RESEARCH ARTICLE

"Willing to Pay?" Tax Compliance in Britain and Italy: An Experimental Analysis

Nan Zhang[1]*, Giulia Andrighetto[1,2], Stefania Ottone[3], Ferruccio Ponzano[4], Sven Steinmo[5,6]

1 Department of Social and Political Sciences, European University Institute, San Domenico di Fiesole (FI), Italy, 2 Institute of Cognitive Science and Technology, National Research Council, Rome, Italy, 3 Department of Economics, Management and Statistics, University of Milano-Bicocca, Milano, Italy, 4 Department of Law, Politics, Economics and Social Sciences, University of Eastern Piedmont, Alessandria, Italy, 5 Robert Schumann Center for Advanced Studies, European University Institute, San Domenico di Fiesole (FI), Italy, 6 Department of Political Science, University of Colorado, Boulder, United States of America

* nan.zhang@eui.eu

Abstract

As shown by the recent crisis, tax evasion poses a significant problem for countries such as Greece, Spain and Italy. While these societies certainly possess weaker fiscal institutions as compared to other EU members, might broader cultural differences between northern and southern Europe also help to explain citizens' (un)willingness to pay their taxes? To address this question, we conduct laboratory experiments in the UK and Italy, two countries which straddle this North-South divide. Our design allows us to examine citizens' willingness to contribute to public goods via taxes while holding institutions constant. We report a surprising result: when faced with identical tax institutions, redistribution rules and audit probabilities, Italian participants are significantly more likely to comply than Britons. Overall, our findings cast doubt upon "culturalist" arguments that would attribute cross-country differences in tax compliance to the lack of morality amongst southern European taxpayers.

Introduction

Modern welfare states face a set of difficult challenges as they adapt to the demographic, economic and political strains of the early 21st century. States must struggle to maintain adequate support for social welfare and educational programs in the face of growing distrust of bureaucratic institutions, intense pressures to cut taxes for politically powerful constituencies, and fiscal burdens arising from an aging population. The ability of governments to collect revenues in an efficient and cost-effective manner is of central importance to how successfully states meet their policy goals. And to ensure a healthy fiscal foundation, states must be able to control (or reduce) tax evasion on the part of their citizens.

Yet, while Western European states generally possess tax systems sharing many of the same formal features [1], actual rates of tax compliance vary widely across these societies [2–7]. Moreover, evasion rates also seem to follow a geographic pattern, with high levels of compliance in northern Europe, and widespread under-reporting in the countries further south.

Using the size of the "shadow economy" as a proxy for tax evasion, Schneider and Enste find the lowest compliance rates in Western Europe in Portugal, Spain, Italy and Greece [4].

The literature has advanced several "institutionalist" theories to account for this cross-national variation in tax compliance [8]. In large part, these explanations focus on the relationship between the quality of government and citizens' willingness to comply with fiscal demands. Specifically, this literature argues that citizens are more likely to pay their taxes if they believe that the government is spending their money honestly and efficiently [3, 7, 9–17]. By contrast, when citizens perceive public institutions as corrupt and wasteful, they are likely to reciprocate by being dishonest in turn [18, 19]. Thus, one explanation for EU-wide differences in levels of tax compliance is that, in southern European countries, people often interact with low quality institutions for which they are (unsurprisingly) unwilling to pay.

A second set of theories links tax compliance to broader cultural norms and values. Within Europe, one important axis of cultural variation concerns how different societies draw the boundaries of moral behavior [18, 20–24]. Specifically, southern European societies are typically characterized as more "familistic" or "collectivist," and ethical conduct is often assumed to apply to only a small circle of familial or personal relationships, while outside of this circumscribed network, selfish or opportunistic behavior is norm. In the public sphere, individuals follow the rules not out of some internalized sense of "right" and "wrong," but only when they are coerced to do so [24]. By contrast, northern European societies are often said to emphasize values of "autonomy" or "individualism," and citizens are presumed to apply the same ethical principles that prevail within familial relations to conduct in the civic realm.

This distinction between what Tabellini has termed "limited" and "generalized" morality has direct implications for the level of tax compliance within a society [24–26]. As many scholars have noted, audit rates and punishment probabilities are insufficient to deter cheating in most cases [10, 27–30], and states must therefore rely upon voluntary compliance to collect fiscal dues. However, if "cheating the system" imposes little moral cost, then the willingness to pay is undermined. Thus, a broader cultural argument would lead us to expect greater tax evasion in southern European countries, *independent of institutional performance*.

Our research attempts to test this hypothesis using cross-cultural behavioral experiments. The advantage of this approach is that it allows us to hold formal institutions (e.g. tax rates, audit probabilities, the efficiency of the state, etc.) constant across countries, and thereby isolate the influence of broader cultural factors on fiscal behavior [9, 31, 32]. In this paper, we report results from two countries—the UK and Italy—which we take as "representatives" of northern and southern Europe. Slemrod estimates the evasion rate in the UK to be around 8% or 9% of GDP [33], while comparative figures for Italy can reach as high as 25% to 30% [34]. Culturally, Italy is often vilified—both in the press as well as in popular opinion—as the quintessential "amoral" society in which people cannot be trusted to behave ethically outside the network of familial and personal relations [20]. By contrast, Britain is rather typical of Protestant northern European societies in terms of cross-national rankings of "autonomy" and "individualism" [21, 22].

Our tax experiment involves over 500 participants across multiple locations in Italy and the UK. The main experimental task consists of a tax compliance scenario in which participants earning real money are asked to report their income under a variety of tax rates and redistributive rules. By comparing income declarations across countries, we are able to investigate whether, independently of the institutions, Italians are indeed less compliant than Britons when faced with identical fiscal choices.

