



UNIVERSITÀ
di **VERONA**

Department
of **ECONOMICS**

Working Paper Series
Department of Economics
University of Verona

What Matters for Whistleblowing on Tax Evaders? Survey and Experimental Evidence

Armenak Antinyan, Luca Corazzini, Filippo Pavesi

WP Number: 7

December 2018

ISSN: 2036-2919 (paper), 2036-4679 (online)

What Matters for Whistleblowing on Tax Evaders? Survey and Experimental Evidence

Armenak Antinyan,¹ Luca Corazzini,² Filippo Pavesi³

Abstract. Whistleblowing is a powerful tool that the tax authorities of various countries use to curb tax evasion. Nonetheless, the determinants shaping one's positive attitude toward whistleblowing on tax evaders are rather understudied. We investigate the relationship between trust in the government and the attitude toward whistleblowing on tax evaders. We use data from two survey experiments conducted in Italy and the US, as well as from a unique national household survey administered in the Republic of Armenia. Our findings indicate that the level of trust in the government positively influences individuals' attitude toward whistleblowing, with this effect being robust across countries and data sources.

Keywords: Government Trust; Whistleblowing; Tax Evasion.

JEL Classifications: H26; G28.

¹ *Corresponding Author.* Wenlan School of Business, Zhongnan University of Economics and Law, Nanhu Avenue 182, Wuhan 430073, P.R. China. E-mail: antinyan.armenak@gmail.com, Tel: +86 027 88387186, Fax: +86 027 88387186.

² Department of Economics and Center for Experimental Research in Management and Economics (CERME), University of Venice "Ca' Foscari," Cannaregio, 821, 30121 Venezia (VE), Italy, and ISLA, Bocconi University, Milan, Italy. Email: luca.corazzini@unive.it.

³ Department of Economics, University of Verona, Via Cantarane, 24, 37129 Verona, Italy, and Stevens Institute of Technology, School of Business, Hoboken, NJ, USA. E-mail: filippo.pavesi@univr.it, Tel: +39 328 3255281.

1. Introduction

“Report someone to HM Revenue & Customs (HMRC) if they’re in trade or running a business but aren’t paying tax on their profits.”⁴ “Whistleblowers have helped the IRS detect and deter tax noncompliance and avoidance, helping to protect both the nation’s revenue collection and the integrity of our voluntary compliance tax system. Indeed, since 2007, information submitted by whistleblowers has assisted the IRS in collecting \$3.4 billion in revenue.”⁵

These statements depict whistleblowing as a crucial channel through which the tax authorities of various countries (in the abovementioned paragraph the tax authorities of the UK and the USA, respectively) intend to curb tax evasion. Indeed, detecting instances of tax fraud by obtaining information from witnesses or accomplices can be (sometimes) more efficient than extracting information through costly audits. Despite its paramount importance for policymakers, the determinants that shape citizens’ attitude toward whistleblowing on tax evaders are rather understudied in the literature. More specifically, to the best of our knowledge, no one has investigated the relationship between the attitude toward whistleblowing and the trust in formal institutions, such as the government and the tax authorities. As far as we are concerned, the personal and contextual correlates of whistleblowing in the context of tax evasion are not extensively studied either. The current paper aims at contributing to these relevant research questions.

We believe that the aim of the present paper is of utmost importance, since the taxation literature puts forth trust in the authorities as an important determinant of individual tax compliance. More specifically, it is illustrated that high trust can increase taxpayers’ intrinsic

⁴ <https://www.gov.uk/report-an-unregistered-trader-or-business> (retrieved on November 2, 2018).

⁵ 2016 Annual Report to the Congress of the Internal Revenue Service (https://www.irs.gov/pub/whistleblower/fy16_wo_annual_report_final.pdf, retrieved on November 2, 2018).

motivation to pay taxes (e.g., Feld & Frey, 2002; Torgler, 2003) and curb tax evasion (e.g., Kirchler, Hoelzl & Wahl, 2008; Muehlbacher, Kirchler & Schwarzenberger, 2011). Considering this evidence, one may conjecture that high trust can positively affect taxpayers' attitude toward whistleblowing as well as increase the frequency of actual reports. Thus, if the authorities of a country strive to establish institutional mechanisms to encourage whistleblowing, they should first invest effort in enhancing citizens' trust in public institutions. Otherwise, if the issue of trust is neglected, even the best legal regimes (e.g., whistleblower protection acts) or incentivizing mechanisms (e.g., financial rewards) may fail to encourage individuals to report fraudulent behavior in the presence of low trust in institutions.

To investigate the causal relationship between the trust in the government and the attitude toward whistleblowing, we administer two survey-experiments in Italy and the US. More specifically, we elicit respondents' attitude toward whistleblowing and study how this is affected by trust in the government. We manipulate trust by presenting vignettes that describe the situation of a hypothetical agent living in either a high-trust or a low-trust environment. Our choice to focus on the American and the Italian contexts is motivated by real-world considerations. The US is characterized by a long-established whistleblowing culture and low public-sector corruption (ranking 16/180 on an increasing scale of perceived corruption according to the Corruption Perception Index 2017 of the Transparency International). Relative to the USA, Italy is characterized by higher rates of tax evasion (e.g., Schneider, 2005; Schneider & Enste, 2013), a more corrupt public sector (ranking 54/180 according to the Corruption Perception Index 2017) and a recently emerging whistleblowing culture. These differences between the American and the Italian contexts allow us to check the robustness of our conclusions in relation to sources of institutional and social heterogeneity.

We believe that survey-experiments represent an ideal methodological framework for our research question, especially for the possibility of establishing clear causal relationship as well as for the possibility of properly framing the decision task in the taxation context and manipulating the trust in the government. Nevertheless, our methodological choice is not exempted from potential issues. Most importantly, the external validity of the results can be questioned by both the representativeness of the experimental population and the ad-hoc vignette description used to manipulate trust. In an attempt to address these issues, we extend our analysis by using a unique nationwide survey data collected by the Caucasus Research and Resource Centers (CRRC) in the framework of the “Armenia Tax Reform Project.” Taxation is the main topic of the survey and the standardized questionnaire includes a broad set of questions about taxes, respondents’ tax morale, the tax administration, and the level of trust in state institutions (including the Government). To the best of our knowledge, no other survey simultaneously elicits respondents’ trust in public institutions and the attitude toward whistleblowing on tax evaders. The Republic of Armenia is a landlocked post-soviet republic in the South Caucasus. The country is characterized by high corruption levels (ranking 107/180 according to the Corruption Perception Index 2017) and low-quality institutions, which are among the main obstacles hindering tax collection (Davoodi & Grigorian, 2007). Given the institutional incapacity to detect all acts of wrongdoing, third-party reporting of malfeasance can serve as a powerful and cost-effective channel for policy makers to improve the tax collection rates. Taking this into account, in recent years, the Government is actively trying to create a favorable institutional environment for whistleblowers in the country.⁶ Nonetheless, trust in state institutions is chronically low (e.g., Antinyan, Baghdasaryan & Grigoryan, 2018).

⁶ For instance, please see <https://news.am/eng/news/389219.html> (retrieved on October 4, 2018).

Anticipating the results, we document a robust relationship between trust and the attitude toward whistleblowing. Individuals with high trust in the authorities are more likely to report acts of wrongdoing. This result is robust across countries (i.e., Armenia, Italy, USA) and study methodologies (i.e., survey experiment, survey). High tax morale is positively related to individuals' attitude toward reporting tax evaders in the survey-experiment, albeit non-significant in the survey. The perception of tax evasion is not related with the attitude toward whistleblowing neither in the survey nor in the survey-experiment. The same refers to such socio-demographic variables as age, marital status, education, employment status and income. Finally, females are more likely to express a positive attitude toward whistleblowing than males in the survey, whereas we do not detect substantial gender differences in the survey-experiment.

