Willing to share? Tax compliance and gender in Europe and America

John D'Attoma[1], Clara Volintiru[2,3] and Sven Steinmo[1,4]

Abstract
Studies examining the effects of gender on honesty, deceptive behavior, pro-sociality, and risk aversion, often find significant differences between men and women. The present study contributes to the debate by exploiting one of the largest tax compliance experiments to date in a highly controlled environment conducted in the United States, the United Kingdom, Sweden, and Italy. Our expectation was that the differences between men's and women's behavior would correlate broadly with the degree of gender equality in each country. Where social, political and cultural gender equality is greater we expected behavioral differences between men and women to be smaller. In contrast, our evidence reveals that women are significantly more compliant than men in all countries. Furthermore, these patterns are quite consistent across countries in our study. In other words, the difference between men's and women's behavior is not significantly different in more gender neutral countries than in more traditional societies.

Introduction

Gender equality is a top priority throughout the world and is often considered a basic feature of human rights. Gender inequality contributes to a large number of social ills, and the evidence is clear that improving gender equality improves standard of living, life expectancy, quality of life, amongst many other benefits. Improving the quality of women's lives and improving the opportunities available to them improves everyone's life.

Progress towards gender equality has certainly been achieved in many societies. But differences persist. Even in Scandinavian countries, where many of the structural and legal barriers against women have been reduced, we still find differences in pay, representation in private business boards, as well as the distribution of household and child-care work/effort.[1]

Institutionalists suggest that citizens in different societies should have different preferences based on their experiences in that society. Institutions and history matter, in this sense, because they can *shape preferences over time* (Steinmo and Thelen, 1992). We have attempted to disentangle the effects of institutions from other structural and social variables by running a simple tax compliance experiment in a set of countries: Sweden, UK, Italy and the US. In this paper we report the results from what we believe to be the largest cross-national tax compliance experiment conducted anywhere in the world.

We believe that this design allows us to explore the issue of gender behavior and gender differences that is relatively unique in both the experimental literature and the field of gender studies. Specifically, we are able to examine whether the differences between the behavior of men and women correspond to the extent of the structural barriers or limitations that women face in a given country. It is reasonable to assume that men's and women's behavior converges (i.e. men and women will behave in a similar way) in

[1]The Robert Schuman Centre for Advanced Studies, European University Institute, Florence, Italy
[2]Bucharest Academy of Economic Studies, Romania
[3]London School of Economics and Political Science, UK
[4]University of Colorado-Boulder, USA

Corresponding author:
John D'Attoma, The Robert Schuman Centre for Advanced Studies, European University Institute, Via dei Roccettini, 9, San Domenico di Fiesole (FI) 50014, Italy.
Email: john.dattoma@eui.eu

contexts where they are in fact treated more equally than they are in countries in which women still face large and significant discrimination. Sweden and Italy, for example, rank 4 and 16 on the gender inequality index, respectively, while the UK and US rank 28 and 43 respectively (Human Development Report, 2016).

Precisely because this experiment was conducted in four countries with over 1500 participants we are able to look at how particular incentives affect individual decisions across and within individual countries. Our study examines how individuals respond to different types of redistribution, different tax rates, and different tax structures. We are thus able to compare how people in different countries, and of different genders, respond to these different incentives and or disincentives (i.e. variations in the level of redistribution, tax rates or structure).

Our study was not initially, nor primarily, intended to study gender differences in tax compliance. Instead, this study was designed to explore how people in different societies would behave under similar conditions (aka institutions). We wanted to know, quite simply, if people in different societies were faced with exactly the same incentives and choices with respect to taxation, would they make different decisions? In this paper, however, we report the fact that men and women significantly differ in their willingness to comply with their taxes across countries and conditions. These differences are remarkably large and are consistent across a wide variety of institutional choices. Simply put, women appear to be much more tax compliant than men in every country and under every condition.

Literature review

There is an extensive body of literature addressing attitudinal and political differences between men and women. Virtually all of this work confirms that men and women have different policy attitudes and preferences. Women tend to be more left-leaning and more likely to support state intervention through the expansion of the welfare state (Inglehart and Norris, 2003; Morgan-Collins, 2013). They are also less likely to condone corruption (Barnes and Beaulieu, 2014; Dollar et al., 2001; Torgler and Valev, 2010), but less politically engaged (Atkeson, 2003; Burns et al., 2001; Lawless and Fox, 2010; Verba et al., 1997).

It is generally assumed that differences in attitudes will translate into difference in behavior. We know, for example, that women in the US are more likely to vote than men (Centre for American Women and Politics (CAWP), 2015). Also women in advanced industrial democracies increasingly tend to vote more to the Left than men (Inglehart and Norris, 2000). But, since behavior is structured by the social, political, and economic environment, it can be difficult to determine the extent to which individual behaviors are the product of individual preferences or that of socially constrained (and constructed) choices. In recent years,

experimental research has blossomed, partly in order to address this issue of identifying the main anchoring point of manifested behaviors.