To preview our main results, we find little evidence to support the contention that the morality of tax compliance is weaker in Italy, compared to the UK. Instead, when we average across all of the institutional scenarios in our experiment, we find that the average compliance rate amongst Italian participants is significantly *higher* than amongst the British. As we

describe in more detail below, the size of the British-Italian gap in tax compliance varies between each of the different scenarios. However, we emphasize that in every scenario of the experiment, Italian participants report a higher percentage of their income, compared to the British. These results remain robust to the inclusion of a host of demographic controls, and are reproduced in multiple experimental locations in the two countries.

In summary, although stereotypes about the "amorality" of Italian (and, more generally, southern European) taxpayers abound in the popular consciousness, our results suggest that cultural values cannot explain the significant cross-national variation in evasion rates that we observe in the real world. In the concluding section, we discuss several implications of our findings for future work.

Methods

Overview

Our experiments were conducted at six universities across the United Kingdom and Italy at various points during the academic year 2013–2014. To respect the anonymity of participants, we do not report the experimental locations in this paper, although details are available from the authors upon request. Our team spent over a year designing and re-designing our experimental protocols to ensure the consistency of the laboratory set-ups and selection pools in each of these locations. Everything from the recruitment methods, to the way participants entered the lab, to the final payments procedure was the same in each session. Our experimental instructions (both oral and on-screen) were also translated and back-translated between Italian and English by different native speakers to ensure consistency in meaning.

Participants were recruited to the experiment using ORSEE [35]. More specifically, each of the universities in which we conducted our study maintains an electronic database of individuals who had expressed interest in participating in behavioral experiments. These participant pools are composed mainly of undergraduate students, but also include a number of non-students and people who had already graduated. Several days prior to the actual sessions, individuals in the database receive an email informing them of the opportunity to take part in an upcoming research project (for a reasonable hourly wage). The email also contained a link where individuals could sign up for one session of their choice from a list of scheduled sessions.

On the day of the experiment, upon arrival in the laboratory, participants were given a randomly-drawn, anonymized ID number and assigned to a corresponding personal computer terminal. Only participants and the researchers were present for each session. The experiments were conducted using zTree [36], participants undertook all experimental tasks via computer, and the terminals were partitioned to ensure that participants could not communicate during the session, nor observe others in the room. Also, to ensure anonymity, we announced that decisions and payments would be linked only to participants' ID-numbers, and not to individual names.

Once all participants were seated at their individual terminals, we began the session by reading a short introductory script. Participants were informed that they would be asked to complete a number of tasks (which we would gradually describe to them) and make a number of choices. Based on their choices and the choices of the other participants, they would earn experimental currency units (ECUs), which would be converted into local currencies (pounds and euros) at the end of the session. The exchange rate was set so that participants would earn approximately twice the average hourly wage for student employment in the local context. Individuals then took part in the tax compliance experiment, which we describe in the following section. Finally, we collected demographic and attitudinal information via an online survey, which was linked to decisions in the experiment using the anonymized ID numbers. In all,

each session of the experiment lasted about 90 minutes, and participants earned an average of 16 euros or 14.5 pounds for their time. All payments were distributed anonymously in cash.

Importantly, prior to beginning the experimental tasks, participants were given no information about the aims of the research project, nor about the decisions we sought to elicit. Also, at no point whatsoever were participants told that they are taking part in a larger study comparing decision-making across national groups. Finally, when moderating each session, we made sure to employ only native speakers (and, in the Italian case, speakers with the "correct" regional accent). These procedures were implemented to ensure that participants would not be subject to national (or group-level) reputational concerns when making their decisions.

Ethics Statement

Our experiments have been approved by the IRB Committee at the University of Colorado, Boulder, where the principal investigator holds a professorship. Our project has also been approved by the Ethics Council of the European Research Council, and the the European University Institute Ethics committee. Finally, our work has been authorized by all of the Italian and UK laboratories we have used, but we did not undergo a separate university-based IRB review in these cases. All participants signed a written consent form prior to taking part in the study. Participants who wish to leave the experiment early could elect to receive a 5 euro / pound show-up fee. In practice, no individuals refused to participate or dropped out.

Participants

The data we present in this paper are drawn from 31 different experimental sessions involving a total of 671 participants from a variety of academic disciplines. Because we are interested in comparing specifically British and Italian participants, we retain the data for only native students, whom we define as those individuals born in Britain (Italy) to British (Italian) parents. Because of a misunderstanding at the recruitment stage, one session in [location anonymized] enrolled many immigrants. While we excluded the immigrants' data, we were also concerned that interacting with a disproportionate number of foreigners may have skewed the behavior of even native-born participants. We therefore drop this session entirely from the analysis. The result leaves us with a subset of 531 participants, of which 281 (52.9%) are from Italy and 250 (47.1%) are from the UK (including the entire sample of 671 participants does not substantively change our results). S1 Table presents descriptive statistics for our participant pools. Overall, 56% of our participants are male, with an average age of 23.8 years (s.d. = 7.7 years).

We note that there are several slight demographic differences between our British and Italian samples. In particular, British participants are significantly more likely to be employed and to report a higher willingness to take risks, while Italians are more likely to study economics. In addition, British participants tend to be more skilled at our real effort task (described below), and hence earn more ECUs in the experiment. On the other hand, we detect no significant differences between the two populations in terms of gender, age, or previous participation in experiments. We control for all of these demographic characteristics in our analyses below.

Tax Compliance Experiment
Design and Procedure

The tax compliance experiment proceeds as follows. Participants begin by completing a clerical task, in which they must copy rows of information from a sheet of paper into the computer. For each correctly copied row, participants earn 10 ECUs. Next, participants are asked to declare this income for taxation purposes under different scenarios (an example of the income

reporting screen is shown in S1 Fig). In the terminology of the experiment, each scenario constitutes a "round." Participants are informed that they are free to declare any amount of their income—from 0% to 100%—in each round, with the knowledge that they would pay taxes on only the reported portion of their incomes.