The paper contributes to the limited empirical literature that studies whistleblowing in the context of tax evasion (e.g., Breuer, 2013). The extant experimental and empirical work mainly focuses on whistleblowing behavior of employees in organizations, of cartel members, and of bribers on corrupt officials. This, in our view, is substantially different from the whistleblowing behavior of individuals in the context of taxation. The former type of whistleblowing mainly deals with conspiracies from within organization and refers to “traitorous whistleblowing” (e.g., Breuer, 2013). Furthermore, in this case, the probability of detection and retaliation is relatively high, since the set of concerned counterparts is rather narrow. In contrast, whistleblowing on tax evaders mainly depicts the situation of a citizen who reports peers' tax frauds without being directly involved in the business. Moreover, due to anonymity protection policies generally implemented by tax authorities, the probability of detection and retaliation in this setting is low. For instance, any potential customer can anonymously report to the tax authority that the seller (or the dentist, the mechanic, etc.) did not issue the tax receipt. These differences suggest that there are sound

reasons to believe that the institutional, contextual and personal determinants that shape whistleblowers' decisions and intentions may substantially vary across the organizational and the taxation settings.

The rest of the paper is structured as follows. Section 2 provides a brief literature review. Section 3 sketches the empirical approach. Section 4 details the description of the survey experiment. Section 5 illustrates the survey data, the empirical specification and reports the results of the estimations. Section 6 concludes the paper and suggests policy recommendations.

2. Literature Review

The (experimental) literature dealing with whistleblowing has mainly focused on five research questions detailed below. To the best of our knowledge and with the only exception of Breuer (2013), existing studies have focused on whistleblowing in contexts that are outside the topic of the present paper, namely tax evasion.

Can whistleblowing curb dishonest activities? To this end, experiments have depicted whistleblowing as a powerful tool to prevent such activities as cartel formation (e.g., Apestegua, Dufwenberg & Selten, 2007; Hinloopen & Onderstal, 2007), bribery (Abbink, 2014), software piracy (Oh & Teo, 2010), tax evasion (Breuer, 2013), managerial wrongdoing (e.g., Wallmeier, 2018), organizational lying (Reuben, Reuben & Stein, mimeo), cheating (Grimm et al, 2016) and the like.

How are the whistleblowers perceived in society and are they intrinsically motivated to report the deceptive behavior of others? Interestingly, many individuals are intrinsically motivated and report wrongdoing. However, the whistleblowers are not perceived well even by honest individuals, and this negative stigma effect can substantially shrink the instances of reporting (e.g., Reuben & Stephenson, 2013).

Which legal regime (belief-based or fact-based) protects the whistleblowers the best and encourages them to report acts of wrongdoing truthfully? In a belief-based jurisdiction (e.g., USA), the whistleblower is not required to provide proof of the allegations made: a “reasonable belief” is sufficient to be protected by law, even if the allegation turned out to be incorrect. In fact-based jurisdictions (e.g., France or Germany), whistleblower protection is generally granted after the validation of the whistleblower’s claim. According to experiments, in belief-based system, the reporting of wrongdoing can be higher, nonetheless the reports can contain a substantial number of fraudulent claims. The latter can diminish the prosecutors’ incentives to investigate and can make the deterrence of wrongdoing more difficult (e.g., Mechtenberg, Roider & Muehlheusser, 2017).

Should the whistleblowers be financially rewarded for reporting wrongdoers? Since, the instances of whistleblowing are not ubiquitous enough, incentivizing the whistleblowers is a hot topic. While in the US jurisdiction, the Dodd-Frank Act allows the whistleblowers to receive financial bounties for reporting the information to the Securities and Exchange Commission (SEC), the regulatory agencies in other countries (e.g., the UK) strongly oppose such a practice. The research illustrates that monetary rewards lead to a significant increase in reporting fraudulent behavior (e.g., Breuer, 2013; Butler, Serra & Spagnolo, 2017; Schmolke & Utikal, 2016).

What are the personal and the situational correlates inducing individuals to blow the whistle? Indeed, understanding the personal and situational correlates may help policymakers to estimate who is more likely to blow the whistle and in which situations. This will contribute to the design of both private (e.g., Bartuli, Djawadi & Fahr, 2016) and public institutions as well as incentive schemes. Based on the whistleblowing literature in the organizational context, it is still unclear whether socio-demographic characteristics influence the decision to blow the whistle and what the

sign of such a relationship is. For instance, while some studies find that whistleblowers are more educated, male and older, others either claim that females are more likely to become whistleblowers or do not document gender effect at all (e.g., Near & Miceli, 1996; Mesmer-Magnus & Viswesvaran, 2005; Cassematis & Wortley, 2013). The same argument refers to other socio-demographic factors (e.g., Bartuli, Djawadi & Fahr, 2016). Several researchers go beyond the socio-demographics and study the impact of personality traits (e.g., Bjorkelo, Einarsen & Matthiesen, 2010; Chiu, 2003; Bartuli, Djawadi & Fahr, 2016).

Our contribution to the abovementioned literature is that we experimentally and empirically investigate the relationship between trust in the government and the attitude toward whistleblowing on tax evaders.

Why should trust in the government represent an important determinant of whistleblowing on tax evaders? The “slippery slope” framework by Kirchler, Hoelzl & Wahl (2008) puts forth two main determinants affecting taxpayers: trust and power. Taxpayers’ high trust in authorities results in a cooperative environment in which taxpayers voluntarily declare their income. In contrast, taxpayers’ distrust in the authorities results in an antagonistic climate in which the taxpayers and the state are in conflict with one other. In this case, the state can enforce tax compliance merely by exercising power. The empirical evidence illustrates that high trust in tax authorities is indeed associated with high levels of voluntary tax compliance and an increased sense of civic duty (Muehlbacher, Kirchler & Schwarzenberger, 2011; Kogler, Muehlbacher & Kirchler, 2015; Feld & Frey, 2002). We believe that the uncovering of wrongdoers can be perceived as a cooperative outcome between individuals and the state, whereby citizens are ready to voluntarily expose tax evaders for the sake of increasing tax compliance.

3. Behavioral Prediction and the Empirical Approach

In light of the extant literature, the main prediction we aim to test is whether trust in authorities stimulates citizens' attitude to cooperate with tax authorities by voluntary reporting tax frauds. Thus, we expect that the positive attitude toward whistleblowing is much more pronounced in a high-trust than in a low-trust environment. It is also reasonable to expect that *ceteris paribus*, in a high-trust environment, the probability of actual whistleblowing will be substantially higher than in a low-trust environment.

To uncover the causal impact of trust on whistleblowing attitudes, we administer a survey-experiment in Italy and the US and manipulate the perception of trust in a hypothetical government with ad hoc vignettes. Vignettes are short descriptions of a person or a social situation, which contains factors that might be relevant and salient for the respondent's decision (e.g., Alexander & Becker, 1978). One of the main advantages of a vignette analysis is that it can narrow the gap between the study and the real world, mimicking real decision tasks and situations (Hainmueller, Hangartner & Yamamoto, 2015). This is important for our setting, in which we need to manipulate the level of trust in institutions as realistically as possible. To do so, we depict loaded vignettes with either a well-functioning or a malfunctioning government. To this date, survey-experiments have been successfully used by economists to investigate such important research questions as the concerns for relative standing (Pingle & Mitchell, 2002; Johansson-Stenman, Carlsson & Daruvala, 2002), the determinants of life satisfaction (Angelini, Bertoni & Corazzini, 2016), the willingness to pay for insurance plans (Krueger & Kuziemko, 2013), the perceptions and concerns for distributional fairness (Faravelli, 2007; Cruces, Perez-Truglia & Tetaz, 2013), the effects of information about inequality and taxes on preferences for redistribution (Kuziemko et al., 2015),

as well as the effects of information about public spending on support for public spending (Lergetporer et al., 2016).

Although survey-experiments are a powerful instrument to establish the causal link between trust and the attitude toward whistleblowing, they can raise important concerns of external validity. For instance, the results may suffer from a hypothetical bias, i.e., the responses to the hypothetical scenarios may substantially differ from real behavior. In this regard, Hainmueller, Hangartner & Yamamoto (2015) illustrate the external validity of survey-experiments: the estimated causal effects obtained via hypothetical choices in a vignette study can hold in natural experiments. One may also argue that the results can be attributed to the specific wording of the vignettes. A third critique can be that, if the sample size is not large enough, the results may not hold with a different sample of respondents.