The evidence regarding gender differences, however, is somewhat contradictory. Some studies have shown that men and women behave differently even when facing abstract choices. For example, women tend to be less competitive and less certain of the quality of their performance (Preece and Stoddard, 2015). Eckel and Grossman (2001) have demonstrated in a variety of experiments that women are more altruistic, but others have shown men to be more willing to contribute to the public good (1998, 2001, 2008; see also Bruner et al., 2017; Brown-Kruse and Hummels, 1993; Sell and Wilson, 1991; Solow and Kirkwood, 2002).

Men and women also appear to have different attitudes and behavior when it comes to taxation specifically. Surveys have shown that in contrast to men, women tend to think that the tax code is fairer, the likelihood of getting caught for evasion is greater, and they overestimate the penalties for evasion (Kinsey, 1992; Smith and Stalans, 1991). In terms of behavior, a number of tax compliance experiments have also shown women to be more compliant than men (Cadsby et al., 2004; Cadsby et al., 2006; Chung and Trivedi, 2003; Gërxhani, 2007; Hasseldine and Hite, 2002; Kastlunger et al., 2010; Lohse and Qari, 2014; Powell and Ansic, 1996; Spicer and Hero, 1985). However, most of these studies treat gender as a residual predictor.

We know that behavior is sensitive to context (Chermak and Krause, 2002; Seguino et al., 1996). As Sequino et al. pointed out, "our environment helps shape how we act and how we see others." And therefore they "suggest that social structures that shape our preferences may differ along gender lines" (p.14–15). Even in the lab, behavior is shaped by broader social norms. The inconsistent role gender plays in many of the existent experiments might be a product of differences in the experiments themselves, or a product of the differences in the social/political context in which the experiments were conducted. Our study attempts to answer this issue of contextual or treatment variations that might confound gender differences. As such, our study examined the gender differences manifested by subjects in the same fiscal compliance experiment, across different countries. We can therefore further the current research by controlling for the potential effects broader social/political contexts might bear on behavior.

Experimental overview

Our experiments were conducted at universities during the 2013/2014 and 2014/2015 academic years.[2] We took great care to assure that all subject pools were demographically very similar and that the selection mechanisms were nearly identical and unbiased (see Tables A1, A2, A3, and A4 in the online appendix). Each university uses an electronic database to which students, or past

Table 1. Summary of tax reporting rounds.

Stage	Task	Description
	Clerical 1	Earn income that is reported in Rounds 1 through 3
Stage 1	Round 1: No redistribution	Flat tax rate of 30% on all reported income
		Tax revenues are not redistributed
	Round 2: Redistribution	Flat tax rate of 30% on all reported income
		Tax revenues are collected into a common fund, which is redistributed on an equal per capita basis to all participants
	Round 3: Redistribution x 2	Flat tax rate of 30% on all reported income
		Tax revenues are collected into a common fund, the amount in the fund is doubled, and then redistributed on an equal per capita basis to all participants
	Clerical 2	Earn income that is reported in Rounds 4 through 6
Stage 2	Round 4: 10% tax rate	Flat tax rate of 10% on all reported income
		Tax revenues are collected into a common fund, the amount in the fund is doubled, and then redistributed on an equal per capita basis to all participants
	Round 5: 30% tax rate	Flat tax rate of 30% on all reported income
		Tax revenues are collected into a common fund, the amount in the fund is doubled, and then redistributed on an equal per capita basis to all participants
	Round 6: 50% tax rate	Flat tax rate of 50% on all reported income
		Tax revenues are collected into a common fund, the amount in the fund is doubled, and then redistributed on an equal per capita basis to all participants
	Clerical 3	Earn income that is reported in Rounds 7 through 9
Stage 3	Round 7: Progressive 1	Top 10% of earners in Clerical 3 pay 50% tax on reported income
		Bottom 10% of earners in Clerical 3 pay 10% tax on reported income
		Everyone else pays 30% tax on reported income
		Tax revenues are collected into a common fund, the amount in the fund is doubled, and then redistributed on an equal per capita basis to all participants.
	Round 8: Progressive 2	Participants pay tax of 10% on all reported income under 50 CU
		Participants pay tax of 30% on all reported income between 50 and 100 CU
		Participants pay tax of 50% on all reported income over 100 CU
		Tax revenues are collected into a common fund, the amount in the fund is doubled, and then redistributed on an equal per capita basis to all participants
	Round 9: Charity	Flat tax rate of 30% on all reported income
		Tax revenues are collected into a common fund, the amount in the fund is doubled, and then donated to charity

ECU:experimental currency unit.

students, voluntarily submit their information for participation in experiments. The participants were then randomly selected and invited by email to participate in the experiment (for more details on the Online Recruitment System for Economic Experiments (ORSEE), see Greiner, 2004). Once the participants arrived at the laboratory they were given an anonymized identification number and assigned to a partitioned computer to limit the interaction between themselves and other participants.[3] We linked participant pay to ID number thus ensuring complete anonymity.

