%) has a positive impact on evasion: the evasion rate increases from 33.53% to 54.14% (p=0.000). Overall, the effect of audits on evasion rates is larger if compared to increases in the redistribution rate of the tax-funded common pool or in the tax rate. Similar increments are present in the two countries individually.

It is worth noting that considering treatment averages obviously ignores the great degree of heterogeneity in our data. *Figure 1* in the Appendix illustrates that, for each country and round, two types of behavior emerge: individuals who declared 100% of their true income and, in contrast, individuals who completely evaded their taxes. *Figure 1* clearly shows that, in both countries, subjects are very likely to be fully compliant in round 9, when the probability of being audited is 50%. Full evasion rates decrease as redistribution and audit probability increases, whereas higher tax rates increase the percentage of full evaders.

Since our main interest is the effect of morale and culture on compliance, we shall compare tax behavior between subject pools, Danes *versus* Italians, in each round (i.e. keeping experimental parameters constant). *Table 3* in the Appendix shows that, for each round (except for rounds 3 "Redistribution X2", 8 "Audit Prob 30%" and 9 "Audit Prob 50%"), the evasion rate is lower for Italians rather than for

Danes. Interestingly, the difference in compliance behaviors between the two countries is no more statistically significant in the presence of high audit rates ("Audit Prob 30%" and "Audit Prob 50%") and a higher redistribution (i.e. "Redistribution X2"). More specifically, Danish participants failed to report roughly the 64.29% of their gross income when the audit probability was at 5%–but this percentage was reduced to 5.45% of income unreported when the audit probability was at 50%. The response of the Italian subjects to increases in audit probabilities was more nuanced. Italians failed to report 41.26% of their income when the audit probability was at 5%; when the audit probability was at 50%, Italians did not report 11.11% of their income. The stronger reaction of Danes to audit probability changes indicates that they are more sensitive than Italians both to enforcement measures, and to opportunities to evade taxes without being detected.

The overall difference between Italian subjects and Danish subjects can be clearly seen in *Figure 2* in the Appendix, which plots the cumulative distribution function of the "Evasion Rate" pooling across treatments. That the cumulative distribution function of the Danish subject pool is lower is consistent with the fact that the average evasion rate in the Danish subject pool is higher than the average reporting compliance rate in the Italian subject pool.[17] The CDFs of evasion rate between countries are almost parallel to each other, indicating that the difference in tax behavior holds true at all evasion levels (i.e. both at the extensive and intensive margins).

Let us now consider tax morale, which is measured in the post-experimental survey as the complement of tolerance for tax evasion. *Figure 3* in the Appendix plots the histogram of the degree of tolerance for tax evasion in Italy and Denmark. The histogram and kernel density for the Italian subject pool show a higher tail on the right, compared to the histogram for the Danish subject pool. This means that cheating on tax is more often justified by Italian subjects rather than by Danish subjects. To test whether the mean of tolerance for tax evasion is different between the two subject pools, we performed a two-sample t-test. We found that average tolerance for tax evasion was higher in the Italian sample than in the Danish sample, and this difference is statistically significant (two-sample t-test, $p = 0.0289$).

In this section, we have shown that average evasion rates differ between rounds and between countries. Moreover, in each round and country, compliance behaviors follow a two-modal distribution with two predominant patterns: zero evasion (or, full compliance) and full evasion. The simple descriptive statistics provide suggestive results, which need to be fully examined by appropriate econometric techniques, as we show in the next section.

3.2. Determinants of Evasion

Table 4 in the Appendix presents the regression analyses regarding the determinants of tax behavior.

Models (1) and (2) show the estimates from pooled and panel DH basic models; Models (3) and (4) add some complexities to the basic models by including interaction effects. Potential dependence for the

[17] The corresponding Kolmogorov-Smirnov and Epps-Singleton tests on the distributions of two samples over average compliance rates are statistically significant, meaning that the two subject pools differ significantly. Combined K-S: D=0.1579; p-value=0.000; corrected=0.000.

repeated decisions made by participants (9 rounds) is controlled by including random effects at the individual level in the panel DH models.

For each regression model, results from estimating the first and second hurdle equations are shown in separate columns. Estimating the first hurdle reveals how the regressors affect the probability that a subject is identified as an evader. Estimating the second hurdle reveals how the regressors affect the amount of evasion, conditional on a subject being an evader.

The dependent variable is the evasion rate, which is computed as 1 minus the reporting compliance rate, and ranges from 0 to 1. The explanatory variables are given by the dummy variable *Italy* and the experimental treatments (*Tax Rate*, *Audit*, *Redistribution*) as regressors, plus the following control variables: *Male* controls for gender; *Age* measures the age of the participant in years; *Education* is a discrete variable for the level of education (0 "junior high school;" 1 "high school;" 2 "bachelor's degree;" 3 "master's degree;" 4 "master;" 5 "Ph.D."); *No Previous Participation* is a dummy variable which takes into account whether the subject has never participated to other experimental studies; *Rows* is a discrete variables which measures the number of rows correctly copied in each clerical task; *Pool Size* is the number of participants per session.

The regression models include also the variables *Risk Aversion*, *Perception*, and *Tolerance of Tax Evasion*. *Risk Aversion* captures individual degree of risk aversion measured as the number of "Option A" choices in the risk aversion elicitation task. *Perception* is a discrete variable—ranging from 0 to 5—which measures individual beliefs about others' compliance choices. This is captured via the following question asked to each participant in each experimental round: "In your opinion, how many participants in this room declared their true earnings for tax purposes?". The possible answers were: none (*Perception=0*); less than a half (*Perception=1*); about a half (*Perception=2*); more than a half (*Perception=3*); almost everyone (*Perception=4*); everyone (*Perception=5*). *Tolerance of Tax Evasion* is a discrete variable which proxies individual tax morale. This is measured via the following question asked in the post-experimental questionnaire: "Can cheating on tax if you have the chance always be justified, never be justified, or something in between?" *Tolerance of Tax Evasion* ranges from 0 (never be justifies) to 10 (always be justified). *Table 5* describes the list of the variables included in the econometric analysis.