To resolve potential issues associated with survey-experiments, as a second step, we utilize a nationwide cross-sectional household survey conducted by the Caucasus Research and Resource Centers (CRRC) in the framework of the “Armenia Tax Reform Project” in November–December 2013. The survey was carried out with USAID support and the World Bank Group technical assistance.⁷ Taxation is the main focus of the survey and the standardized questionnaire includes a broad set of topics regarding taxes and the tax administration. A typical face-to-face interview lasted approximately an hour.

4. The Survey-Experiment

4.1. Design

Our survey-experiment consists of *High Trust-Low Trust* (HT-LT) and *Low Trust-High Trust* (LT-HT) treatments. In both treatments, the respondents are invited to complete an online

⁷ <http://blogs.worldbank.org/voices/armenia-perception-matters-tax-reforms> (retrieved November 2, 2018).

questionnaire that consists of three sections. The first and the third sections, kept constant across treatments, include general questions about the demographic and socio-economic conditions of the respondents, as well as questions regarding their opinions on taxation. The second section includes vignettes, depicting a third person, who lives in a hypothetical country, in which trust in the government is manipulated depicting it as either high or low. The third person gets to know about instances of tax evasion, which he or she does not report to tax authorities. The respondents are asked to judge the decision of the third person not to report the tax evaders on a scale from 1 (Completely Agree) to 4 (Completely Disagree). The responses to this question constitute the main variable of our interest. All participants answer this question for both high-trust and low-trust vignettes. To avoid order effects, in randomly selected 50% of the cases, high-trust vignette appears first (HT-LT treatment), while in the remaining 50% of the cases the low-trust vignette appears first (LT-HT treatment). For the sake of brevity of the text, the vignettes that appear first are called the first stage vignettes, while the vignettes that appear second are called the second stage vignettes.

Few design issues regarding the vignettes are worth noting. First, the trust manipulation stems from the discussion of Kirchler, Hoelzl and Wahl (2008). More specifically, the government in the high-trust vignette is depicted as one that aims for transparent procedures and for respectful and supportive treatment of taxpayers. Furthermore, the government is described to perform a service for the community by not wasting taxpayers' money and providing all the necessary public goods.

Second, a hypothetical third person living in a hypothetical country is portrayed in the vignettes. Whistleblowing is a sensitive topic, and whistleblowers may be associated with sneaks, spies, rats and the like (e.g., Jubb, 1999). Under these circumstances, a study participant may not report her true preferences when answering whistleblowing questions on her own behalf in order

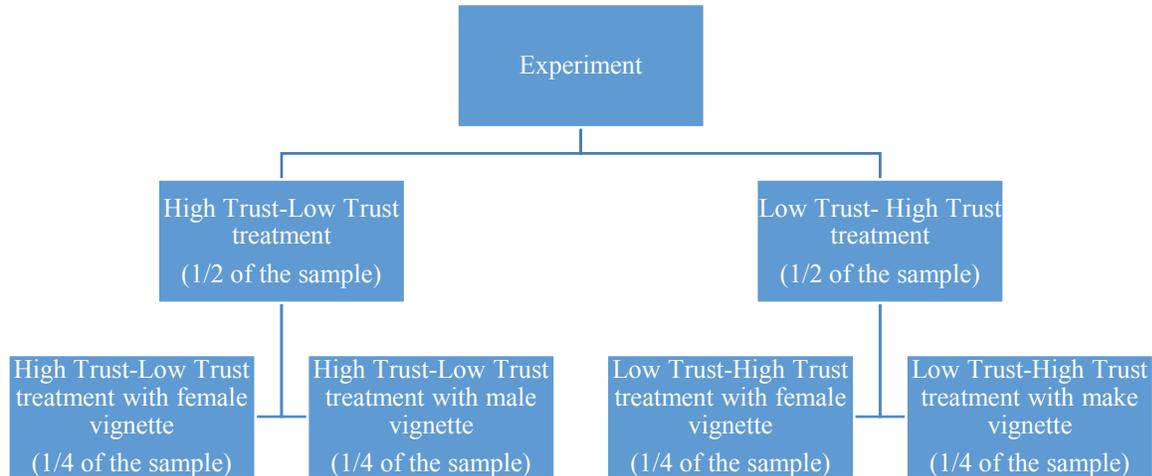
not to harm her identity because of concerns about social approval. By using a third person, we intend to liberate the respondents from their own circumstances. We assume that the respondents will apply their own preferences when making choices on behalf of the third person. Projecting the vignettes onto a third person is an accepted approach in the experimental literature. For instance, Alpizar, Carlsson & Johansson-Stenman (2005), Carlsson, Johansson-Stenman & Martinsson (2007) and Johansson-Stenman, Carlsson & Daruvala (2002) administer survey experiments in which respondents make decisions on behalf of their future grandchildren.

Third, we alter the gender in the vignettes to avoid any possible interaction between the gender of the third person and the responses of the participants. In 50% of the cases, randomly selected, the hypothetical third person is a male, while in the remaining 50% of the cases the hypothetical third person is a female. A respondent faces either male or female vignettes. We use male and female names that are common in the countries where the experiment takes place.⁸ Figure 1 provides a graphical illustration of the structure of the survey-experiment, while Table A1 in Appendix A details the vignettes utilized in the study.

Fourth, our experiment allows both for between-subjects as well as within-subjects comparisons, which, we believe, adds to the generalizability and the robustness of our findings. More specifically, by comparing the responses to the first and the second stage vignettes across HT-LT and LT-HT treatments, we can estimate the between-subjects effect of trust in the government on the attitude toward whistleblowing. In the same vein, by comparing the responses to the first and the second stage vignettes within HT-LT and LT-HT treatments, we can estimate the within-subjects effect of trust in the government on the attitude toward whistleblowing.

⁸ We used Marco and Sarah for Italy, while Robert and Anna for the USA.

Figure 1: The Structure of the Survey-Experiment



Note: The structure of the survey-experiment. The allocation to various treatments and vignettes is randomly determined by the computer.

4.2. Procedures

The survey-experiment was administered via Qualtrics (www.qualtrics.com) and took place in Italy in January 2018 and in the USA in August 2018.

The Italian subject pool was recruited through various Italian student groups on Facebook. The Qualtrics link of the survey-experiment was placed in the respective groups. Once a study participant entered into the study by clicking on the link, she was randomly and anonymously assigned to (only) one of the two treatments.

The American subject pool was recruited via Amazon Mechanical Turk (AMT henceforth). AMT has recently turned into a popular means of experimental data generation for social scientists. AMT is a crowdsourcing marketplace which allows individuals (and businesses) to assign paid tasks to a large population of workers all over the world. The advantage of AMT is that it is a relatively cheap and fast device for data collection (Rand, 2012) with a median hourly wage of

\$1.38 (Horton & Chilton, 2010). The researchers are able to replicate (even complex) experimental studies conducted in the physical laboratory (e.g., Paolacci, Chandler & Ipeirotis, 2010). Overall, 200 subjects were recruited, and each subject was paid 0.6\$ for a 10-minute survey, which results in an hourly wage of 3.6\$.

4.3. Results

In total, 278 subjects out of 350 in Italy (86% response rate) and 172 subjects out of 200 in the USA (79.429% response rate) completed the questionnaire. Table 1 shows some socio-demographic characteristics of the participants in both countries.

Table 1: The Socio-Demographic Characteristics of the Study Participants

	Italy	USA
Mean Age and Standard Deviation	30.417 (11.389)	35.436 (10.900)
Male	138 (49.64%)	99 (57.56%)
Employment Status:	163	152
Working	(58.63%)	(88.37%)
Education		
No education	3 (1.08%)	0 (0%)
School Certificate	122 (43.88%)	36 (20.93%)
Bachelor's Degree	77 (27.7%)	83 (48.26%)
Master's Degree	45 (16.19%)	46 (26.74%)
PhD	31 (11.15%)	7 (4.07%)
Marital Status: Couple	109 (39.21%)	114 (66.28%)
N	278	172

Note: The socio-demographic characteristics of the participants in the survey-experiments. *Employment Status: Working*– A binary variable that is equal to 1 if the respondent is a hired worker, individual entrepreneur/employer or self-employed, and the value of 0 otherwise; *Marital Status: Couple*– A binary variable that is equal to 1 if the respondent is married or cohabitating and 0 otherwise.