The experiment consisted of three stages with three income reporting rounds within each stage.[4] In total, subjects report their income nine times for tax purposes. Subjects were given a different experimental parameter with each income reporting decision. In the first stage we varied the size of the general fund; in stage two we altered the tax rates; and in stage three we changed the tax structure. At the beginning of each stage, subjects were asked to perform a simple clerical task for which they received experimental currency (EC) that would be converted into real money at the end of the experiment. Subjects were paid 10 experimental currency units (ECU) for each correctly copied line of text which was then exchanged for domestic currency at the rate of 0.01 per token.

At the beginning of each income reporting round, subjects were given specific examples similar to the decisions that they would make in that particular round (see Table 1). Specifically, in stage one, the tax rate was 30% of reported income, the audit probability was 5%, and we varied the amount of redistribution. In round one (Round 1: no redistribution) there was no general fund, and thus no redistribution. In round two (Round 2: redistribution) the tax revenue was placed in a general fund and redistributed equally to all participants. For round three (Round 3: redistribution x 2) we doubled the general fund and divided it equally amongst all subjects. In stage two, the redistribution of the general

Table 2. Summary statistics.

Variable	N	Mean	SD	Min	Max
Y (Fraction of reported income for each subject in each round)	12,497	0.623	0.432	0	1
Italy	12,512	0.198	0.399	0	1
UK	12,512	0.230	0.421	0	1
US	12,512	0.362	0.481	0	1
Sweden	12,512	0.209	0.407	0	1
Female	12,344	0.496	0.500	0	1
Standardized risk	12,240	−0.004	0.998	−2.171	1.851
Economics major	12,208	0.205	0.404	0	1
Employed	12,328	0.384	0.486	0	1
Past-participation	12,328	0.715	0.451	0	1
Standardized income	12,512	−0.001	0.999	−3.179	5.189
Trust in Government	12,336	−0.001	0.569	−1.67	2.119
Duty to Pay	12,336	0.002	0.584	−2.866	1.222
Pro-redistribution	12,336	−0.006	0.619	−2.719	1.605
SVO angle	12,512	21.964	14.959	−16.260	61.389
Father's country of birth	12,512	0.566	0.496	0	1
Mother's country of birth	12,512	0.696	0.460	0	1

SVO: social value orientation.

fund remained constant, there was still a 5% chance of being audited, but we varied the tax rates in each round. In round four (Round 4: 10% tax rate) there was a 10% flat tax; for round five (Round 5: 30% tax rate) there was a 30% flat tax; and in round six (Round 6: 50% tax rate) there was a 50% flat tax. Lastly, in stage three the redistribution and audit rate remained constant, while we adjusted the tax structure. In round seven (Round 7: progressive 1) the top 10% of declared incomes payed a 50% tax rate; the bottom 10% of declared incomes payed a 10% tax rate; and everyone else payed a 30% rate. In round eight (round 8: progressive 2), all income over 100 ECU was taxed at a 50% rate; income between 50 and 100 ECU was taxed at a 30% rate; and all income below 50 ECU was taxed at a 10% rate.[5]

At the end of the tax compliance experiment, participants were asked to participate in a simple iterated dictator game designed by Ryan Murphy and Kurt Ackermann (2011) to assess one's level of prosociality. Once all experiments were complete, we asked the subjects to complete a fifteen-minute survey regarding certain demographic and attitudinal characteristics.

Altogether there were a total of 1564 subjects: 311 (Italy), 360 (UK), 566 (US), and 327 (Sweden). In our pool 50% were female, 38% were employed, and 21% were economics majors. The vast majority, 72%, of our subjects had participated in experiments before (see Table 2).

Methods and results

Given the literature and common expectations about the differences observed between men and women's behavior we draw the hypotheses:

H1: Tax compliance among women will be higher than men across rounds.

H2: The differences observed between men and women's responses should be smaller in countries that have achieved greater levels of legal and social equality.