In all, the experiment is composed of nine separate income declaration rounds. Participants have no prior knowledge of the total number of rounds. In each round, we specify different rules for the taxation and redistribution of declared incomes (which we describe shortly). In doing so, we wanted to explore the effects of different causal variables from the tax policy literature. As a consequence, readers will note that our design is similar to previous tax compliance experiments that have also investigated the influence of these variables [27, 37–43]. Importantly, our primary contribution examines how, holding these institutional features constant, tax compliance varies across countries [9, 31].

In Rounds 1 through 3, we implement a flat 30% tax rate, and vary how tax revenues are redistributed to participants, thus simulating behavior under different levels of "efficiency" in providing public goods. In Round 1, there is no redistribution, and all tax revenues are "wasted." In Round 2, all taxes collected are redistributed on an equal basis to all participants. Finally, in Round 3, all taxes collected are first *doubled* and then redistributed equally to all participants.

By contrast, in Rounds 4 through 6, we hold redistribution constant and instead vary the tax rate. We use a flat tax rate of 10% in Round 4, of 30% in Round 5, and of 50% in Round 6. In each of these rounds, the revenues collected are doubled and redistributed equally to all participants.

Rounds 7 through 9 are more eclectic. In Rounds 7 and 8, we introduce two different progressive taxation schemes. In the first scheme, participants falling within the top 10% of declared incomes pay a 50% tax rate, participants in the bottom 10% of declared incomes pay a 10% tax rate, and everyone else pays a 30% rate. However, participants do not know exactly where they themselves fall in the overall distribution of declared incomes. In the second progressive scheme, all income over 100 ECU is taxed at a 50% rate, income between 50 and 100 ECU is taxed at a 30% rate, and all income below 50 ECU is taxed at a 10% rate. Again, under these two schemes, all taxes collected are doubled and redistributed on an equal basis.

Finally, in Round 9, we donate all tax revenues (collected under a flat 30% tax rate, and then doubled) to a real world charity, rather than redistributing revenues to the group. We selected Oxfam for UK participants, and the UNICEF for Italian participants. The order of experimental scenarios, as well as the rules in each round, are summarized in S2 Table.

In each round, participants were informed that they faced an (independent) 5% probability of being audited, in which case those who have under-reported their income must pay a fine equal to twice the amount of uncollected taxes. Importantly, we reveal the results of any audits only at the conclusion of the experiment. Also, at no point during the experiment do participants have information about whether other participants are audited, nor indeed, whether other participants are honestly declaring their own incomes. In fact, we continuously reminded participants at various points throughout the session that all decisions would be treated anonymously. These procedures were implemented to mitigate the influence of reciprocity, conditional cooperation, reputation or wealth effects.

We also took great care to ensure that the experiment would simulate, as much as possible within a laboratory setting, the private decision problem facing an individual taxpayer. Along these lines, we intentionally incorporate tax language in our protocols, using words such as "income," "taxes," and "audit" [9, 44]. While the issue of framing effects in tax experiments is far from settled [27, 45], we believe this design choice offers an improvement over the use of neutrally-framed compliance games in terms of the ability to stimulate taxpayer motivations.

In summary, by comparing how income is reported across our nine taxation and redistribution scenarios, we are able to investigate differences in tax compliance across a range of parameters. Furthermore, because other researchers have also employed similar experimental designs, we are able to use previous studies as an external check on the validity of our results.

Results and Discussion

Fig 1 displays the average percentage of earned income that is reported in each of the nine rounds, broken down between British and Italian participants. The vertical axis displays the *average tax compliance rate*, defined as the percentage of total earned income that is declared in each round. Several points stand out from the graph. First, comparing Rounds 1 through 3, we see that compliance responds positively to the efficiency of redistribution: in both countries, individuals are more willing to declare a larger percentage of their income when they know that tax revenues produce more public goods. Secondly, individuals respond to higher tax rates by evading their fiscal obligations: compliance falls as we move from Rounds 4 through 6. These results are in line with previous studies [27, 37–43], and provide us with some assurance about the validity of our experimental design.

Turning now to our main results, we document a surprising cross-national difference in compliance rates: on average, British participants reported a smaller share of their total income in every round as compared to Italians. As shown in S2 Fig, this finding is also fairly consistent across multiple experimental locations in each country.

While our visual inspection of the the average compliance rate already yields some interesting patterns, such a statistic also hides substantial nuance in participants' decision-making. Specifically, if we examine the distribution of compliance rates across *all* reporting decisions, we see that the data are not normal (see Fig 2). Rather, the average compliance rate actually aggregates three different outcomes:
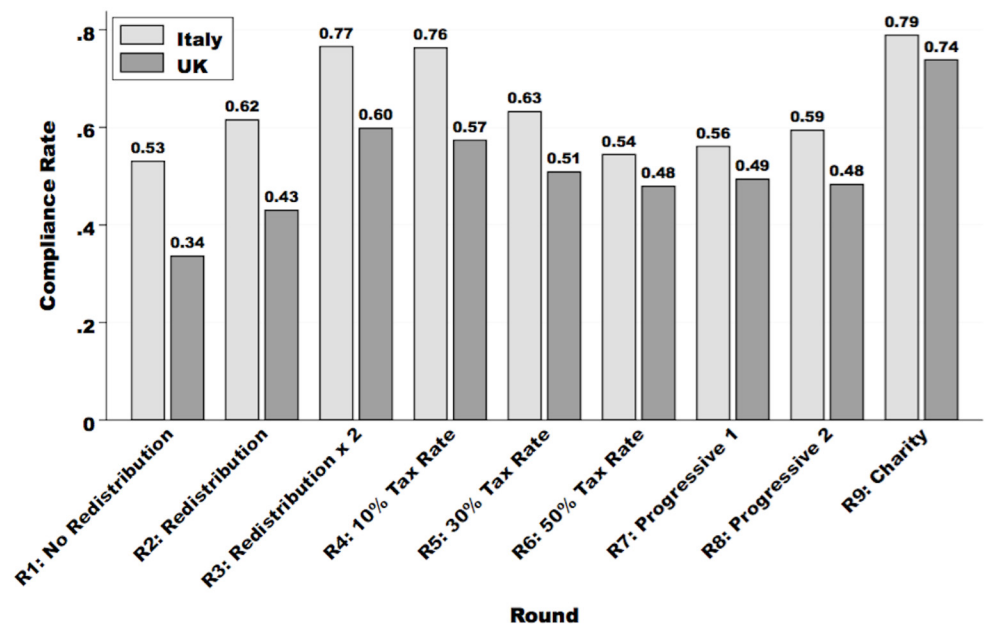