The US sample is older than the Italian one, more educated with a larger fraction of non-single participants. These considerations are in line with general demographics of “turkers” in the US who are usually found to be educated and with a median age of 30 (Ross et al., 2010).

4.3.1. Between-Subjects Analysis

We first analyze the respondents’ attitude toward whistleblowing by comparing responses to the first and the second stage vignettes across treatments. Since the respondents’ judgements may depend on the gender of the third person depicted in the vignettes, we check for a possible interaction between the gender of the vignette and the attitude toward whistleblowing. For each country and stage, we compare the responses to the high-trust (low-trust) vignette using a male third person with the responses to the high-trust (low-trust) vignette using a female third person.

The non-parametric Mann-Whitney U test illustrates that the gender in the vignettes does not exert any effect on respondents’ judgments, both when considering the Italian and the American samples in the first stage. Interestingly, we document a gender effect for the American sample in the second stage.

The results are shown in Table 2. Thus, for the first stage analysis we can pool the responses to the high-trust (low-trust) vignette using a male third person with the responses to the high-trust (low-trust) vignette using a female third person. For the second stage analysis, we only pool the data for the Italian sample and for the high-trust vignette of the American sample.

Table 2: Non-parametric analysis for gender effects

	Italy		USA	
	First Stage	Second Stage	First Stage	Second Stage
Marco HT vs. Sarah HT	Z=0.712, p=0.476	Z=-0.198, p=0.843		
Marco LT vs. Sarah LT	Z=0.281, p=0.779	Z=-0.734, p=0.463		
Robert HT vs. Anna HT			Z=-1.038, p=0.299	Z=0.852, p=0.394
Robert LT vs. Anna LT			Z=-0.052, p=0.958	Z=-1.997, p=0.046

Note: Non-parametric Mann-Whitney U tests to check for gender effects in the first and second stage vignettes of the experiment. Marco (Sarah) HT - Marco (Sarah) high-trust vignette that appears in the Italian experiment. Marco (Sarah) LT - Marco (Sarah) low-trust vignette that appears in the Italian experiment. Robert (Anna) HT - Robert (Anna) high-trust vignette that appears in the American experiment. Robert (Anna) LT - Robert (Anna) low-trust vignette that appears in the American experiment.

We also check whether the sequence of the vignettes creates order effects. For instance, the responses to the high-trust vignette question may differ depending on whether the high-trust vignette appears in the first or the second stage (i.e., low-trust vignette appears in the first stage). The same refers to the answers to the low-trust vignette question. The non-parametric Mann-Whitney U test documents no order effects. Only in the comparison of the vignettes with a female third person in the American and the Italian samples we detect marginally significant differences. The results are depicted in Table 3.

Table 3: Non-parametric analysis for order effects

	Italy	USA
Male 1 st Stage HT vs. Male 2 nd Stage HT	Z=-1.128, p=0.260	Z=-1.058, p=0.290
Male 1 st Stage LT vs. Male 2 nd Stage LT	Z=-0.723, p=0.470	Z=-0.033, p=0.974
Female 1 st Stage HT vs. Female 2 nd Stage HT	Z=-1.888, p=0.06	Z=0.821, p=0.412
Female 1 st Stage LT vs. Female 2 nd Stage LT	Z=0.296, p=0.768	Z=1.747, p=0.081

Note: Non-parametric Mann-Whitney U tests to check for order effects in the experiment. We compare Marco high-trust (low-trust) 1st stage vignette with Marco high-trust (low-trust) 2nd stage vignette, Sarah high-trust (low-trust) 1st stage vignette with Sarah high-trust (low-trust) 2nd stage vignette, Robert high-trust (low-trust) 1st stage vignette with Robert high-trust (low-trust) 2nd stage vignette and Anna high-trust (low-trust) 1st stage vignette with Anna high-trust (low-trust) 2nd stage vignette.

Next, we focus on the impact of trust in the government on the attitude toward whistleblowing, utilizing the pooled dataset wherever applicable. As a reminder, the respondents had to state how much they agree or disagree with a third person's decision not to report the instances of tax evasion around him or her on a scale from 1 (Completely Agree) to 4 (Completely Disagree). Table 4

reports the mean level of agreement across different treatments. The standard deviation is reported in parentheses.

Table 4: The Mean Level of Agreement in the Treatments

	Italy		USA	
	First Stage	Second Stage	First Stage	Second Stage
HT	2.862 (0.785)	3.057 (0.820)	2.5 (0.935)	2.536 (1.069)
LT	2.579 (0.769)	2.551 (0.855)	2.083 (0.921)	
Robert LT (*)				2.048 (0.825)
Anna LT (*)				2.391 (0.829)
Mann-Whitney U test	Z=2.918, p=0.004	Z=4.854, p=0.000	Z=2.981, p=0.003	

Note: The mean level of agreement or disagreement with the decision not to report tax evaders. The question is measured on a scale from 1 (Completely Agree) to 4 (Completely Disagree), thus high values suggest a higher level of disagreement. Standard deviations are reported in the parentheses.

(*) As already discussed, due to the effect of the gender of the vignette, we cannot pool the data for Robert Low-Trust and Anna Low-Trust vignettes in the second stage of the American Sample.

The table suggests that, in both countries, respondents' disagreement is higher in the high-trust vignette than in the low-trust vignette. The non-parametric Mann-Whitney U test detects significant treatment effect in Italy both in the first stage ($Z=2.918$, $p=0.004$) and in the second stage ($Z=4.854$, $p=0.000$) vignettes. Likewise, significant treatment effect is present in the first stage vignettes in the USA ($Z=2.981$, $p\text{-value}=0.003$). In the second stage, we document a significant treatment difference between Robert Low-Trust and the pooled High-Trust vignettes only.⁹

The formal regression analysis confirms the results of the non-parametric tests. As our dependent variable is measured on a scale from '1' to '4,' we treat it as an interval variable and

⁹ Recall, that in the American sample we can pool the responses to the high-trust vignette with a male third person with the responses to the high-trust vignette with a female third person. However, we cannot pool the responses to the low-trust vignette with a male third person with the responses to the low-trust vignette with a female third person. Thus, we separately compare the responses to Robert Low-Trust and Anna Low-Trust vignettes with those of the pooled High-Trust vignette.

estimate OLS and tobit models for each country.¹⁰ Throughout the text we comment on those variables that are robust across different models. Table 5 depicts the results of the estimations. We regress the respondents' agreement not to blow the whistle on tax evaders on a *High-Trust dummy*, which assumes the value of 1 if the vignette refers to the high-trust scenario and of 0 otherwise, as well as on a set of personal and contextual controls (tax morale, tax evasion perception and the variables reported in Table 1). Tax evasion perception is an integer that assesses respondents' perception of tax evasion in the country on a scale from 1 (almost no one evades taxes) to 10 (almost everyone evades taxes).

In the first stage, we utilize the pooled dataset for both countries. In the second stage, we use the pooled dataset for the Italian sample only. For the American sample, we distinguish between low-trust vignettes with male and female third persons. To do so, we introduce a dummy variable *Anna Low Trust*, which assumes the value of 1 in the low-trust vignette with a female third person and 0 otherwise. Thus, *Robert Low Trust* is the omitted category.

¹⁰ To check the robustness of our results, we preserve the ordinal nature of the scale and estimate ordered probit models. Conclusions remain qualitatively unchanged. Results are available upon request.