Let us consider the pooled DH model (1) in *Table 4*. The estimates show that the effects of tax rate and redistribution are statistically significant only in the first hurdle. This means that increasing the tax rate increases the probability of a subject being an evader but has no impact on the level of evasion. Similarly, increasing redistribution of the tax-funded common pool reduces the probability of a subject evading tax, but has no effect on the amount of evasion. In other words, increasing redistribution makes subjects less likely to evade, but has no effect on subjects who do evade. Increasing audit probability has a negative and statistically significant in both the hurdles, with a greater impact on the first one.

The effect of the dummy variable for Italy is negative and statistically significant in both hurdles, with a greater impact on the first one (the decision to evade). Italian subjects were less likely to evade tax than

Danish subjects and, among those classified as evaders, Italians evaded by a lesser amount. More generally, we found that the amount of evasion decreases if subjects believed that the other subjects were truthfully declaring their income (see the variable *Perception*). This is in line with previous survey-based research on the effect of perceived tax evasion by other citizens on individuals' tax morale (e.g. Torgler and Schneider, 2007).

More interestingly, we found that the impact of tax morale on actual evasion choices is relatively small in magnitude compared to that of perceived compliance by others and the experimental treatments (see the variable *Tolerance of Tax Evasion*). Moreover, the coefficient is even smaller in magnitude and not statistically significant in the second hurdle. This suggests that self-reported tax morale do not substantially predict actual tax behavior, if compared to the other explanatory variables.

In addition, the decision to evade was affected by risk aversion: the higher an individuals' risk aversion, the less likely s/he was to evade. Despite being statistically significant, this effect was small in magnitude. Our finding here is similar to that of Cummings et al. (2009) from experiments in Botswana and South Africa. We also found that education level positively affects only the first hurdle, but not the second. This is partially consistent with Torgler et al. (2007), showing that higher education is positively correlated with tax evasion. Being female has a negative effect on both the probability and the level of evasion, whereas age has a negative and relatively small effect only on the first hurdle.

It is worth noting that all the regression models controlled for individual performance in the clerical task and subject pool size (see the variables *Rows* and *Pool Size*). We found these variables to have no impact on tax behavior. We also controlled for individuals' previous participation in experimental studies and found that subjects with no past experience in experiments were less likely to evade taxes, but, among evaders, they evaded to a greater extent (see the variable *No Previous Participation*).

The estimates for the panel DH model (2) are almost similar to those from the pooled DH model (1), even if the effects are not always as sharp as in the pooled DH.[18] In particular, in the panel DH model (2) the coefficient of *Italy* is still negative in both hurdles, but no more statistically significant. Similarly, in addition of being relatively small in magnitude, the coefficient of *Tolerance of Tax Evasion* is no more statistically significant. The overall fit is better in the panel DH model (2) than in the pooled DH model (1), as indicated by the log likelihoods and the values of σ_α and σ_ε.[19]

Models (3) and (4) show the estimates from additional regressions that add complexity to the DH pooled model by considering the interaction terms *Perception*Italy*, *Tolerance Of Tax Evasion*Italy*, and *Risk Aversion*Italy*. Let us consider the pooled DH model (3). The coefficient of the dummy variable Italy is negative and statistically significant in the first hurdle. We see that the dummy for *Italy* interacted with *Perception* shows a statistically significant coefficient positive in the first hurdle and negative in the

[18] The regressors in the first hurdle of the panel DH models do not include the variables *Tax Rate*, *Audit*, *Redistribution*, *Rows* and *Perception* (because the panel DH regression model has only one outcome per subject).

[19] Residual analysis in the pooled DH model (1) shows the importance of individual-specific random effects α_i, which contributes to 50% of the total variance (or $\sigma_\alpha^2/(\sigma_\alpha^2+\sigma_\varepsilon^2)=0.5$). This measure improves in the panel DH model (2), where individual-specific random effects contributes to 40% of the total variance.

second hurdle. This suggests that an increase in individual beliefs about others' compliance behavior had a smaller effect on evasion decisions for Italian subjects than for Danish subjects. In other words, interpersonal trust measured as individual perceived compliance of the others had a greater negative effect on individual evasion choices among Danish subjects than Italians. This is consistent with survey-based studies revealing that Northern European countries show higher levels of interpersonal trust than Southern European countries, and this is related to their unwillingness to evade taxes.

The dummy for *Italy* interacted with *Tolerance of Tax Evasion* shows a positive coefficient, which is relatively small in magnitude and statistically significant only in the first hurdle. This suggests that individual self-reported attitudes towards tax evasion do not substantially predict actual evasion choices.

Finally, the interaction between *Risk Aversion* and *Italy* has a negative, statistically significant impact on evasion choice in the second hurdle, but the effect is very small in magnitude. More interestingly, the coefficient for the variable *Risk Aversion* is no more statistically significant when the model includes the interaction term *Italy*Risk Aversion*. This is particularly relevant if compared to Cummings et al. (2009): since they found no differences in the degree of risk aversion between Botswana and South Africa, they argued that discrepancies in tax behavior revealed in the tax compliance experiment were only driven by cultural differences. In contrast, we cannot state the same conclusion when comparing Italy and Denmark: differences in tax behavior between Italian subjects and Danish subjects are also driven by different degrees of risk aversion.

The estimates for the panel DH model (4) are almost similar to those from the pooled DH model (3), although the coefficients of the interaction factors are not statistically significant. Possibly because of the difficulty in estimating the interaction effects given the size of the sample and the length of the panel, the overall fit is better in the pooled DH model (4) than in the panel DH model (3), as indicated by the log likelihoods and the values of σ_α and σ_ε.[20]

Consistent across all the four DH models we tested is that self-reported attitudes towards tax evasion have a small, not significant impact on actual evasion choices in most specifications, and Italians were less likely to evade taxes than Danes (although this effect is not statistically significant in the panel DH models). In contrast, the variables which have a significant and substantial effect on individual evasion choices are the audit probability, tax rates and gender. These results are largely discussed in the next section.