Fig 1. Average compliance rates, by country, in rounds 1 through 9. Bar heights represent the average percentage of earned income that is reported. The compliance rate is lower amongst British participants in every single round. This difference is statistically significant at the 5% level in rounds 1 through 5, and round 8. In addition, it is statistically significant at the 10% level in round 6 (results available from the authors, upon request).

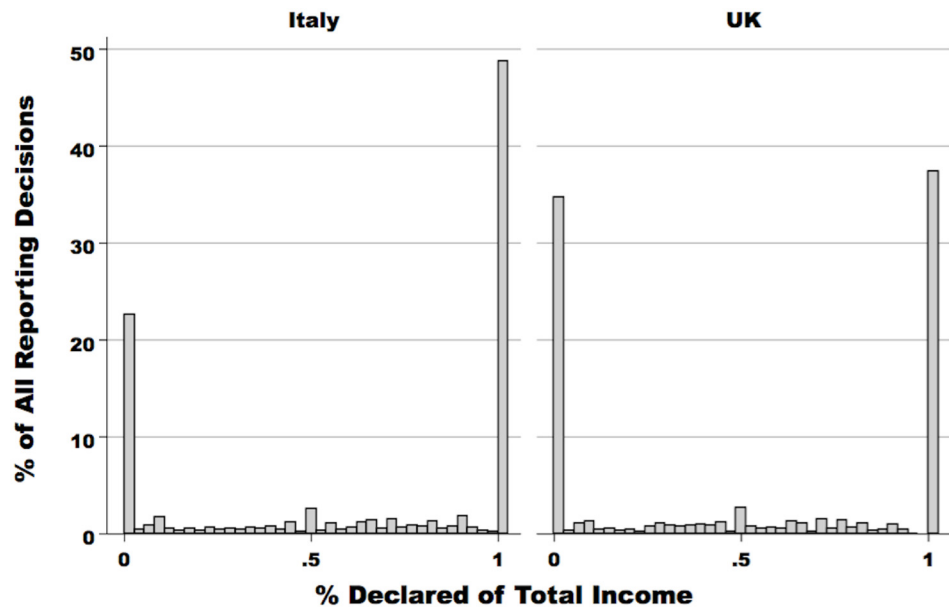doi:10.1371/journal.pone.0150277.g001

Fig 2. Distribution of compliance decisions, all rounds, by country. Bar heights represent the percentage of reporting decisions in which compliance falls within the ranges [0%–5%], [5%–10%]. . .[95%–100%]. The distribution is predominately bimodal: in 44.1% of all decisions, participants declare 100% of their earned income, while in 28.4% of all decisions, participants report that they earned 0 income.

doi:10.1371/journal.pone.0150277.g002

1. **Complete Compliance:** In 44.1% of all decisions, participants declare 100% of their earned income.

2. **Partial Evasion:** In 27.5% of all decisions, participants under-report their income to some degree.

3. **Complete Evasion:** In 28.4% of all decisions, participants report that they earned 0 income.

We wish to emphasize that we define these three categories of compliance at the *decision level*, and not at the *individual level*. We adopt this operationalization because individuals' choices may vary across the different rounds of the experiment (e.g. a participant could declare 0% in round 1, and 100% in round 2), and *a priori*, we had no strong priors about how to classify individuals. For example, if participant A reports 50% of his income in all rounds, while participant B reports 100% of her income in half of the rounds, and 0% in the other half, it was unclear to us that both A and B could be classified as "dishonest" to the same degree. To avoid making (arbitrary) decisions about how to define honesty at the individual level, we opt instead to define categories of compliance at the decision level.

The distributions in Fig 2 are similar to the patterns observed in other experimental studies of tax compliance [27, 46], as well as in "real world" tax declarations. In particular, Alm et al. examined a sample of 1,673 randomly-audited tax returns from the Internal Revenue Service (IRS), focusing solely on self-declared sole-proprietorship income [46]. This income source was selected because it generates no third-party reporting information (e.g. Form W-2 for wage income), and therefore mimics the declaration decision in the laboratory. These authors find that the compliance behavior of real taxpayers also conforms to the U-shaped patterns displayed in Fig 2. In other words, the similarity between our results and Alm et al.'s "taxpayer sample" suggests that our participants are treating the experimental decision-problem as a "real life" problem of tax compliance.

Table 1. Distribution of Complete Evasion, Complete Compliance and Partial Evasion Decisions.