Table 5: Between-Subjects Regression Analysis

	Italy				USA			
	First Stage		Second Stage		First Stage		Second Stage	
	(1) OLS	(2) Tobit	(3) OLS	(4) Tobit	(5) OLS	(6) Tobit	(7) OLS	(8) Tobit
Constant	1.384*** (0.290)	1.121*** (0.379)	1.425*** (0.298)	1.030** (0.464)	1.086** (0.449)	0.381 (0.714)	1.526*** (0.507)	1.204 (0.824)
High-Trust Dummy	0.252*** (0.087)	0.326*** (0.109)	0.558*** (0.096)	0.763*** (0.135)	0.385*** (0.133)	0.560*** (0.197)	0.507*** (0.172)	0.702*** (0.272)
Anna Low Trust							0.319* (0.176)	0.492* (0.300)
Male	-0.040 (0.094)	-0.017 (0.114)	-0.062 (0.099)	-0.082 (0.140)	0.234* (0.130)	0.355* (0.201)	0.072 (0.146)	0.079 (0.223)
Working	0.102 (0.089)	0.139 (0.118)	0.150 (0.097)	0.174 (0.145)	-0.254 (0.222)	-0.456 (0.307)	0.205 (0.218)	0.298 (0.346)
Age	0.003 (0.005)	0.004 (0.006)	0.006 (0.005)	0.010 (0.007)	-0.005 (0.005)	-0.005 (0.009)	-0.007 (0.007)	-0.010 (0.010)
Education	0.042 (0.055)	0.052 (0.061)	-0.083 (0.057)	-0.115 (0.075)	0.147* (0.089)	0.242* (0.128)	0.008 (0.087)	-0.005 (0.141)
Tax Morale	0.097*** (0.018)	0.117*** (0.022)	0.102*** (0.020)	0.135*** (0.026)	0.114*** (0.022)	0.174*** (0.033)	0.062** (0.025)	0.096*** (0.036)
Tax Evasion Rate Perception	0.033 (0.028)	0.040 (0.034)	0.049* (0.028)	0.067 (0.041)	0.011 (0.028)	-0.001 (0.045)	0.020 (0.036)	0.016 (0.049)
N	278	278	278	278	172	172	172	172
F or χ^2	8.81	46.60	10.24	56.84	7.44	41.08	2.58	14.88
P>F or χ^2	0.000	0.000	0.000	0.000	0.000	0.000	0.011	0.062
R ² or pseudo R ²	0.161	0.062	0.196	0.071	0.208	0.079	0.091	0.028

Note: Results from OLS (robust standard errors in parentheses) and Tobit (standard errors in parentheses) models. Dependent variable: Respondents attitude about not reporting the tax evaders measured on a scale from 1 (Completely Agree) to 4 (Completely Disagree). Independent Variables: *High-Trust Dummy*- Dummy variable which equals to 1 in high-trust vignettes and 0 otherwise; *Male*- Dummy variable which equals to 1 if the respondent is male and 0 otherwise; *Working*- Dummy variable which equals to 1 if the respondent is employed either full-time or part-time or self-employed and 0 otherwise; *Age*- An integer, indicating the age of the respondent; *Education*- An integer that indicates the highest level of the respondent's education and assumes the values of 1 (No education), 2 (School certificate); 3 (Bachelor's degree); 4 (Master's degree) and 5 (PhD); *Tax Morale*- An integer that elicits the tax morale of the respondent on a scale from 1 (Completely justified to evade taxes) to 10 (Completely unjustified to evade taxes); *Tax Evasion Rate Perception*- An integer that elicits respondent's perception of tax evasion on a scale from 1 (Almost no one evades taxes) to 10 (Almost everyone evades taxes).

The positive and highly significant coefficient of the *High-Trust Dummy* suggests that the disagreement with not reporting the tax evaders is higher in the high-trust vignette than in the low-trust vignette. One can also evidence an interesting relationship between tax morale and the attitude toward whistleblowing. *Ceteris paribus*, the higher the respondent's tax morale the higher the disagreement with not reporting the tax evaders. As for the relationship between socio-demographic characteristics and the attitude toward whistleblowing, gender, education, the employment status, and age do not seem to play a key role. Tax evasion perception is not associated with the attitude toward whistleblowing either. We summarize the previous results in the following statement.

Result 1. *Trust toward the government has a positive impact on individuals' attitude toward blowing the whistle on tax evaders.*

4.3.2. Within-Subjects Analysis

In this section, we take advantage of the longitudinal dimension of the survey-experiment to study the effect of trust in the government on whistleblowing attitude at the within-subjects level. For the Italian sample, we disregard the gender and compare the pooled high-trust data with the pooled low-trust data. For the American sample, we undertake separate analysis for male and female vignettes, since we detected an interaction between the gender of the third person in the vignette and the attitude toward whistleblowing in the second stage of the experiment.

The non-parametric Wilcoxon signed-rank test for related samples suggests a significant treatment effect both in LT-HT and HT-LT treatments. The results are depicted in Table 6.

Table 6: Wilcoxon signed-rank test

	Italy	USA
<i>HT-LT</i>		
High-Trust Vignette vs. Low-Trust Vignette	Z=4.418, p=0.000	
Robert High-Trust vs. Robert Low-Trust		Z=2.722, p=0.007
Anna High-Trust vs. Anna Low-Trust		Z=1.789, p=0.074
<i>LT-HT</i>		
Low-Trust Vignette vs. High-Trust Vignette	Z=-6.474, p=0.000	
Robert Low-Trust vs. Robert High-Trust		Z=-3.106, p=0.002
Anna Low-Trust vs. Anna High-Trust		Z=-1.808, p=0.071

Note: Non-parametric Wilcoxon signed-rank test for related samples to make within-subjects comparisons in HT-LT and LT-HT treatments.

The non-parametric results are confirmed by parametric analysis. Given the structure of the data, we estimate GLS random-effects models. More specifically, we regress the respondents' agreement to blow the whistle on tax evaders on a *Second Stage Vignette* dummy, which is equal to 1 if the vignette comes from the second stage and 0 otherwise, and on the set of personal and contextual controls utilized for between-subjects analysis.

Table 7: Within-Subjects Regression Analysis

	Italy			USA		
	(1) LT-HT	(2) HT-LT	(3) Robert: LT-HT	(4) Robert: HT-LT	(5) Anna: LT-HT	(6) Anna: HT-LT
Constant	1.475*** (0.313)	1.647*** (0.393)	1.125 (1.037)	1.336* (0.697)	1.777** (0.786)	1.993* (1.065)
Second Stage Vignette	0.479*** (0.073)	-0.312*** (0.071)	0.535*** (0.161)	-0.357*** (0.128)	0.366** (0.182)	-0.196* (0.119)
Male	0.044 (0.118)	-0.111 (0.118)	0.412* (0.247)	0.124 (0.221)	-0.277 (0.208)	0.244 (0.228)
Working	0.031 (0.108)	0.188* (0.115)	-0.110 (0.582)	0.005 (0.397)	-0.029 (0.240)	-0.114 (0.478)
Age	0.006 (0.005)	0.004 (0.007)	-0.011 (0.018)	-0.007 (0.008)	-0.008 (0.011)	-0.004 (0.007)
Education	-0.104* (0.056)	0.062 (0.073)	0.198 (0.134)	-0.012 (0.120)	0.135 (0.161)	0.023 (0.186)
Tax Morale	0.114*** (0.022)	0.092*** (0.022)	0.049 (0.046)	0.149*** (0.041)	0.111*** (0.035)	0.061 (0.043)
Tax Evasion Rate Perception	0.052* (0.029)	0.023 (0.039)	0.057 (0.070)	0.048 (0.048)	-0.064 (0.045)	0.036 (0.073)
N	280	276	86	84	82	92
χ^2	81.59	52.61	33.52	25.42	18.57	8.12
$P > \chi^2$	0.000	0.000	0.000	0.001	0.010	0.405
R^2	0.226	0.151	0.182	0.288	0.195	0.323

Note: Results from GLS (robust standard errors in parentheses) random-effects models. *Second Stage Vignette*–Dummy variable which equals 1 if the vignette comes from the second stage and 0 otherwise. All other remarks of Table 5 apply.