First, we are interested in whether females comply more than males across countries and rounds. Figure 1 demonstrates prima facie evidence suggesting women are more compliant than men across countries. From Figure 2, we can observe that there are large gender differences across treatments. Although it is not the central component of this study, it should be noted, however, that the gender gap does vary between decisions. The gender gap decreases, for example, when we increase the return on the public good. Similarly, the gender gap increases when we increase tax rates. In Italy and Sweden, these differences are mainly being driven by changes in men's behavior. More specifically, men are more responsive to the incentives in each decision in Sweden and Italy, than in the US. Moreover, whereas women tend to respond less to the experimental treatment in Sweden and Italy, in the UK and US women are only slightly less reponsive to the treatment. For a more thorough analysis of the effects of specific treatments on gender in each country see Bruner et al. (2017).

We now examine how gender differences affect tax compliance within and across countries. We treat tax reporting as a single variable with distinctive values for subject and experimental period. We perform a series of Ordinary

Least Squares (OLS) analyses represented by the following equation:

OLS Regression

$$Y_{ij} = \alpha + \beta_1 female_i + \beta_2 Sweden_i + \beta_3 Sweden_i * $$
$$female + \beta_4 UK_i + \beta_5 UK * female_i + \beta_6 Italy_i + $$
$$\beta_7 Italy * female_i + \Theta X_i + \epsilon_{ij}$$

where:

Y_{ij}	=	the fraction of reported income by each participant i in each decision round j
$Female_i$	=	dummy variable for female subjects
$Sweden_i$	=	country dummy for participants who participated in the experiment in Sweden
$Sweden*female_i$	=	country interaction term for Sweden
UK_i	=	country dummy for participants in the UK
$UK*female_i$	=	country interaction term for UK
$Italy_i$	=	dummy variable for participants in Italy
$Italy*female_i$	=	country interaction term for Italy
X	=	vector of individual characteristics
E_{ij}	=	individual-specific error term, clustered at the individual level
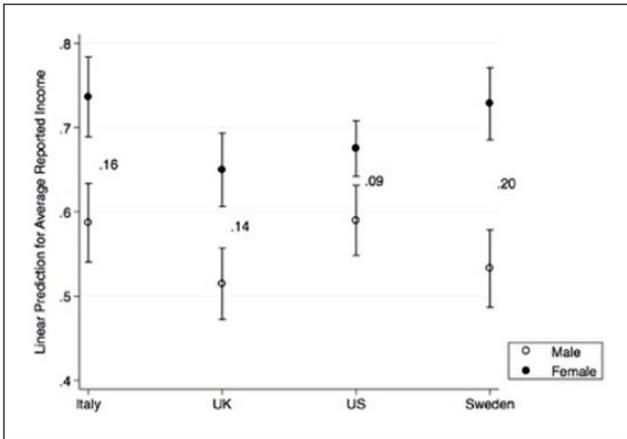
Figure 1. Average compliance rate by gender.

In Table 3 we present the results of our OLS analyses. Economics is a dummy variable for economics majors, and risk is an individual risk assessment measure. Past-participation is a dummy variable for whether participants have participated in experiments in the past, and employed is a variable for whether the subjects are employed. Each of these variables was taken from our attitudinal survey, completed at the end of the experiment. Income (standardized) is participants' income in the first eight rounds. Pro-redistribution, duty to pay, and trust are factors produced from an orthogonal rotated principal components factor analysis, also from the attitudinal survey (for more details about the factor analysis, see Pampel et al., 2017). Pro-redistribution

is a factor with self-placement on a left-right scale as the key component. Duty represents participants' sense of responsibility to the state, such as the extent to which cheating on one's taxes, cheating on government benefits, and not paying taxes for a variety of reasons are justifiable. Trust characterizes participants' level of confidence in government. SVO angle is measured from an iterated dictator game in which one person is asked to allocate their endowment to an unknown partner in the room as a test of altruism (for measurement of the SVO see Murphy and Ackerman (2011). Native-born mother and father are variables which control for the birthplace of the subjects' parents. Finally, we control for each individual reporting round.

In column 1 of Table 3, we estimate the effect of gender on tax compliance in our pooled-country dataset with a host of control variables. We determine that being female is statistically significant and positively correlated with tax compliance, generating a large effect. Economics majors, past-participation, and risk are all negatively correlated with tax compliance, whereas pro-redistribution, duty to pay, and SVO are all positive. In columns 2 to 5, we estimate the effect of gender on tax compliance in each individual country. The effects of our control variables do depend slightly on the country context, but the effect of being female is robust in each individual country. We determine that the effect of being female on tax compliance ranges from an 11% increase in the US to a 20% increase in Sweden, all else being equal.