4. Discussion and Implications

This study contributes to a growing stream of research questioning long-held assumptions about attitude-behavior consistency in the realm of taxation (Halla, 2012). In fact, our experimental and survey data show that tax behaviors, far from being congruent with attitudes, may actually oppose them. The results indicate that populations with high levels of tax morale exhibit higher evasion rates than those

[20] As in the panel DH basic model (2), the regressors in the first hurdle of the panel DH model (4) do not include the variables *Tax Rate*, *Audit*, *Redistribution*, *Rows* and *Perception* (because the panel DH regression model has only one outcome per subject), and the interaction term *Risk Aversion*Italy* (for reasons of convergence).

with low levels of tax morale. In comparing participants from Denmark and Italy, this study also contributes to an emerging literature challenging "culturalist" arguments about tax morale and compliance (e.g., Zhang et al., 2016). By showing that Danes (characterized by high levels of tax morale) are highly non-compliant as taxpayers, and that Italians (characterized by low levels of tax morale) are more compliant than Danes, this paper illustrates how misleading national stereotypes can be for both scholars and policy-makers (Andrighetto et al., 2016).

Methodologically, the paper innovates in three ways. It provides the first experimental data on taxation gathered in Denmark, and sets it in cross-national comparison. Second, it overcomes a methodological flaw found in previous studies, by measuring risk aversion—a factor known to influence tax evasion, but often confounded with national stereotypes—through observed behavior rather than through self-reporting. Third, it contextualizes tax compliance (measured in the experimental rounds) with survey data on perceptions of *others*' compliance levels on both the extensive and intensive margins.

In the following sections, we elaborate on the study's main results and policy implications, as well as the generalizability of our findings, and possible directions for future research.

4.1. Implications for Theory and Practice

This study is positioned at a point of conflict between two streams of research literature. In answer to the question "are individuals' attitudes about paying taxes consistent with their behavior?" taxation research has posited continuity—e.g., that high tax morale drives high levels of tax compliance, and vice versa (Riahi-Belkaoui, 2004; Cummings et al., 2009; Halla, 2012). But this runs contrary to decades' worth of social scientific research indicating that in most contexts, attitude-behavior consistency is weak (LaPiere, 1934; Kutner et al., 1952; Deutscher, 1966; Liska, 1974). Why should we expect the empirical case of tax behavior to deviate from this long-standing pattern?

Our primary contribution to theory consists in examining this attitude-behavior linkage more closely, to see whether tax morale and tax compliance are related in the direction predicted by previous taxation research, or in the direction implied by the broader social science literature. Our results align with the latter, adding to an emerging stream of work that challenges long-standing assumptions in tax research (Barone and Mocetti, 2009; Casal and Mittone, 2016). Secondarily, our work calls into question the "culturalist" approach to taxation, joining a group of very recent studies that undermine national stereotypes that have historically classified Italians as cheaters and Danes as honest taxpayers (Andrighetto et al., 2016; Zhang et al., 2016; Alm et al., 2017).

From a practical point of view, our findings are quite encouraging for policy-makers. They suggest that traditional tax enforcement measures—including audits—along with redistribution policies and perception of others' compliance rates, have the most significant effects on individuals deciding whether to fulfill their fiscal obligations to states. Coupled with other recent studies, particularly Kleven et al. (2011), the results of this study support continued investment in detecting and sanctioning tax evasion.

Denmark, it seems, is not a special case of tax compliance culture that cannot be replicated elsewhere; nor is Italy a hopeless case. Kleven et al. (2011) found in a field experiment that Danes substantially evade their taxes when given the opportunity to do so; the study showed that what makes Denmark exceptional is not its culture or its tax morale, but the dominance of third-party reporting of income, which makes evasion nearly impossible. As for Italians, both Alm et al. (2017) and Zhang et al. (2016) found no differences between Italians, on the one hand, and American or British subjects on the other, in terms of propensity to evade tax in an experimental setting. This further supports the conclusion, implied by our own results, that police instruments (such as redistribution and the probability of being audited) are far more significant in explaining Danes' and Italians' real-world patterns of tax compliance than culture or tax morale.

4.2. External Validity and Generalizability

With all economics laboratory experiments, questions arise as to the validity of extrapolating findings to the real world.[21] Our research design sought to address these concerns. For example, to contend with the argument that lab settings are too artificial and abstract, our design included two features to make the tax evasion decision as realistic as possible. First, we used actual tax terminology and non-neutral terms to describe evasion. Second, we ensured that participants expended real effort in the experiment, and were rewarded with real money in return—as well as being fined in real money when they were caught evading tax. Thus, we created a reasonable proxy of real-world labor (Carpenter and Huet-Vaughn, 2017) and tax reporting (Doerrenberg and Duncan, 2014).

Some might question whether the stakes in the experiment were perceived by participants as being too small to elicit realistic responses. However, as Doerrenberg and Duncan (2014) have shown, many real-world evasion decisions made by individual involve relatively small amounts of money. Moreover, the average payoffs in both Italy and Denmark were calibrated to approximate the average hourly wage for student employment in the local context. Specifically, the average earnings of EUR 10.64 in Bologna and DKK 191.80 in Copenhagen correspond to at least two full lunch meals in the student canteens respectively at the University of Bologna and Copenhagen University. So while those amounts might not be significant for some individuals, they were meaningful in the context of our sample.

In relation to the sample, size and representativeness might be a concern. But considered against other experimental research on tax evasion, the subject pool of 180 participants used in this study—106 from Italy and 74 from Denmark—is at or above average. For comparison: Coricelli et al. (2010) recruited 48 subjects at the Groupe d'Analyse et de Théorie Economique (GATE) in Lyon (France); Kogler et al. (2016) recruited 126 students at the Social Science Research Lab at the Faculty of Psychology at the University of Vienna; and Alm et al. (2017) recruited 170 subjects, of which 92 in the U.S. and 78 in Italy. Finally, concerns about the use of students as representatives of taxpayer behavior have been addressed in a stream of recent research showing that behavioral responses of students are largely the

[21] Among others, see Doerrenberg and Duncan (2014) which offer a defense of most concerns.

same as those of nonstudents in identical tax experiments (e.g. Doerrenberg and Duncan, 2014; Alm et al., 2015; Alm et al., 2017).