As shown by the positive and significant coefficient of the *Second Stage Vignette* in columns 1, 3, and 5, the level of respondents' disagreement when others choose not to report the tax evaders, increases if one moves from a low-trust to a high-trust environment. On the contrary, the evidence in columns 2, 4, and 6 suggests that individuals agree more with hiding tax evaders if one moves from a high-trust to a low-trust environment. These results lead us to the following conclusion.

Result 2. *Even at the within-subjects level, trust in the government exerts a positive effect on attitude toward blowing the whistle on tax evaders.*

5. The National Survey in Armenia

The survey covered all the regions of the Republic of Armenia. The goal of the initiative was to assess the public awareness of the tax system in Armenia, to understand citizens' perceptions of the efficiency of the tax policy and of the services provided by the tax administration, as well as to capture the main reasons of taxpayer dissatisfaction with taxation.¹¹ The sampling frame of the survey was built from the database of electricity users provided by the Electric Networks of Armenia, which is the only electricity provider in the country. Given that roughly 100% of Armenian households use electric power and that the electricity is provided by a single company, this database is amongst the most exhaustive and up-to-date lists of households available for the country. Regarding sampling strategy, multistage cluster sampling with stratification by administrative regions and the settlement type (i.e., the capital, other urban and rural areas) was implemented. Random sampling of clusters within each stratum was conducted and a random sample of households within each cluster was selected. The margin of error was specified at the 95% confidence interval. Moreover, a reserve sample of equal size was constructed with the same methodology, to cover the non-response rate where necessary. The respondent in each household was chosen among the members of age 18 and

¹¹ It is important to notice that the survey was not explicitly conducted to understand the factors influencing taxpayers' attitude toward whistleblowing.

above by the next birthday method of respondent selection, whereby the household member whose birthday is the closest is selected for an interview. The data were collected via face-to-face interviews with the help of standardized questionnaires and show cards, which is a common mode for data collection used in influential surveys (e.g., the World Value Survey, among others). Weights were introduced to ensure the representativeness of the survey. The household and individual weights were calculated with the non-response rate taken into consideration.

5.1. The Empirical Specification

We estimate a regression equation of the following form:

$$Y_i = \beta_0 + \beta_1 \times Trust_i + \beta_2 \times X_i + \varepsilon_i. \quad (1)$$

The dependent variable, Y_i , is individual i 's attitude toward whistleblowing, built as a dummy variable that takes the value of 1 if the respondent answered positively to the question: *“Is it proper to inform the tax authorities in case the requirements of the tax legislation are violated (for example, business entities do not provide cash receipts)?”*.¹² $Trust_i$ indicates individual i 's trust in formal authorities and the government. Specifically, the trust in the government is built upon respondents' answer to the question: *“Using a scale from 1–5, where ‘1’ means ‘Fully Distrust’ and ‘5’ means ‘Fully Trust’, please tell me to what extent do you trust the following institutions?”* The executive government is included in the list of institutions the respondents had to state their trust in.¹³ X_i contains personal and contextual control variables that were used in the vignette study. In all regressions, we account for the potential

¹² Unlike in the USA, the whistleblowers are neither materially rewarded nor fully protected for exposing tax evaders in the Republic of Armenia.

¹³ As a robustness check we concentrated our attention on the three principal decision-making institutions in the country: the parliament, the president, and the executive government (prime minister and the other ministers). We constructed a combined index by averaging the trust toward the seminal institutions and treating it as our variable of interest. The estimation results remain intact.

heteroskedasticity of residuals by introducing White robust standard errors. Table 8 details the variables under consideration and provides the descriptive statistics of the data.

Table 8: Variables and Descriptive Statistics

Variable	Description	Frequency, Mean and Standard Deviation (*)	Missing Values
Whistleblowing: Yes	Is it proper to inform the tax authorities in case the requirements of the tax legislation are violated (for example, business entities do not provide cash receipts)?	624/1363	85
Evasion Perception	Do you think that people cheat on taxes whenever possible? Use a scale from 1-10, where 1 means “Nearly no one” and 10 means “Nearly everyone.”	6.563 (2.737)	133
Tax morale	Using a scale from 1 “Fully justified” to 10 “Not justified at all,” please tell me how justified it is to cheat on taxes whenever possible.	7.818 (2.801)	58
Trust Government	Tell me please to what extent you trust the following institutions on a scale from 1 “Fully distrust” to 5 “Fully trust.” <ul style="list-style-type: none"> • Government 	2.130 (1.245)	32
Age		47.424 (18.056)	0
Education	Please tell me what the highest level of formal education you have completed so far is. <ol style="list-style-type: none"> 1. No formal education 2. Primary education 3. Basic Education 4. Secondary education 5. Vocational Education 6. Bachelor’s degree 7. Master’s degree 8. PhD 		5
Marital Status: Single		524/1443	5
Male		524/1448	0
Employment Status: Working		445/1444	4
Income	In total, how much was your personal net monetary income (after the taxes are paid) last month? <ol style="list-style-type: none"> 1. More than 600001 AMD 2. 380001-600000 AMD 3. 190001-380000 AMD 4. 106001-190000 AMD 5. 40001-106000 AMD 6. 24001-40000 AMD 7. Up to 24000 AMD 8. 0 		36

Note: The descriptive statistics of the variables used in the analysis. (*) In case of Whistleblowing, Education, Marital Status, Male, Employment Status and Income frequencies are provided.

Around 36% of the sample is male. Approximately 36% of the respondents are single and 31% are working (either entrepreneur, or self-employed or blue-, white-collar worker).¹⁴ The mean respondents' age is 47 years old. Roughly 23% of the respondents have a university education. The trust in the government seems to be quite low. Interestingly, the views on blowing the whistle on fraudulent behavior are rather diverse. Most of the respondents (54.22%) seem to be against reporting instances of tax evasion.

The survey includes responses from 1448 subjects. Nonetheless, the dataset contains missing values. While listwise deletion is one of the most frequently used methods in handling missing data, it reduces the sample size and the precision of the results. In our case, the listwise deletion of the missing values would reduce the sample size by 258 observations.

To overcome the problem, we conduct multiple imputation using chained equations (MICE) since missing values occur in the majority of the variables. We used 20 imputations. An underlying assumption of MICE is that the values in the dataset are missing at random (MAR), which means that the probability that a value is missing depends only on observed values and not on unobserved values (e.g., Azur et al., 2011; Schafer & Graham, 2002).¹⁵ In the MICE procedure, several regression models are run whereby every variable with missing data is modeled conditional upon the other variables in the dataset. If a binary variable contains missing values logistic regressions are run, if a continuous variable contains missing values linear regressions are utilized, and if a categorical variable contains missing values ordered logistic regressions are run.

¹⁴ We distinguish between those who actually work and those who do not work. Our definition diverges from the official definition of unemployment, which can create discrepancy with the official statistics. Please also note that there is active debate whether the official unemployment statistics corresponds to the real unemployment in the country.

¹⁵ In order to provide suggestive evidence on this assumption, we run formal regression analysis and illustrate that the missingness pattern of the variables in the dataset either depends on observed characteristics (e.g., age, education, income, marital status and the like) or does not depend on observed characteristics at all. The results of the regressions are available upon request.

5.2.Results

Given the binary nature of the dependent variable, we estimate (1) by running logistic regressions. For the sake of robustness of our results, we estimate equation (1) either by imputing missing data through MICE or by using the original dataset after dropping the missing values.

Table 9 reports the estimates. The results in columns 1 and 2 are obtained with the imputed dataset, whereas the results in column 3 are obtained with the reduced dataset. The estimated odds ratios are reported for the ease of interpreting the results. A value above one indicates that the covariate increases the likelihood of reporting a positive attitude toward whistleblowing, while a value below one decreases the same likelihood. Moreover, the further the value is from one, the stronger the impact of the independent variable.