Amongst our subjects we find moderate support for the idea that women are slightly more risk averse than men, but still we find that being female remains highly correlated with tax compliance even when controlling for risk acceptance, meaning that risk acceptance is not the variable driving these gender differences. The effect does soften some, but maintains a large effect on tax compliance. This result is especially important, because it establishes that women are profoundly more compliant even when their degree of risk acceptance is kept at an identical degree to their male cohorts. Confirming previous literature, risk acceptance is also negatively correlated with tax compliance. Female remains highly significant and positive, holding all else constant. In column 6, we are mainly interested in the interaction terms to examine the gender tax gap between countries, with the US as our baseline case. Here the results are unexpected: the gender tax gap in the US is significantly smaller than in Sweden at the .001 level. The gender tax gap is also significantly smaller in the US than in Italy and the UK, although the significance is weak.

Finally, in column 7, we examine our full model. Economics majors, income, willingness to accept risk, and past-participation are all statistically significant. The effect of participating in experiments in the past has a large effect, decreasing tax compliance by approximately 8% when holding all other variables at their mean. Moreover, our attitudinal variables such as support for the welfare state
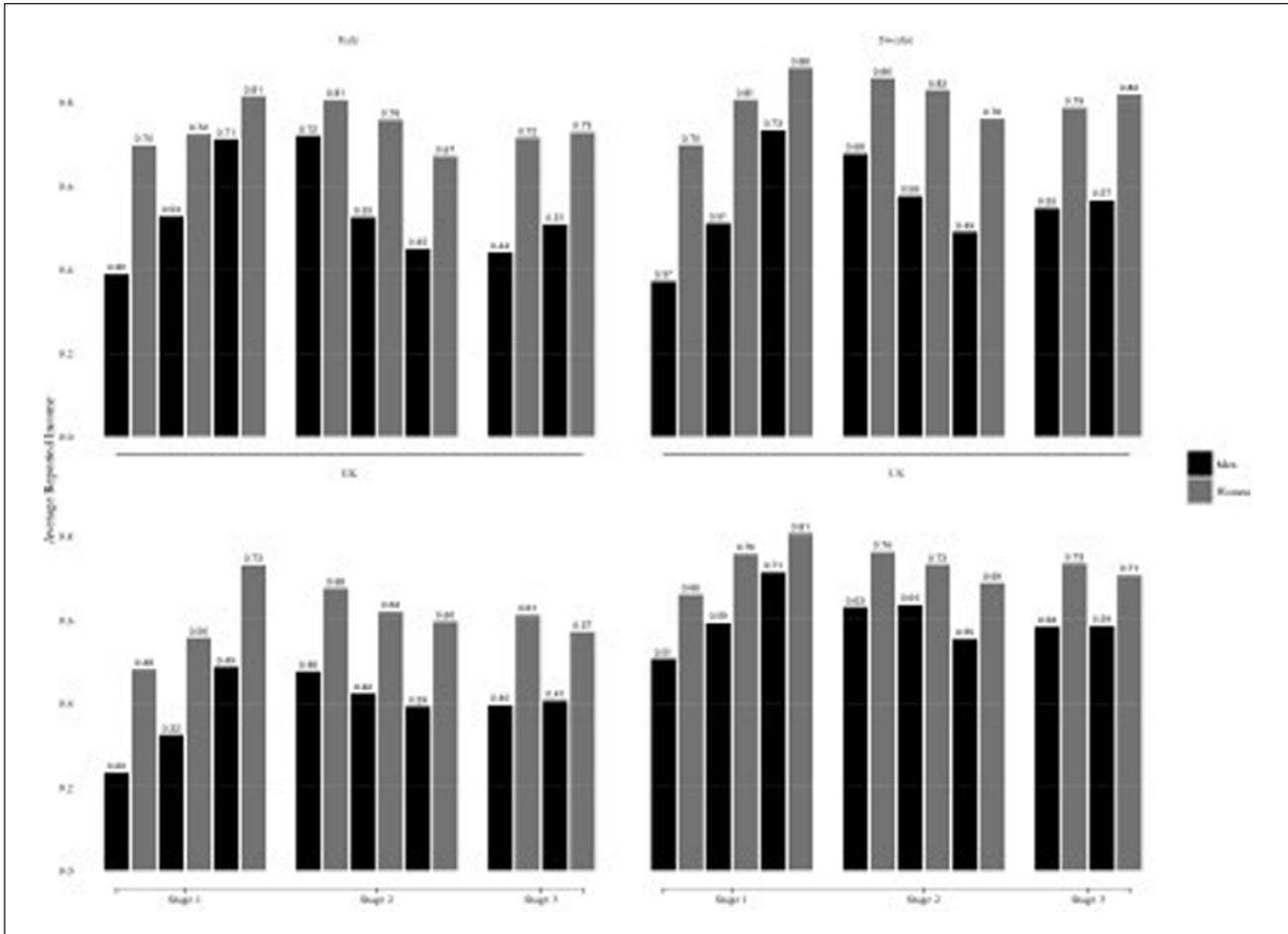
Figure 2. Average compliance rate by gender.
Stage 1: Variation in the level of redistribution; Stage 2: Variation in the tax rate; Stage 3: Variation in the tax structure. Within each stage there were three experimental treatments (see Table 1).