4.3. Directions for Future Research

Some unexpected findings and alternative interpretations of our results suggest directions for future research. For example, we were surprised to find that participants with no previous participation in experimental research showed a reduced propensity to become a tax evader; yet the magnitude of evasion was not affected by this variable. A second surprise was the finding that the size of our experimental groups had no impact on evasion—a result that contrasts with the group size effect identified in experiments on public goods (e.g. Isaac and Walker, 1988). Both of these findings deserve further investigation, to strengthen experimental research on taxation and refine researcher control over contextual factors.

Another promising avenue for future research would be to delve further into the implications of our experimental findings compared to those of Kleven et al.'s (2011) field experiment. Taken together, they suggest that even individuals from countries with vanishingly small rates of income tax evasion will evade to a great extent when given the opportunity. A valuable contribution to knowledge would lay in replicating such results in other countries with low tax evasion rates: conduct field and laboratory experiments providing opportunities to cheat, and see whether they are taken up as robustly as they were in Denmark. Other Scandinavian countries, such as Norway and Sweden, which have evasion rates of similarly small magnitude to Denmark's (Bäckman, 2008), would offer an obvious starting point.

An even more valuable contribution would be to apply such an experimental strategy to countries where evasion and self-employment (rather than third-party reporting of income) are more common, such as the US and Italy (Rubin, 2011). If our policy-centric hypothesis is correct, then increasing opportunities to cheat should work the same way in those environments as in Denmark. The lab experiment (coupled with a post-experimental questionnaire, as in this study) would establish a baseline of tax morale and compliance, while the field experiment would consist of interventions to verify income or increase perceived audit probability, as in Kleven et al. (2011). The objective would to see whether the field interventions could "move the needle" on compliance, morale, or both. In addition to advancing scholarly knowledge, this would further validate the case for policy-based interpretations of tax behavior, as opposed to "culturalist" accounts.

Generally, taxation research would benefit from continued cross-national comparative research designed to tease apart the sources of tax compliance. Our study represents an important first step in moving away from assumptions about attitude-behavior compliance. We hope this will provide a foundation for future studies of the social bases for fiscal support of government, clarifying how states sustain themselves and garner cooperation from the governed.

Funding

The research has received funding from the Danish Council for Independent Research, External grant no. 4003-00026B, “Mind the Gaps: Explaining Variation in Nordic Tax Collection and Compliance”, Principal Investigator Professor Brooke Harrington. The Danish Council for Independent Research has approved the experimental protocol in accordance with ethical standards for studies with human subjects. The experimental laboratories in Bologna and Copenhagen have approved the experimental methodology in accordance with the laboratory policies for recruitment and participations of human subjects. The funder and the laboratories have no role in study design, data collection and analysis, decision to publish and preparation of the manuscript.

Acknowledgements

The authors would like to thank Ronen Avraham, Oren Bar-Gill, Maria Bigoni, Emanuela Carbonara, Marco Casari, Manuele Citi, John D'Attoma, Diego Gambetta, Orsola Garofalo, Agne Kajackaite, Tamar Kricheli-Katz, Jacob Hasselbalch, Mart Laatsit, Henrik Lando, Shay Lavie, Tore Nilssen, Sanne Maria Noyon, Tone Ognedal, Enrico Santarelli, Herman Mark Schwartz, Alexander Christopher Sebald, Sven Steinmo, Francesco Parisi, Lene Holm Pedersen, Ariel Porat, Avraham Tabbach, and Nan Zhang for helpful comments and suggestions. The authors are particularly grateful to John D'Attoma, Sven Steinmo, Nan Zhang and their research team for sharing their z-tree code and providing their clerical task (Andrighetto et al., 2016; Zhang et al., 2016). The authors would also like to thank the seminar audiences at the Buchmann Faculty of Law at Tel Aviv University (November 2017); the Centre for the Study of Equality, Social Organization and Performance of the University of Oslo (October 2017); the Department of Business and Politics of Copenhagen Business School, Denmark (Spring 2017); the SKAT Danish Tax Authority, Denmark (September 2017); the 34th Annual Conference of the European Association of Law and Economics held in London, UK (September 2017); the XXIX Annual Conference of the Italian Society of Public Economics held in Catania, Italy (September 2017). The authors are grateful to the laboratory coordinators Alexander Christopher Sebald for the sessions in Copenhagen and Marco Casari for the sessions in Bologna, as well as the laboratory staff members Andreas Gotfredsen, Valeria Minervini and Hauke Roggenkamp for the sessions in Copenhagen and Lorenzo Golinelli for the sessions in Bologna. We would also like to thank Jens Olav Dahlgaard, Stine Haakonsson, Sine Nørholm Just, Andreas Møller Mulvad and Giulio Zichella for their valuable help in recruiting Danish participants. This article was drafted when Alice Guerra was a Postdoc Research Fellow at Copenhagen Business School financed by the Danish Council for Independent Research (grant No 4003-00026B) and it was concluded during her Marie-Curie fellowship at Copenhagen Business School (grant agreement No 748062). The financial support of these two funding bodies is gratefully acknowledged.

References

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

Alm J. and Torgler B. (2006). Culture Differences and Tax Morale in the United States and in Europe. *Journal of Economic Psychology*; 27(2):224–246.

Alm, J. (2010). Testing behavioral public economics theories in the laboratory. *National Tax Journal* 63 (4, Part 1): 635-658.

Alm, J. (2012). Measuring, explaining, and controlling tax evasion: Lessons from theory, experiments, and field studies. International Tax and Public Finance, 19(1), 54–77. https://doi.org/10.1007/s10797-011-9171-2

Alm, J., and Jacobson, S. (2007). Using laboratory experiments in public economics. *National Tax Journal* 60 (1): 129-152.

Alm, J., Bernasconi, M., Laury, S., Lee, D. J., and Wallace, S. (2017). Culture, compliance and confidentiality: A study of taxpayer behavior in the United States and Italy. *Journal of Economic Behavior & Organization*, 1–37.

Alm, J., Bloomquist, K. M., and McKee, M. (2015). On the external validity of laboratory tax compliance experiments. *Economic Inquiry*, 53(2), 1170-1186.