Table 9: The Results of the Logistic Regression

	(1) Logit with Imputed Data	(2) Logit with Imputed Data	(3) Logit with Reduced Data
Intercept	0.591*** (0.074)	0.799 (0.317)	0.603 (0.220)
Trust Government	1.217*** (0.061)	1.198*** (0.061)	1.164*** (0.056)
Male		0.703*** (0.096)	0.766** (0.099)
Age		0.996 (0.004)	0.995 (0.004)
Single		0.930 (0.123)	0.879 (0.110)
Working		0.916 (0.156)	1.077 (0.171)
Vocational Education		0.873 (0.136)	0.929 (0.135)
Bachelor's degree		0.757 (0.219)	0.646* (0.168)
Graduate Education		0.998 (0.190)	1.121 (0.192)
Low-Income		0.928 (0.170)	1.159 (0.186)
High-Income		1.127 (0.487)	1.605 (0.602)
Perception of Tax Evasion		1.009 (0.024)	1.007 (0.022)
Tax Morale		1.019 (0.023)	1.034 (0.022)
N	1448	1448	1190
Imputations	20	20	
F or Wald- χ^2	15.31	2.12	27.30
p>F	0.000	0.013	0.007

Note. Logistic Regression. Dependent variable. *Whistleblow*– Binary variable that equals to 1 if the respondent thinks it is proper to report tax evaders, and to 0 if the respondent thinks it is not proper to report tax evaders. Independent variables. *Single*– Dummy variable that equals to 1 if the respondent has never been married, is divorced, separated or widowed, and to 0 otherwise; *Vocational Education*– Dummy variable that equals to 1 if the respondent has vocational education and to 0 otherwise; *Bachelor's degree*– Dummy variable that equals to 1 if the bachelor's degree is the highest degree of the respondent, and to 0 otherwise; *Graduate Education*– Dummy variable that equals to 1 if either the master's degree or the doctoral degree is the highest degree of the respondent, and to 0 otherwise (*No Education or School Education Only* is the reference category); *Low Income*– Dummy variable that equals to 1, if the respondent indicates either no income or income in the intervals of up to 20.000 AMD, 20.000-34.000 AMD and to 0 otherwise; *High Income*– Dummy variable that equals to 1, if the respondent indicates income in the intervals 162.001-325.000 AMD, 325.001-512.000 AMD or above 512.000 AMD and to 0 otherwise (*Average Income* is the reference category); *Trust Government*– Trust toward the government measured on a scale from 1 (Fully Distrust) to 4 (Fully Trust); In 2013, based on the reports of the Central Bank of Armenia 1 USD=409.588 AMD on average. Robust standard errors in parentheses. All other remarks of Table 5 apply. Significance Levels: * p<10, ** p<5%, *** p<1%.

Table 9 suggests a significant relationship between trust in the government and the attitude toward whistleblowing. More specifically, the odds ratio of *Trust Government* greater than one indicates that trust increases the likelihood of a positive attitude toward whistleblowing. This result is in line with the findings of the vignette study. We also detect a gender effect, whereby the odds ratio of *Male* (smaller than one) indicates that being male decreases the likelihood of

expressing a positive attitude for whistleblowing. An intuitive justification for the gender effect is that whistleblowers may be associated with sneaks, and a whistleblower (if uncovered) may lose her reputation among her peers. The masculine identity in Armenia (like other patriarchal societies) is closely linked to reputation, status and power and there are a number of activities in which men are unlikely to be engaged, including reporting corruption (e.g., Duban, 2010). In line with the results drawn from the survey-experiment, the perception of tax evasion does not turn out to be significantly related with the attitude toward whistleblowing. The same refers to such socio-demographic variables as age, education, employment status and income.

In sum, the survey evidence from Armenia confirms the positive relationship between trust in authorities and whistleblowing documented in the survey-experiment.

6. Discussion and Concluding Remarks

We have investigated the relationship between trust in the government and the respondents' attitude toward reporting tax evasion. To establish causality and guarantee external validity, we have presented results from a survey-experiment conducted in Italy and the US and from a nationwide household survey data from the Republic of Armenia. Our results indicate that citizens with high trust in formal authorities are more likely to report taxpayers' fraudulent behavior under different methodological approaches and cultural contexts. Interestingly, we detect some gender effect, whereby females are more prone to whistleblowing than males in Armenia.

Our results contribute to the literature on the institutional determinants of taxpayer behavior. Indeed, our results complement the empirically validated positive relationship between trust in the government, tax morale and tax compliance (Kirchler, Hoelzl and Wahl, 2008; Feld & Frey, 2002), suggesting that trust is likely to strengthen the effectiveness of whistleblowing, which is an innovative and relatively cheap instrument to contrast tax evasion.

From a policy perspective, our findings are of special relevance for third-world countries like Armenia, which suffer from chronic distrust in public institutions, albeit strive to establish

effective whistleblowing mechanisms to encourage reporting fraudulent tax behaviors. Enhancing taxpayer trust in the government and in other formal institutions may represent the first crucial step to encourage whistleblowing on tax evaders, before designing complex legal regimes and incentive schemes. In this regard, it may be of vital importance to make sure that the public is aware of the government spending as well as convinced in its efficacy and appropriateness (at least on average). The policymakers may have a myriad of choices here, starting from involving taxpayers in defining the public spending in cooperation with the government (e.g., Lambertson, 2013) and ending with taxpayer calculators (e.g., HMRC tax calculator) or targeted letters sent to taxpayers about the public goods financed by the taxes (e.g., Castro & Scartascini, 2015; Doerrenberg & Schmitz, 2015).

The policymakers may also want to consider the influence of personal variables to enhance reporting activities. In patriarchal societies like Armenia, targeted campaigns can be run to emphasize the positive aspects of whistleblowing and to create a positive image of whistleblowers in the eyes of males. Furthermore, in some countries, the campaigns may also be targeted to increasing individuals' intrinsic motivation to pay taxes. The results of the survey-experiment in Italy and the USA provide suggestive evidence of this claim.

References

- Abbink, K., Dasgupta, U., Gangadharan, L., & Jain, T. (2014). Letting the briber go free: An experiment on mitigating harassment bribes. *Journal of Public Economics*, *111*, 17-28.
- Alexander, C. S., & Becker, H. J. (1978). The use of vignettes in survey research. *Public opinion quarterly*, *42*(1), 93-104.
- Alpizar, F., Carlsson, F., & Johansson-Stenman, O. (2005). How much do we care about absolute versus relative income and consumption?. *Journal of Economic Behavior & Organization*, *56*(3), 405-421.
- Angelini, V., Bertoni, M., & Corazzini, L. (2017). Unpacking the determinants of life satisfaction: a survey experiment. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *180*(1), 225-246.
- Antinyan, A., Baghdasaryan, V., & Grigoryan, A. (2018). *Social Preferences, Public Good Provision, Social Capital and Positional Concerns: Empirical Evidence from the South Caucasus* (No. wp625). The Center for Economic Research and Graduate Education-Economics Institute, Prague.
- Apesteguia, J., Dufwenberg, M., & Selten, R. (2007). Blowing the whistle. *Economic Theory*, *31*(1), 143-166.
- Azur, M. J., Stuart, E. A., Frangakis, C., & Leaf, P. J. (2011). Multiple imputation by chained equations: what is it and how does it work?. *International journal of methods in psychiatric research*, *20*(1), 40-49.
- Bartuli, J., Mir Djawadi, B., & Fahr, R. (2016). Business ethics in organizations: an experimental examination of whistleblowing and personality (No. wp1090). *IZA Discussion Papers*.
- Bjørkelo, B., Einarsen, S., & Matthiesen, S. B. (2010). Predicting proactive behaviour at work: Exploring the role of personality as an antecedent of whistleblowing behaviour. *Journal of Occupational and Organizational Psychology*, *83*(2), 371-394.

Breuer, L. (2013). Tax compliance and whistleblowing—the role of incentives. *The Bonn Journal of Economics*, 2(2), 7-44.

Butler, J., Serra, D., & Spagnolo, G. (2017). Motivating whistleblowers. *CEIS Tor Vergata Research Paper Series*, 15(9).

Carlsson, F., Johansson-Stenman, O. L. O. F., & Martinsson, P. (2007). Do you enjoy having more than others? Survey evidence of positional goods. *Economica*, 74(296), 586-598.