and duty to pay taxes are also statistically significant. Trust in authority is significant, but not in the expected direction. This could largely be due to the relatively small variation across countries in tax compliance, combined with significant variation in trust in government. Swedes, for example, demonstrate high trust, whereas Italians demonstrate low trust. Furthermore, having native-born parents has no effect on our model.[6]

Against our expectations, we discover that Sweden, a country which has achieved one of the highest levels of gender equality in the world, demonstrates the largest tax compliance gap, and that gap is statistically greater than the US, the country with the largest level of gender inequality, according to the gender inequality index.[7] In fact, Sweden is the only country with a gender gap that is significantly larger than the gap in the US. Figure 3 displays the predicted probabilities from column 7 of Table 3 for the compliance rate by gender in each country. What stands out from the figure is the fact that the gender gap is quite large in all countries, but especially large in Sweden, the UK, and Italy.
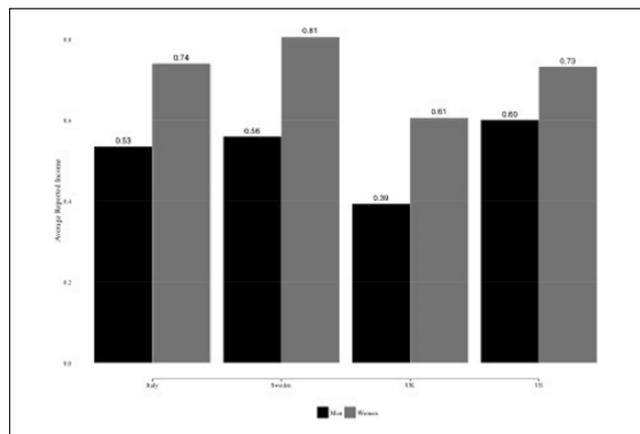
Figure 3. Predicted probabilities for the compliance rate.

Robustness checks

In this section we examine the robustness of our results. Specifically, we are concerned with whether our results are robust to treatment order. We ran an additional six sessions in

Table 3. Ordinary least squares regression for average compliance rate.

Variables	Pooled	Italy	UK	US	Sweden	Interactions	Full Model
Female	0.133***	0.150***	0.124***	0.092***	0.188***	0.131***	0.086***
	(0.016)	(0.033)	(0.033)	(0.027)	(0.032)	(0.028)	(0.027)
Italy	0.031					−0.066*	−0.003
	(0.024)					(0.034)	(0.033)
UK	−0.050**					−0.208***	−0.077**
	(0.022)					(0.033)	(0.031)
Sweden	−0.006					−0.041	−0.060*
	(0.026)					(0.035)	(0.034)
Economics major	−0.052**	−0.050	−0.078**	−0.023	−0.027		−0.052**
	(0.021)	(0.035)	(0.040)	(0.047)	(0.047)		(0.021)
Risk (standardized)	−0.045***	−0.034**	−0.068***	−0.046***	−0.033*		−0.045***
	(0.008)	(0.017)	(0.018)	(0.013)	(0.018)		(0.008)
Income (standardized)	−0.031***	−0.013	−0.065***	−0.021	−0.027		−0.031***
	(0.008)	(0.015)	(0.015)	(0.014)	(0.020)		(0.008)
Past-participation	−0.077***	−0.039	−0.128***	−0.078***	−0.047		−0.076***
	(0.016)	(0.043)	(0.037)	(0.025)	(0.036)		(0.016)
Employed	0.007	0.023	0.020	−0.005	−0.013		0.008
	(0.016)	(0.041)	(0.034)	(0.025)	(0.030)		(0.016)
Pro-redistribution	0.045***	0.029	0.036	0.015	0.090***		0.045***
	(0.014)	(0.033)	(0.032)	(0.024)	(0.031)		(0.014)
Duty to pay	0.070***	0.115***	0.078**	0.032	0.077**		0.070***
	(0.015)	(0.032)	(0.031)	(0.025)	(0.034)		(0.015)
Trust in government	−0.044***	−0.062*	−0.032	−0.059**	−0.043		−0.046***
	(0.015)	(0.031)	(0.033)	(0.027)	(0.035)		(0.016)
SVO angle	0.007***	0.006***	0.007***	0.007***	0.009***		0.007***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)		(0.001)
Father's birth place	0.002	0.028	−0.034	−0.001	−0.010		−0.001
	(0.022)	(0.054)	(0.039)	(0.041)	(0.054)		(0.021)
Mother's birth place	−0.015	−0.084	0.008	0.029	−0.058*		−0.011
	(0.020)	(0.063)	(0.037)	(0.042)	(0.033)		(0.020)
Redistribution	0.097***	0.085***	0.083***	0.093***	0.130***	0.096***	0.097***
	(0.010)	(0.023)	(0.019)	(0.017)	(0.022)	(0.010)	(0.010)
Redistribution x 2	0.229***	0.228***	0.259***	0.173***	0.288***	0.226***	0.229***
	(0.011)	(0.025)	(0.024)	(0.018)	(0.024)	(0.011)	(0.011)
10% Tax	0.206***	0.236***	0.258***	0.120***	0.259***	0.189***	0.205***
	(0.012)	(0.026)	(0.026)	(0.021)	(0.028)	(0.012)	(0.012)
30% Tax	0.146***	0.105***	0.201***	0.106***	0.188***	0.130***	0.146***
	(0.011)	(0.023)	(0.023)	(0.018)	(0.026)	(0.010)	(0.011)
50% Tax	0.083***	0.022	0.169***	0.046**	0.106***	0.069***	0.083***
	(0.012)	(0.025)	(0.024)	(0.020)	(0.025)	(0.011)	(0.012)
Progressive 1	0.114***	0.037*	0.188***	0.085***	0.153***	0.098***	0.114***
	(0.011)	(0.022)	(0.022)	(0.019)	(0.027)	(0.010)	(0.011)
Progressive 2	0.120***	0.082***	0.176***	0.069***	0.179***	0.103***	0.120***
	(0.011)	(0.023)	(0.021)	(0.018)	(0.025)	(0.010)	(0.011)
Female * Italy						0.082*	0.049
						(0.045)	(0.040)
Female * UK						0.075*	0.062
						(0.045)	(0.042)
Female * Sweden						0.116***	0.107***
						(0.044)	(0.040)
Constant	0.355***	0.420***	0.337***	0.393***	0.247***	0.486***	0.379***
	(0.030)	(0.070)	(0.052)	(0.044)	(0.061)	(0.024)	(0.032)
N	11849	2352	2760	4172	2565	12129	11849
r2	0.212	0.199	0.243	0.136	0.272	0.104	0.215