Andrighetto, G., Zhang, N., Ottone, S., Ponzano, F., D'Attoma, J., and Steinmo, S. (2016). Are Some Countries More Honest than Others? Evidence from a Tax Compliance Experiment in Sweden and Italy. *Frontiers in Psychology*, 7.

Bäckman, J-E. (2008) Tax gap map for Sweden: How was it created and how can it be used. *Swedish National Tax Agency Report: 1B.*

Bandura, A. (2002) Selective moral disengagement in the exercise of moral agency. *Journal of Moral Education* 31: 101-119.

Barone, G. and Mocetti, S. (2009) Tax morale and public spending inefficiency. Temi di discussion (Economic working papers) 732, Bank of Italy, Economic Research and International Relations Area.

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

Bühren, C., and Kundt, T. C. (2014). Does the Level of Work Effort Influence Tax Evasion? Experimental Evidence. *Review of Economics*, 65(2), 137-158.

Bühren, C., Pleßner, M., (2014). The trophy effect. *Journal of Behavioral Decision Making*, 27 (4), 363–377.

Carpenter, J., Huet-Vaughn, E., 2017. Real Effort Tasks, forthcoming in: Schram, A., Ule, A. (Eds.), *Handbook of Research Methods and Applications in Experimental Economics*. Edward Elgar, Cheltenham, UK.

Casal, S. and Mittone, L. (2016) Social esteem versus social stigma: The role of anonymity in an income reporting game. *Journal of Economic Behavior & Organization*, 124: 55-66.

Choo, C. L., Fonseca, M. A., and Myles, G. D. (2016). Do students behave like real taxpayers in the lab? Evidence from a real effort tax compliance experiment. *Journal of Economic Behavior & Organization*, 124, 102-114.

Coricelli, G., Joffily, M., Montmarquette, C., & Villeval, M. C. (2010). Cheating, emotions, and rationality: an experiment on tax evasion. *Experimental Economics*, 13(2), 226-247.

Cragg, J. G. 1971. Some statistical models for limited dependent variables with application to the demand for durable goods. *Econometrica* 39: 829-844.

Cummings R, 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(3):447–457.

Deutscher, I. (1966) Words and deeds: Social science and social policy. *Social Problems*, 13:235-65.

Doerrenberg, P., and Duncan, D. (2014). Experimental evidence on the relationship between tax evasion opportunities and labor supply. *European Economic Review*, 68, 48-70.

Eckel, C. C., and Grossman, P. J. (2008). Forecasting risk attitudes: An experimental study using actual and forecast gamble choices. *Journal of Economic Behavior & Organization*, 68(1), 1-17.

Engel, C., and P. G. Moffatt. 2014. dhreg, xtdhreg, and bootdhreg: Commands to implement double-hurdle regression. *Stata Journal* 14: 778-797.

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

Floris, G. (2010). *Separati in patria: Nord contro Sud, perché l'Italia è sempre più divisa*. Milan: Rizzoli.

Greiner B. (2004). The Online Recruitment System ORSEE 2.0.—A Guide for the Organization of Experiments in Economics. University of Cologne, *Working Paper Series in Economics*. 2004; 10(23):63–104.

Greiner, B. (2015). Subject pool recruitment procedures: organizing experiments with ORSEE. *Journal of the Economic Science Association*, 1(1), 114-125.

Halla, M. (2012). Tax morale and compliance behavior: First evidence on a causal link. *The BE Journal of Economic Analysis & Policy*, 12(1).

Holt, C. A., and Laury, S. K. (2002). Risk aversion and incentive effects. *American Economic Review*, 92(5), 1644-1655.

Isaac, R. M., and Walker, J. M. (1988). Group size effects in public goods provision: The voluntary contributions mechanism. *The Quarterly Journal of Economics*, 103(1), 179-199.

Kleven, H. J., Knudsen, M. B., Kreiner, C. T., Pedersen, S., & Saez, E. (2011). Unwilling or unable to cheat? Evidence from a tax audit experiment in Denmark. *Econometrica*, 79(3), 651-692.

Kogler, C., Mittone, L., & Kirchler, E. (2016). Delayed feedback on tax audits affects compliance and fairness perceptions. *Journal of Economic Behavior & Organization*, 124, 81-87.

Kugler, T., Connolly, T., and Ordóñez, L. D. (2012). Emotion, decision, and risk: Betting on gambles versus betting on people. *Journal of Behavioral Decision Making*, 25(2), 123-134.

Kutner, B., Wilkins, C. and Yarrow, P. (1952) Verbal attitudes and overt behavior involving racial prejudice. *Journal of Abnormal and Social Psychology*, 47: 649-52.

LaPiere, R. (1934) Attitudes vs. actions. *Social Forces*, 13:230-237.

Liska, A. (1974) Emergent issues in the attitude-behavior consistency controversy. *American Sociological Review*, 39: 261-272.

Rabin, M. (2000). Risk aversion and expected-utility theory: A calibration theorem. *Econometrica*, 68(5), 1281-1292.

Riahi-Belkaoui, A. (2004). Relationship between tax compliance internationally and selected determinants of tax morale. *Journal of international accounting, auditing and taxation*, 13(2), 135-143.

Rothstein, B., and Uslaner, E. M. (2005). All for all: Equality, corruption, and social trust. *World Politics*, 58(1), 41-72.

Rubin, M. (2011). The practicality of a top down approach to estimating the direct tax gap. Paper presented *New Perspectives on Tax Administration: An IRS-TPC Research Conference*, Washington, June 22.

Torgler, B. (2005) Tax morale in Latin America. *Public Choice*, 122: 133-157.

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

Torgler, B., Schaffner, M., and Macintyre, A. (2007). Tax compliance, tax morale, and governance quality. *International Studies Program Working Paper*, Andrew Young School of Policy Studies, Georgia State University.

White Jr, R. (2002) Do employees act like they think? Exploring the dichotomy between moral judgment and ethical behavior. *Public Administration Quarterly*, 25: 391-412.

Yitzhaki, S. (1974). Income tax evasion: A theoretical analysis. *Journal of Public Economics*, 3(2), 201-202.