Casematis, P. G., & Wortley, R. (2013). Prediction of whistleblowing or non-reporting observation: The role of personal and situational factors. *Journal of business ethics*, 117(3), 615-634.

Castro, L., & Scartascini, C. (2015). Tax compliance and enforcement in the pampas evidence from a field experiment. *Journal of Economic Behavior & Organization*, 116, 65-82.

Chiu, R. K. (2003). Ethical judgment and whistleblowing intention: Examining the moderating role of locus of control. *Journal of Business Ethics*, 43(1-2), 65-74.

Choo, L., Grimm, V., Horvath, G., & Nitta, K. (2016). Whistleblowing and Diffusion of Responsibility: An Experimental Investigation.

Cruces, G., Perez-Truglia, R., & Tetaz, M. (2013). Biased perceptions of income distribution and preferences for redistribution: Evidence from a survey experiment. *Journal of Public Economics*, 98, 100-112.

Davoodi, H. R., & Grigorian, D. A. (2007). *Tax potential vs. tax effort: a cross-country analysis of Armenia's stubbornly low tax collection* (No. 7-106). International Monetary Fund.

Doerrenberg, P., & Schmitz, J. (2015). Tax compliance and information provision—A field experiment with small firms. *IZA Discussion Papers No. 9013*.

Duban, E. (2010). Gender assessment. *Yerevan: USAID/Armenia*, 3.

Faravelli, M. (2007). How context matters: A survey based experiment on distributive justice. *Journal of Public Economics*, 91(7-8), 1399-1422.

Feld, L. P., & Frey, B. S. (2002). Trust breeds trust: How taxpayers are treated. *Economics of Governance*, 3(2), 87-99.

Hainmueller, J., Hangartner, D., & Yamamoto, T. (2015). Validating vignette and conjoint survey experiments against real-world behavior. *Proceedings of the National Academy of Sciences*, 112(8), 2395-2400.

Hinloopen, J., & Onderstal, S. (2014). Going once, going twice, reported! Cartel activity and the effectiveness of antitrust policies in experimental auctions. *European Economic Review*, 70, 317-336.

Horton, J. J., & Chilton, L. B. (2010, June). The labor economics of paid crowdsourcing. In *Proceedings of the 11th ACM conference on Electronic commerce* (pp. 209-218). ACM.

Johansson-Stenman, O., Carlsson, F., & Daruvala, D. (2002). Measuring future grand parents' preferences for equality and relative standing. *The Economic Journal*, 112(479), 362-383.

Jubb, P. B. (1999). Whistleblowing: A restrictive definition and interpretation. *Journal of Business Ethics*, 21(1), 77-94.

Krueger, A. B., & Kuziemko, I. (2013). The demand for health insurance among uninsured Americans: Results of a survey experiment and implications for policy. *Journal of Health Economics*, 32(5), 780-793.

Kirchler, E., Hoelzl, E., & Wahl, I. (2008). Enforced versus voluntary tax compliance: The 'slippery slope' framework. *Journal of Economic Psychology*, 29(2), 210-225.

Kogler, C., Muehlbacher, S., & Kirchler, E. (2015). Testing the "slippery slope framework" among self-employed taxpayers. *Economics of Governance*, 16(2), 125-142.

Kuziemko, I., Norton, M. I., Saez, E., & Stantcheva, S. (2015). How elastic are preferences for redistribution? Evidence from randomized survey experiments. *American Economic Review*, 105(4), 1478-1508.

Lamberton, C. (2013). A spoonful of choice: How allocation increases satisfaction with tax payments. *Journal of Public Policy & Marketing*, 32(2), 223-238.

Lergetporer, P., Schwerdt, G., Werner, K., & Woessmann, L. (2016). Information and preferences for public spending: Evidence from representative survey experiments.

Mechtenberg, L., Muehlheusser, G., & Roider, A. (2017). Whistle-blower protection: theory and experimental evidence.

Mesmer-Magnus, J. R., & Viswesvaran, C. (2005). Whistleblowing in organizations: An examination of correlates of whistleblowing intentions, actions, and retaliation. *Journal of business ethics*, 62(3), 277-297.

Muehlbacher, S., Kirchler, E., & Schwarzenberger, H. (2011). Voluntary versus enforced tax compliance: Empirical evidence for the ‘slippery slope’ framework. *European Journal of Law and Economics*, 32(1), 89-97.

Near, J. P., & Miceli, M. P. (1996). Whistle-blowing: Myth and reality. *Journal of management*, 22(3), 507-526.

Oh, L. B., & Teo, H. H. (2010). To blow or not to blow: An experimental study on the intention to whistleblow on software piracy. *Journal of Organizational Computing and Electronic Commerce*, 20(4), 347-369.

Paolacci, G., Chandler, J., & Ipeirotis, P. (2010). Running experiments on amazon mechanical turk. *Judgment and Decision Making*, 5(5), 411-419.

Pingle, M., & Mitchell, M. (2002). What motivates positional concerns for income?. *Journal of Economic Psychology*, 23(1), 127-148.

Rand, D. G. (2012). The promise of Mechanical Turk: How online labor markets can help theorists run behavioral experiments. *Journal of theoretical biology*, 299, 172-179.

Reuben, E., & Stephenson, M. (2013). Nobody likes a rat: On the willingness to report lies and the consequences thereof. *Journal of Economic Behavior & Organization*, 93, 384-391.

Reuben, A., Reuben, E., & Stein, C. Whistleblowing and Incentives: The Role of Incentive Schemes on the Willingness to Blow the Whistle.

Schafer, J. L., & Graham, J. W. (2002). Missing data: our view of the state of the art. *Psychological methods*, 7(2), 147.

Schmolke, K. U., & Utikal, V. (2016). Whistleblowing: Incentives and situational determinants (No. wp9). FAU Discussion Papers in Economics.

Schneider, F. (2005). Shadow economies around the world: what do we really know? *European Journal of Political Economy*, 21(3), 598-642.

Schneider, F., & Enste, D. H. (2013). *The shadow economy: An international survey*. Cambridge University Press.

Torgler, B. (2003). To evade taxes or not to evade: that is the question. *The Journal of Socio-Economics*, 32(3), 283-302.

Wallmeier, N. (2018). The Hidden Costs of Whistleblower Protection.

Appendices

Appendix A: Vignettes

Table A1: The vignettes used in the survey experiment

High-Trust Vignette
Robert lives in country X. The taxpayers in country X exhibit very high trust toward their Government for several reasons. First, no tax money is wasted. The Government operates transparently and makes efficient use of the taxes, providing the necessary public goods and services to the citizens. Second, all taxpayers are equal against the law for the Government and the Government treats all the taxpayers in a respectful manner. Recently Robert discovered that citizens around him are evading taxes. For instance, an acquaintance of his did not declare the extra income earned. A shop owner did not provide a cash receipt. Robert did not report the tax evaders to the authorities. Hence, the tax evaders will keep all the money for themselves and pay no taxes to the Government. On a scale from 1 (“Completely Agree”) to 4 (“Completely Disagree”), please indicate how much you agree or disagree with Robert's decision?
Low-Trust Vignette
Robert lives in country X. The taxpayers in country X exhibit very low trust toward their Government for several reasons. First, substantial amount of tax money is wasted. The Government operates opaquely (not transparently) and makes inefficient use of the taxes, not providing the necessary public goods and services to the citizens. Second, not all taxpayers are equal against the law for the Government and the Government does not treat all the taxpayers in a respectful manner. There are certain privileged groups. Recently Robert discovered that citizens around him are evading taxes. For instance, an acquaintance of his did not declare the extra income earned. A shop owner did not provide a cash receipt. Robert did not report the tax evaders to the authorities. Hence, the tax evaders will keep all the money for themselves and pay no taxes to the Government. On a scale from 1 (“Completely Agree”) to 4 (“Completely Disagree”), please indicate how much you agree or disagree with Robert's decision?

Note: The vignettes used in the survey-experiment. In the vignette, where the portrayed third person is female, Robert is replaced with Anna.