	No. of Decisions: Complete Evasion			No. of Decisions: Complete Compliance			No. of Decisions: Partial Evasion			% of Income Declared by Partial Evaders:		
	(1) Italy	(2) UK	(3) Diff.	(4) Italy	(5) UK	(6) Diff.	(7) Italy	(8) UK	(9) Diff.	(10) Italy	(11) UK	(12) Diff.
R1: No Redistribuiton	0.33	0.50	-0.17*	0.33	0.19	0.14*	0.34	0.30	0.03	0.59	0.47	0.11*
R2: Redistribution	0.26	0.44	-0.18*	0.48	0.31	0.17*	0.26	0.25	0.01	0.52	0.48	0.03
R3: Redistribution x 2	0.13	0.30	-0.17*	0.66	0.48	0.18*	0.21	0.22	-0.01	0.51	0.55	-0.04
R4: 10% Tax Rate	0.16	0.32	-0.17*	0.64	0.48	0.17*	0.20	0.20	0.00	0.60	0.49	0.11
R5: 30% Tax Rate	0.23	0.34	-0.11*	0.47	0.37	0.10*	0.30	0.29	0.01	0.55	0.49	0.06
R6: 50% Tax Rate	0.31	0.38	-0.07	0.39	0.35	0.04	0.30	0.27	0.03	0.51	0.48	0.03
R7: Progressive 1	0.27	0.35	-0.08*	0.39	0.35	0.04	0.34	0.30	0.04	0.50	0.47	0.02
R8: Progressive 2	0.21	0.32	-0.11*	0.39	0.29	0.10*	0.40	0.39	0.01	0.51	0.49	0.02
R9: Charity	0.14	0.17	-0.03	0.69	0.63	0.06	0.17	0.02	-0.03	0.55	0.53	0.02

N Italy = 281; N UK = 250.

We employed Schlag's Z-test to test for country-level differences in columns (3), (6) and (9), and Mann-Whitney tests in column (12).

* indicates whether differences between countries are statistically significant at the 5% level.

Table 1 examines how this U-shaped pattern differs between the UK and Italy. Columns (1) through (3) of Table 1 display the proportion of decisions in each round characterized by complete evasion. We see that in almost all rounds (except 50% Tax Rate and Charity), a significantly greater percentage of decisions made by British participants result in a declaration of 0. The UK-Italy gaps are substantively large, ranging from about 8% of decisions in Round 7 to almost 18% of decisions in Round 2. In columns (4) through (6), we see the corresponding totals for the proportion of decisions in each round that are completely compliant. Here, the data tell a similar story: in the majority of rounds, significantly more Italians report their entire income. The cross-country gaps range from 10% to 18%. Finally, columns (7) though (9) examine the proportion of partial evasion decisions in each country. Here, however, we see that there are no statistically significant differences. In short, we see that British participants are (a) less likely to report 100% of their incomes, and (b) more likely to report exactly 0 income. By contrast, the proportion of decisions in which individuals engage in partial evasion is almost identical in the two countries.

However, in addition to considering the raw proportion of partial evasions decisions in each round, we must also examine the percentage of taxes declared in this subset. We consider this possibility in columns (10) through (12) of Table 1. Here, we analyze *only* the subset of partial evasion decisions and ask: *conditional* upon engaging in partial evasion, how much income is actually reported? Interestingly, here as well we detect almost no statistically significant differences between Italian and British participants (with the exception of the very first round). In other words, neither the proportion of partial evasion decisions, nor the percentage of income declared in these partial evasion decisions, is very different across our two countries. Rather, it appears that the Italy-UK compliance gap is driven almost entirely by differences in complete compliance and complete evasion.

To what extent can our findings be explained by demographic differences between our participant pools? Controlling for individual-level characteristics that may be correlated with the compliance rate, do Britons still declare less than Italians? To address these questions, we estimate the conditional effect of an Italy country dummy on the compliance rate, holding other

factors constant. Following [46], our basic specification is:

$$Y_{i,t} = \beta_0 + \beta_1 Italy_i + \beta_2 Income_{i,t} + \beta_3 X_i + \psi_t + u_i + \epsilon_{i,t}$$

where the dependent variable $Y_{i,t}$ denotes the percentage of income reported by participant i in round t; $Italy_i$ is a dummy variable for Italian participants; $Income_{i,t}$ is participant i's earned income (standardized) in round t; X_i denotes a vector of demographic variables including gender, age, employment status, economics training, previous participation in behavioral experiments, self-reported risk attitudes, and self-reported beliefs about the behavior of other participants; ψ_t is a set of $T-1$ dummies that capture round fixed effects; u_i are random effects that control for unobservable individual characteristics; and $\epsilon_{i,t}$ is an individual-round error term. A more detailed description of these variables is available in the supplementary materials. Following [46], we report results for a (participant) random effects generalized least squares estimation with standard errors corrected for clustering at the individual level. Results are presented in Table 2.

Table 2 shows that gender, economics training, past participation in experiments, risk attitudes, and beliefs about others' behavior are all correlated with compliance. Several results are worth emphasizing: first, unsurprisingly, participants who are more risk-taking also comply less. In addition, British participants self-rate as more risk-taking on average. Thus, the inclusion of the risk variable in the regression with only the $Italy_i$ dummy tends to shrink the cross-country compliance gap from 13% to 8% (results not shown). However, the coefficient on $Italy_i$ is still significant at the 1% level.

Table 2. Estimates of the Compliance Rate.

	(1)	(2)	(3)
Italy	0.13** (0.03)	0.12** (0.03)	0.10** (0.03)
Income (standardized)		-0.03** (0.01)	-0.02 (0.01)
Male			-0.17** (0.03)
Age (standardized)			0.01 (0.01)
Employed			-0.04 (0.03)
Economics Training			-0.09** (0.03)
Previous Participation			-0.09** (0.03)
Risk (standardized)			-0.06** (0.01)
Others Report: "Less"			-0.12** (0.04)
Others Report: "Much Less"			-0.26** (0.04)
Constant	0.52** (0.02)	0.37** (0.02)	0.74** (0.04)
Wald χ^2	20.94**	460.2**	871.8**
Number of Participants	531	531	512
Number of Decisions	4779	4779	4608
Round Fixed Effects	No	Yes	Yes

Panel estimations with participant random effects and clustered (participant level) standard errors. The dependent variable is the percentage of earned income that is declared for tax purposes by individual i in round t.