Zhang, N., Andrighetto, G., Ottone, S., Ponzano, F., & Steinmo, S. (2016). " Willing to Pay?" Tax Compliance in Britain and Italy: An Experimental Analysis. *PloS one*, 11(2).

Appendix: Tables and Figures

Table 1: Experimental Design

Stage	Round	Treatments		
		Tax Rate	**Audit Probability**	**Tax Redistribution**
1	1	30 %	5 %	**No**
	2	30 %	5 %	**Redistribution**
	3	30 %	5 %	**Redistribution x 2**
2	4	**5 %**	5 %	Redistribution
	5	**30 %**	5 %	Redistribution
	6	**50 %**	5 %	Redistribution
3	7	30 %	**5 %**	Redistribution
	8	30 %	**30 %**	Redistribution
	9	30 %	**50 %**	Redistribution

Table 2: Risk Attitude Test (adapted from Holt and Laury, 2002)

Decision	Option A			Option B			EV(B) - EV(A)
	Pr. of winning 20 points	**Pr. of winning 16 points**	**EV(A) (σ^2)**	**Pr. of winning 38.5 points**	**Pr. of winning 1 points**	**EV(B) (σ^2)**	
1	10%	90%	16.4 (1.44)	10%	90%	4.75 (126.6)	-11.65
2	20%	80%	16.8 (2.56)	20%	80%	8.5 (225)	-8.3
3	30%	70%	17.2 (3.36)	30%	70%	12.25 (295.3)	-4.95
4	40%	60%	17.6 (3.84)	40%	60%	16 (337.5)	-1.6
5	50%	50%	18 (4)	50%	50%	19.75 (351.6)	1.75
6	60%	40%	18.4 (3.84)	60%	40%	23.5 (337.5)	5.1
7	70%	30%	18.8 (3.36)	70%	30%	27.25 (295.3)	8.45
8	80%	20%	19.2 (2.56)	80%	20%	31 (225)	11.8
9	90%	10%	19.6 (1.449)	90%	10%	34.75 (126.6)	15.15
10	100%	0%	20 (0)	100%	0%	38.5 (0)	18.5

Notes: Columns 4, 7, and 8 did not appear in the instructions handed out to participants. "Pr" is an abbreviation that stands for "Probability". "EV()" stands for Expected Value and "σ^2" is the variance of a specific lottery.

Table 3: Summary Statistics of Evasion Rates and Wilcoxon Rank Sum test, by Country and Round

Stage	**Round**	**Pooled (%)**	**Denmark (%)**	**Italy (%)**	**Difference (%)**	**Non-parametric test Wilcoxon Rank Sum test**	
						z	**p>|z|**
1	No Redistribution	58.85 (43.10)	71.65 (39.78)	49.92 (43.24)	21.73 (6.340)	3.435	0.0006
	Redistribution	46.24 (44.82)	58.35 (44.03)	37.79 (43.60)	20.55 (6.63)	3.157	0.0016
	Redistribution X2	25.21 (39.66)	28.68 (41.96)	22.78 (37.98)	5.89 (6.00)	0.628	0.5301
2	Tax Rate 5%	33.53 (44.24)	40.96 (48.07)	28.34 (40.79)	12.62 (6.65)	1.715	0.0864
	Tax Rate 30%	52.82 (44.05)	67.04 (42.21)	42.89 (42.74)	24.15 (6.44)	3.765	0.0002
	Tax Rate 50%	54.14 (44.60)	64.26 (42.37)	47.08 (44.95)	17.18 (6.65)	2.723	0.0065
3	Audit Prob 5%	50.73 (45.52)	64.29 (44.88)	41.26 (43.72)	23.02 (6.69)	3.434	0.0006
	Audit Prob 30%	19.24 (33.98)	23.61 (35.92)	16.20 (32.37)	7.41 (5.13)	1.489	0.1364
	Audit Prob 50%	8.783 (26.14)	5.454 (19.67)	11.11 (29.71)	-5.65 (3.94)	-1.035	0.3006
Average		38.83 *44.33*	47.14 1.78	33.04 1.36	14.10 (2.21)	6.245	0.000
Observations		*180*	*74*	*106*			

Notes: Mean and standard deviations in bracket of evasion rates in percentage by country and round. p>|z| reports the p-value of the non-parametric Wilcoxon rank-sum test for differences in distributions (H0: no differences) between the two subject pools. *Observations* indicates the number of observations in the subject pools.