SVO: social value orientation.
Robust standard errors in parentheses.
***p<0.01, **p<0.05, *p<0.1.

Table 4. Difference in means t-test for gender by round in treatment order B.

	Gender	2 x Redistribution	T	Single Distribution	T	No Redistribution	Means Test
Stage 1			−3.46***		−3.073***		−2.839***
	Male (54)	0.6		0.63		0.549	
	Female (66)	0.833		0.83		0.746	
	Gender	**50% Tax**	**T**	**30% Tax**	**T**	**10% Tax**	**Means test**
Stage 2	Male	0.61	−4.226***	0.655	−3.125***	0.747	−1.540
	Female	.87		0.848		0.843	
	Gender	**Progressive 1**	**T**	**Progressive 2**	**T**	**Charity**	**Means test**
Stage 3	Male	0.657	−2.553***	0.7	−2.618***	0.616	−4.388***
	Female	0.84		0.86		0.892	

Notes: The t-test for equal is reported in the third, fifth, and seventh column with statistical significance indicated by asterisks: *** indicates the difference is significant at the 1% level.

Italy varying the treatments. First, we ran a series of difference in means t-tests in each round to test if there are gender differences in compliance in each round for treatment order B (see Table 4). The difference in means t-test suggests that there are significant differences between men and women in each round, with the exception of the 10% tax rate round.

Finally, we estimated an OLS with clustered standard errors and dummy variables for rounds 2 through 8 (see Table 5). The coefficients for the dummy variables show the increase or decrease in compliance relative to the first round, which serves as the reference. Indeed, we still uncover a large gender gap on average in treatment order B.

Discussion and conclusion

The first and obvious result drawn from this experiment is that women behave differently from men in all conditions and in all countries studied here. Interestingly, the "gender gap" differs greatly between the US and Sweden, and in a direction that was completely unexpected when embarking on this research. Whereas we expected the gender gap to be substantially smaller in more gender egalitarian countries, we find instead that gender is a powerful and robust variable across societies.

Although gender differences in tax compliance have been reported in previous articles (see more recently in Brockman et al., 2016), these works generally treat gender as a residual predictor. The present study demonstrates clearly and systematically why this is a major omission: gender is strongly associated with an individual's tax behavior, when controlling for institutional, cultural, and social factors. Our experiments show substantial gender variation in tax compliance across the four countries, given varying levels of social and legal gender equality.