** indicates significance at the 1% level.

The number of observations drops slightly once we include demographic covariates in model (3). This is because the experimental tasks were implemented in zTree, while the demographic information was collected separately using Qualtrics survey software. This necessitated that participants enter their anonymous IDs twice: once into zTree, and once again into Qualtrics. Because some participants accidentally entered different IDs into the two systems, we were unable to match their experimental decisions with their demographic data.

doi:10.1371/journal.pone.0150277.t002

Second, individuals who believe that others are dishonest are themselves less likely to comply. This echoes other findings on the importance of social norms for tax compliance [47]. We also note that Italians are on average more pessimistic about the behavior of their counterparts. Thus, controlling for beliefs in the regression slightly increases the size of the cross-country gap from 13% to 14% (results not shown). We also tested whether the relationship between beliefs and compliance differs across our two countries by interacting beliefs with a country dummy (results not shown), but the coefficient on the interaction term was not statistically significant.

Third, in line with earlier studies [48–52], we find that men and economists are less compliant. Since there are slightly more economists amongst Italian participants, controlling for this variable also inflates the cross-country compliance gap from 13% to 15% (results not shown). Fourth, we note that individuals who have participated in experiments in the past report less of their income on average, compared to first-time participants. We believe this result reflects the fact that first-time participants may take the experimental tax frame more seriously, while individuals who participate in multiple experiments may treat the experience as more of a game, and may therefore report less of their income. Finally, we find some evidence that higher earners report less of their income. In model (2) with only the Italy dummy and round fixed effects, a one SD increase in income earned in the clerical task is associated with a 3% drop in the compliance rate (p-value = 0.009). With additional controls (model 3), the association remains negative (2% drop), but is no longer statistically significant at conventional levels (p-value = 0.071).

Most importantly, we find that even controlling for all of these characteristics, the Italian country dummy remains substantively strong and statistically significant. While a simple comparison of the average compliance rate across countries (model 1) shows that Italians report about 13% more of their total income than British participants, this estimate falls only slightly once we include round fixed effects and a control for earned incomes (model 2), as well as other individual level controls (model 3). We also note that our results remain substantively unchanged when the analysis is run using a series of tobit models (available upon request). Importantly, the core finding that Italian participants are systematically more compliant in their fiscal decision-making remains robust in our regression models. In model 3, with the full host of controls, the country gap remains at 10%.

To what extent can we interpret these results as evidence that Italians possess a stronger ethics of tax compliance? Are "culturalist" arguments about the (lack of) morality amongst southern European taxpayers simply wrong? Or might other factors that we have not controlled for account for our surprising results? Here, we discuss two potential confounds: trust in anonymity, and group-level reputational concerns.

First, although we emphasized the anonymity of decisions at multiple points throughout the experiment, the propensity of participants to trust our assurances may vary across countries. For instance, Italian participants may be more suspicious that their anonymity could somehow be compromised, and therefore act more compliantly than they would otherwise.

While we obviously cannot directly control what participants in different societies believe about lab experiments, we have worked to the best of our ability to guarantee to participants that their decisions would be treated anonymously. In particular, information about confidentiality was prominently displayed on the consent forms that participants signed prior to taking part in the experiments. On these forms, it was also stated that participants would be given anonymous ID numbers and randomly assigned to computer stations. Participants also knew that, after the experiment, they would be paid anonymously in sealed envelopes linked only to their ID number. The consent form itself was approved by institutional review boards in two countries, and our confidentiality procedures were scrupulously vetted. We strongly

believe that these procedures made it obvious to participants in both Italy and the UK that their anonymity would not be compromised.

A second potential confound relates to a concern we have already noted about maintaining a positive reputation. Specifically, participants in the lab know that they are being observed, and this feeling of being "under the microscope" (especially with respect to a sensitive topic such as tax evasion) may bring about more compliance than would otherwise prevail in the "real world." Furthermore, the size of this effect may differ between Italy and the UK. In particular, Italians may be more concerned about their collective reputation as a "high evasion" nation. They may therefore act more compliantly in the experiment in order to show their rejection of this stereotype. Importantly, this dynamic may still obtain *even though* (a) our moderators were all native speakers, and (b) participants were not informed that they would be compared to individuals in another country.

While we cannot rule out this alternative explanation for our results, we note that this dynamic is also present in other cross-national experimental studies on tax compliance [9, 31, 53], and does not seem to compromise the validity of earlier findings. Therefore, although we acknowledge that reputation concerns may be potential confounding variable, we believe that these concerns do not play a prominent role in accounting for our results.

Conclusions and Implications

Turning now to the original question animating our study, what light can our experiment shed on the current fiscal crises facing European countries? Clearly, many northern Europeans (and Americans) hold the view that the difficulties facing Greece, Italy, Spain and others are not simply the product of poor institutions or badly designed public policies, but are instead symptomatic of a far more fundamental problem—one of culture. In a nutshell, southern Europe's fiscal conundrum is said to stem from the fact that, in such "limited morality" societies, the "moral costs" of non-compliance are insufficient to ensure effective tax collection.

The problem, analytically, is that culture and institutions are difficult to disentangle and are almost certainly interdependent. For example, Italians may readily cheat on their taxes in "real life" precisely *because* they believe that the state—in corruptly and inefficiently using their tax revenue—is also cheating them. In fact, recent public opinion polls have shown that southern Europeans give their governments consistently low scores on control of corruption and the quality of service delivery [54, 55]. By contrast, northern Europeans' high willingness to pay may be a direct reflection of their belief that taxes go to support important public services which they value and personally consume.

In this article we have tried to analyze the specific question of whether southern Europeans (in this case, Italians) would behave differently than northern Europeans (in this case, the British) when faced with exactly the *same* institutions. We discovered, much to our own surprise, that British participants are more likely to under-declare their incomes in a tax / public goods experiment than Italians. While we do not pretend that our study has isolated all of the cultural variables that may influence tax compliance decision-making, at a minimum, our results cast doubt on the above-mentioned "culturalist" arguments.

We find these results encouraging. If it were indeed the case that culture dominates institutional structure, then the prospects for Europe—and for much of the developing world today—would be grim indeed. Instead, we show that when given the opportunity to contribute to and share in collective goods on an equal institutional playing field, Italians behave no worse than their British counterparts. Thus, we believe that a specific focus on institutional reforms and improving the "Quality of Government" [19] is likely to yield more significant results than the cultural blame game that too often seeps into the policy debate.