Table 4: Determinants of Evasion

	Basic DH Models				DH Models with Interaction Effects			
	Model (1) Pooled DH		Model (2) Panel DH with random effects		Model (3) Pooled DH		Model (4) Panel DH with random effects	
Variable	**1st Hurdle**	**2nd Hurdle**	**1st Hurdle**	**2nd Hurdle**	**1st Hurdle**	**2nd Hurdle**	**1st Hurdle**	**2nd Hurdle**
Italy	-0.403*	-0.093*	-1.090	-0.082	-0.974**	0.139	-0.988	-0.200
	(0.211)	(0.052)	(0.832)	(0.167)	(0.430)	(0.097)	(0.960)	(0.288)
Tax Rate	0.838**	0.028		0.386***	0.819**	0.042		0.385***
	(0.341)	(0.090)		(0.107)	(0.343)	(0.090)		(0.107)
Audit	-1.557***	-0.227**		-0.958***	-1.582***	-0.239**		-0.958***
	(0.294)	(0.109)		(0.113)	(0.296)	(0.110)		(0.113)
Redistribution	-0.264***	-0.006		-0.111***	-0.265***	-0.007		-0.111***
	(0.082)	(0.020)		(0.026)	(0.083)	(0.020)		(0.026)
Male	0.519***	0.160***	0.801**	0.194***	0.476***	0.170***	0.786**	0.191***
	(0.079)	(0.021)	(0.319)	(0.063)	(0.081)	(0.021)	(0.315)	(0.063)
Age	-0.042***	-0.004	-0.076*	-0.003	-0.045***	-0.005*	-0.075*	-0.004
	(0.012)	(0.003)	(0.040)	(0.009)	(0.012)	(0.003)	(0.040)	(0.009)
Education	0.243***	0.023	0.681***	0.023	0.240***	0.028*	0.654**	0.030
	(0.061)	(0.015)	(0.253)	(0.050)	(0.062)	(0.015)	(0.255)	(0.048)
No Previous Participation	-0.259***	-0.008	-1.186***	0.143*	-0.272***	-0.013	-1.221***	0.156**
	(0.094)	(0.028)	(0.318)	(0.078)	(0.094)	(0.029)	(0.315)	(0.078)
Rows	-0.014	-0.001		-0.003	-0.011	-0.000		-0.002
	(0.023)	(0.005)		(0.011)	(0.024)	(0.005)		(0.011)
Risk Aversion	-0.073***	-0.012**	-0.196*	-0.022	-0.039	0.002	-0.188*	-0.022
	(0.025)	(0.006)	(0.102)	(0.020)	(0.038)	(0.008)	(0.102)	(0.027)
Perception	-0.543***	-0.058***		-0.245***	-0.657***	-0.038***		-0.249***
	(0.034)	(0.010)		(0.014)	(0.053)	(0.014)		(0.017)
Tolerance Of Tax Evasion	0.071***	0.003	0.061	0.007	0.016	0.003	0.146	-0.016
	(0.019)	(0.005)	(0.099)	(0.014)	(0.034)	(0.008)	(0.236)	(0.022)
Pool Size	0.007	0.003	0.035	0.003	0.001	0.003	0.041	0.002
	(0.015)	(0.004)	(0.058)	(0.011)	(0.015)	(0.004)	(0.059)	(0.011)
Tolerance Of Tax Evasion * Italy					0.082**	0.000	-0.079	0.036
					(0.042)	(0.011)	(0.252)	(0.031)
Perception * Italy					0.191***	-0.030*		0.006
					(0.065)	(0.018)		(0.022)
Risk Aversion * Italy					-0.052	-0.030**		0.001
					(0.050)	(0.012)		(0.037)
Constant	2.841***	0.963***	2.756**	1.263***	3.394***	0.856***	2.475*	1.344***
	(0.420)	(0.098)	(1.301)	(0.246)	(0.478)	(0.107)	(1.467)	(0.259)
σ_α		0.267***		0.309***		0.266***		0.307***
		(0.007)		(0.023)		(0.007)		(0.022)
σ_ε		0.267***		0.374***		0.266***		0.374***
		(0.007)		(0.011)		(0.007)		(0.011)
Transformed ρ				--0.163				-0.185
				(0.373)				(0.371)
χ^2 (Prob>χ^2)	673.873	163.688	42.657	880.409	685.946	173.701	46.7784	885.993
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
χ^2 overall (Prob>χ^2)	694.657		881.743		707.332		889.392	
	(0.000)		(0.000)		(0.000)		(0.000)	
Log likelihood	-808.128		-803.899		-796.001		-802.984	
Num. obs.	1620		1620		1620		1620	
Num. groups	NA		180		NA		180	

Standard errors are in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table 5: List of Variables

Variable	Description	Value Labels
Dependent Variable		
ER	Evasion Rate = 1 – reporting compliance rate, where the latter is the ratio of declared income to earned income.	[0,1]
Independent Variables		
Italy	Country of origin	1 if Italy, 0 if Denmark
Tax Rate	Tax rate	{0.05, 0.3, 0.5}
Audit	Probability of being audited	{0.05, 0.3, 0.5}
Redistribution	Redistribution of taxes paid by participants	{0, 1, 2}
Male	Gender	1 if male, 0 if female
Age	Age	positive integer
Education	Level of Education	discrete, [0,5]: 0 = junior high school 1 = high school 2 = Bachelor's Degree 3 = Master's Degree 4 = Master 5 = Ph.D.
No Previous Participation	No previous participation in experimental studies	1 if no previous participation, 0 otherwise
Rows	Number of rows correctly copied in each clerical task	discrete
Risk Aversion	Individual degree of risk aversion, measured as the total number of "Option A" choices in the post-experimental risk elicitation task.	discrete, [0,10]
Perception	Individual belief about others' compliance choices, measured in each round via the following question: "In your opinion, how many participants in this room declared their true earnings for tax purposes?"	discrete, [0,5]: 0=none; 1=less than a half; 2=about a half; 3=more than a half; 4=almost everyone; 5=everyone
Tolerance Of Tax Evasion	Self-reported attitude towards tax evasion, measured in the post-experimental questionnaire via the following question: "Can cheating on tax if you have the chance always be justified, never be justified, or something in between?"	discrete, [0 (never be justified),10 (always be justified)]
Pool Size	Number of participants per session	discrete