Supporting Information

S1 Fig. Example screenshot: Earnings reporting screen.
(PNG)

S2 Fig. Compliance rate across all locations.
(PNG)

S1 Table. Summary of participant characteristics: Italy and the UK.
(PDF)

S2 Table. Summary of tax reporting rounds.
(PDF)

S1 Text. English Language Experimental Instructions.
(PDF)

S2 Text. English Language Questionnaire.
(PDF)

S3 Text. IRB Approval Letters.
(PDF)

S4 Text. Detailed Description of Selected Covariates.
(PDF)

Acknowledgments

Author Contributions

Conceived and designed the experiments: GA SO FP SS. Performed the experiments: NZ GA SO FP SS. Analyzed the data: NZ. Wrote the paper: NZ.

References

1. Alm J, Torgler B. Culture Differences and Tax Morale in the United States and in Europe. Journal of Economic Psychology. 2006; 27(2):224–246. doi: 10.1016/j.joep.2005.09.002

2. Alm J, Martinez-Vazquez J, Schneider F. "Sizing" the Problem of the Hard-to-Tax. In: Alm J, Martinez-Vazquez J, Wallace S, editors. Taxing the Hard-to-Tax: Lessons from Theory and Practice. Amsterdam: Elsevier/North Holland; 2004. p. 11–75.

3. Edlund J. Trust in Government and Welfare Regimes: Attitudes to Redistribution and Financial Cheating in the USA and Norway. European Journal of Political Research. 1999; 35(3):341–370. doi: 10.1111/1475-6765.00452

4. Schneider F, Enste D. The Shadow Economy: An International Survey. New York, NY: Cambridge University Press; 2013.

5. Svallfors S. Worlds of Welfare and Attitudes to Redistribution: A Comparison of Eight Western Nations. European Sociological Review. 1997; 13(3):283–304. doi: 10.1093/oxfordjournals.esr.a018219

6. Taylor-Gooby P. Who Wants the Welfare State? Support for State Welfare Provision in European Countries. In: Kangas O, Svallfors S, editors. In the Eye of the Beholder: Opinions on Welfare and Justice in Comparative Perspective. Umeå: Bank of Sweden Tercentenary Foundation; 1995. p. 11–51.

7. Torgler B, Schneider F. What Shapes Attitudes Toward Paying Taxes? Evidence from Multicultural European Countries. Social Science Quarterly. 2007; 88(2):443–470. doi: 10.1111/j.1540-6237.2007.00466.x

8. Steinmo S. Political Institutions and Tax Policy in the United States, Sweden, and Britain. World Politics. 1989; 41(4):500–535. doi: 10.2307/2010528

9. Cummings R, Martinez-Vazquez J, McKee M, Torgler B. Tax Morale Affects Tax Compliance: Evidence from Surveys and an Artefactual Field Experiment. Journal of Economic Behavior & Organization. 2009; 70(3):447–457. doi: 10.1016/j.jebo.2008.02.010

10. Frey B, Feld L. Deterrence and Morale in Taxation: An Empirical Analysis. CESifo working paper Number 760; 2002.

11. Frey B, Torgler B. Tax Morale and Conditional Cooperation. Journal of Comparative Economics. 2007; 35(1):136–159. doi: 10.1016/j.jce.2006.10.006

12. Levi M. Of Rule and Revenue. Berkeley: University of California Press; 1989.

13. Levi M, Sacks A, Tyler T. Conceptualizing Legitimacy, Measuring Legitimating Beliefs. American Behavioral Scientist. 2009; 53(3):354–375. doi: 10.1177/0002764209338797

14. Pommerehne W, Hart A, Frey B. Tax Morale, Tax Evasion and the Choice of Policy Instruments in Different Political Systems. Public Finance. 1994; 49(Supplement):52–69.

15. Scholz J, Lubell M. Trust and Taxpaying: Testing the Heuristic Approach to Collective Action. American Journal of Political Science. 1998; 42(2):398–417. doi: 10.2307/2991764

16. Smith K. Reciprocity and Fairness: Positive Incentives for Tax Compliance. In: Slemroad J, editor. Why People Pay Taxes. Tax Compliance and Enforcement. Ann Arbor: University of Michigan Press; 1992. p. 223–250.

17. Smith K, Stalans L. Encouraging Tax Compliance With Positive Incentives: A Conceptual Framework and Research Directions. Law & Policy. 1991; 13(1):35–53. doi: 10.1111/j.1467-9930.1991.tb00056.x

18. Bergman M. Tax Evasion and the Rule of Law in Latin America: The Political Culture of Cheating and Compliance in Argentina and Chile. University Park, PA: Pennsylvania State University Press; 2009.

19. Rothstein B. The Quality of Government: Corruption, Social Trust, and Inequality in International Perspective. University of Chicago Press; 2011.

20. Banfield E. The Moral Basis of a Backward Society. Glencoe: Free Press; 1967.

21. Hofstede G. Culture's Consequences: Comparing Values, Behaviors, Institutions and Organizations across Nations. New York: Sage; 2001.