Figure 1: Distribution of Evasion Rates, by Country and Rounds

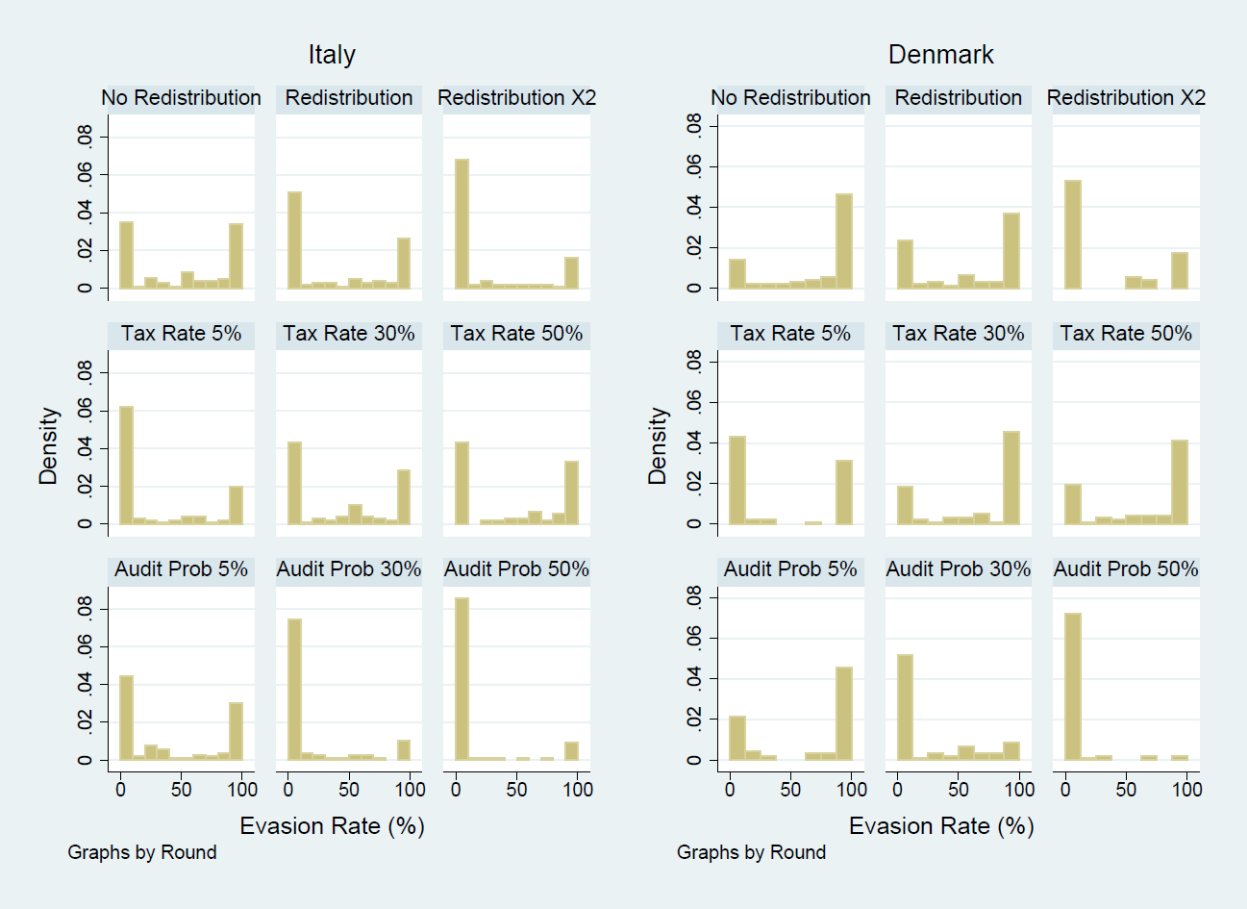

Figure 2: Cumulative Distributions of Evasion Rate, by Country

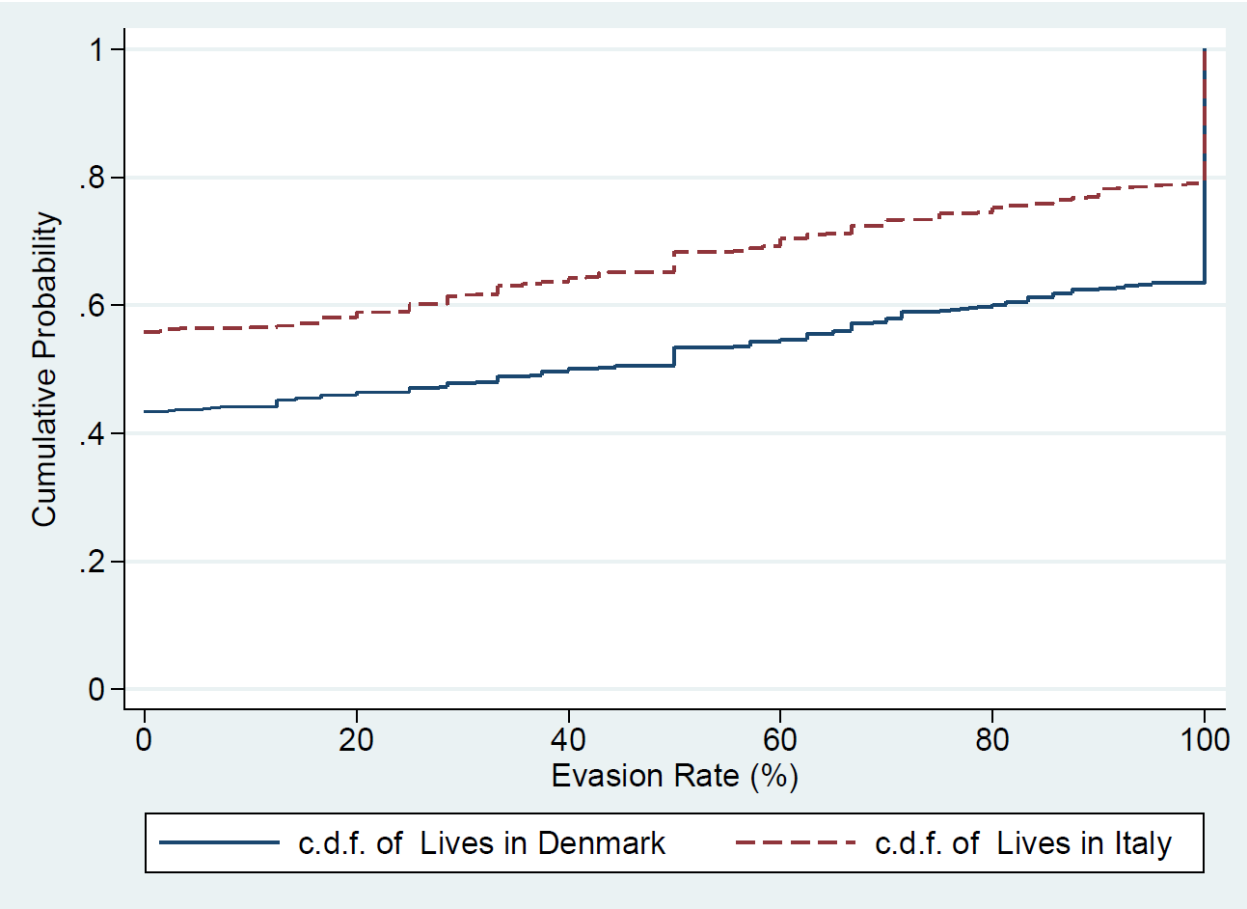

Figure 3: Tolerance of Tax Evasion, by Country