22. Inglehart R, Welzel C. Changing Mass Priorities: The Link between Modernization and Democracy. Perspectives on Politics. 2010; 8(2):551–567 doi: 10.1017/S1537592710001258

23. Platteau JP. Institutions, Social Norms, and Economic Development. Amsterdam: Psychology Press; 2000.

24. Tabellini G. Culture and Institutions: Economic Development in the Regions of Europe. Journal of the European Economic Association. 2010; 8(4):677–716. doi: 10.1111/j.1542-4774.2010.tb00537.x

25. Slemrod J. On Voluntary Compliance, Voluntary Taxes, and Social Capital. National Tax Journal. 1998; 51(3):485–491.

26. Putnam R. Making Democracy Work. Princeton: Princeton University Press; 1993.

27. Alm J, McClelland G, Schulze W. Why Do People Pay Taxes? Journal of Public Economics. 1992; 48(1):21–38.

28. Andreoni J, Erard B, Feinstein J. Tax Compliance. Journal of Economic Literature. 1998; 36(2):818–860.

29. Graetz M, Wilde L. The Economics of Tax Compliance: Fact and Fantasy. National Tax Journal. 1985; 38(3):355–363.

30. Torgler B. Speaking to Theorists and Searching for Facts: Tax Morale and Tax Compliance in Experiments. Journal of Economic Surveys. 2002; 16(5):657–683. doi: 10.1111/1467-6419.00185

31. Alm J, Sanchez I, De Juan A. Economic and Noneconomic Factors in Tax Compliance. Kyklos. 1995; 48(1):3–18. doi: 10.1111/j.1467-6435.1995.tb02312.x

32. Torgler B, Schaltegger C. Tax Amnesties and Political Participation. Public Finance Review. 2005; 33(3):403–431. doi: 10.1177/1091142105275438

33. Slemrod J. Cheating Ourselves: The Economics of Tax Evasion. The Journal of Economic Perspectives. 2007; 21(1):25–48. doi: 10.1257/jep.21.1.25

34. Santoro A. L'Evasione Fiscale. Bologna: Mulino; 2010.

35. Greiner B. 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.

36. Fischbacher U. z-Tree: Zurich Toolbox for Ready-made Economic Experiments. Experimental Economics. 2007; 10(2):171–178. doi: 10.1007/s10683-006-9159-4

37. Alm J, Jackson B, McKee M. Deterrence and Beyond: Toward a Kinder, Gentler IRS. In: Slemrod J, editor. Why People Pay Taxes: Tax Compliance and Enforcement. Ann Arbor, MI: The University of Michigan Press; 1992. pp. 311–329.

38. Alm J, Jackson B, McKee M. Estimating the Determinants of Taxpayer Compliance with Experimental Data. National Tax Journal. 1992; 45(1):107–114.

39. Alm J, Jackson B, McKee M. Institutional Uncertainty and Taxpayer Compliance. American Economic Review. 1992; 82(4):1018–1026.

41. Bernasconi M, Corazzini L, Seri R. Reference Dependent Preferences, Hedonic Adaptation and Tax Evasion: Does the Tax Burden Matter? Journal of Economic Psychology. 2014; 40:103–118. doi: 10.1016/j.joep.2013.01.005

40. Becker W, Büchner H, Sleeking S. The Impact of Public Transfer Expenditures on Tax Evasion: An Experimental Approach. Journal of Public Economics. 1987; 34(2):243–252. doi: 10.1016/0047-2727(87)90022-3

43. Friedland N, Maital S, Rutenberg A. A Simulation Study of Income Tax Evasion. Journal of Public Economics. 1978; 10(1):107–116. doi: 10.1016/0047-2727(78)90008-7

42. Bosco L, Mittone L. Tax Evasion and Moral Constraints: Some Experimental Evidence. Kyklos. 1997; 50(3):297–324. doi: 10.1111/1467-6435.00018

44. Calvet R, Alm J. Empathy, Sympathy and Tax Compliance. Journal of Economic Psychology. 2014; 40:62–82. doi: 10.1016/j.joep.2012.10.001

45. Wartick M, Madeo S, Vines C. Reward Dominance in Tax-Reporting Experiments: The Role of Context. Journal of the American Taxation Association. 1999; 21(1):20–31. doi: 10.2308/jata.1999.21.1.20

46. Alm J, Bloomquist K, McKee M. On the External Validity of Laboratory Tax Compliance Experiments Economic Inquiry. 2015; 53(2):1170–1186.

47. Alm J, McClelland G, Schulze W. Changing the Social Norm of Tax Compliance by Voting. Kyklos. 1999; 52(2):141–171. doi: 10.1111/j.1467-6435.1999.tb01440.x

48. Chung J, Trivedi V. The Effect of Friendly Persuasion and Gender on Tax Compliance Behavior. Journal of Business Ethics. 2003; 47(2):133–145. doi: 10.1023/A:1026004716676

49. Cullis J, Jones P, Lewis A. Tax Framing, Instrumentality and Individual Differences: Are There Two Different Cultures? Journal of Economic Psychology. 2006; 27(2):304–320. doi: 10.1016/j.joep.2005.07.003

50. Gërxhani K. Explaining Gender Differences in Tax Evasion: The Case of Tirana, Albania. Feminist Economics. 2007; 13(2):119–155. doi: 10.1080/13545700601184856

51. Kastlunger B, Dressler S, Kirchler E, Mittone L, Voracek M. Sex Differences in Tax Compliance: Differentiating between Demographic Sex, Gender-role Orientation, and Prenatal Masculinization (2D:4D) Journal of Economic Psychology. 2010; 31(4):542–552. doi: 10.1016/j.joep.2010.03.015

52. Torgler B, Valev N. Gender and Public Attitudes toward Corruption and Tax Evasion. Contemporary Economic Policy. 2010; 28(4):554–568. doi: 10.1111/j.1465-7287.2009.00188.x

53. Lewis A, Carrera S, Cullis J, Jones P. Individual, Cognitive and Cultural Differences in Tax Compliance: UK and Italy Compared. Journal of Economic Psychology. 2009; 30(3):431–445. doi: 10.1016/j.joep.2008.11.002

54. Charron N, Dijkstra L, Lapuente V. Regional Governance Matters: Quality of Government within European Union Member States. Regional Studies. 2013; 48(1):68–90. doi: 10.1080/00343404.2013.770141

55. Charron N, Dijkstra L, Lapuente V. Mapping the Regional Divide in Europe: A Measure for Assessing Quality of Government in 206 European Regions. Social Indicators Research. 2014:1–32.