Indeed, the original assumption that gender differences should be substantially minimized in more gender neutral societies assumed a temporal connection between norms, institutions and behavior that may be unrealistic. Norms are

Table 5. Regression analysis for compliance in treatment B.

Variables	Model 1
Female	0.186***
	(0.051)
Econ major	−0.066
	(0.076)
Risk	−0.003
	(0.012)
Income	−0.001
	(0.002)
Past-participation	−0.150***
	(0.057)
Employed	−0.076
	(0.060)
Pro-redistribution	−0.095
	(0.058)
Duty to pay	−0.050
	(0.042)
Trust in government	−0.016
	(0.051)
Redistribution	0.020
	(0.030)
Redistribution x 2	−0.072*
	(0.038)
10% Tax	0.021
	(0.033)
30% Tax	0.037
	(0.032)
50% Tax	0.079**
	(0.040)
Progressive 1	0.043
	(0.038)
Progressive 2	0.074**
	(0.033)
Constant	0.873***
	(0.165)
N	800
r2	0.170

Robust standard errors in parentheses.
***p<0.01, **p<0.05, *p<0.1.

sticky. This, we believe, can help explain why what Preece calls "gendered psyches" continue to have strong effects on behavior that may not correspond immediately to recently effected institutional changes. As Inglehart and Norris (2003: 79) noted "structural developments lead to, and interact with, cultural shifts that tend to reshape political values." Thus, we should not be so surprised to see that behavioral change lags behind the institutional change.

Our findings thus call into question the liberal assumption that given equal conditions, individuals behave in the same ways. Instead, our research demonstrates, even where structural reforms have actively targeted and effectively reduced legal and formal gender gaps, the behavioral differences between males and females persist. We find only weak evidence of intrinsic behavioral differences, such as attitudes towards risk, and as such our study strengthens the argument on variation in gendered perceptions (Fox and Lawless, 2011; Lawless and Fox, 2010.

In sum, our study uncovers significant gender variation in tax compliance across tax conditions and countries. Furthermore, this study makes the claim that albeit more social and legal gender equality, large behavioral differences between genders still persist – even in Sweden. The results do leave room for alternative interpretations and require further study. We welcome and encourage scholars to utilize our data in combination with other studies to further our understanding of these behavioral differences.

Declaration of conflicting interest

The authors declare that there is no conflict of interest.

Supplementary materials

All data can be downloaded from www.willingtopay.eu.

Notes

1. The World Economic Forum Global Gender Gap rankings place the Nordic countries as some of the most gender equal countries in the world (Hausmann et al., 2009).
2. The experimental sites included Bologna Laboratory for Experiments in Social Sciences, Centro d'Economia Sperimentale A Roma Est, and Experimental Economics Lab of the University of Milano Bicocca in Italy; Oxford Experimental Laboratory, Experimental Economics Laboratory–Royal Holloway in London, Finance and Economics Experimental Laboratory at Exeter, and ESSEXLab at Essex in the UK; Learning and Experimental Economics Projects at University of California–Santa Cruz, Social Science Experiments Lab at University of Colorado–Boulder, Appalachian Experimental Economics Laboratory in Boone, North Carolina, Center for Behavioral Political Economy in Stony Brook, New York, University of Hawaii Laboratory for Computer-Mediated Experiments and the Study of Culture in Honolulu, Hawaii, in the US; the Behavioural Lab in Stockholm and Behavioural and Experimental Economics in Gothenburg, Sweden.
3. The experiments were conducted in z-tree and a newer web-based experimental software called Bahavery (see Fischbacher, 2007).
4. For the purpose of this study, we chose to use only the first eight rounds of the experiments.
5. Round 9 was a charity round in which all tax revenue is donated to a specific charity. We have chosen not to include round 9 in the analysis, because it is not necessarily tax, but rather a donation. However, we did run preliminary analysis including the charity round, and it does not affect our results.
6. We ran three robustness tests on these data. We examined the intensity of evasion by gender using an ordered probit model on a variable which represents those that report 100%, a variable which measures those who evaded 100%, and another which represents those who report something in between. We also ran the above models in each individual round. First, men tend to cheat 100% significantly more than women, and women are 100% compliant significantly more than men. Furthermore, women are more likely to cheat in small amounts. We also ran the above model in each individual round, with very little change in the overall results. Finally, we ran our full model for only the first round of the data to eliminate any confound from potential learning/order effects. The main results still hold if we run the model for only round 1.
7. A Wald test confirms that the gender gaps between the U.K., Sweden, and Italy are insignificant.

References

Andrighetto G, Zhang N, Ottone S, et al. (2016) Are some countries more honest than others? Evidence from a tax compliance experiment in Sweden and Italy. *Frontiers in Psychology* 7: 472. DOI:10.3389/fpsyg.2016.00472

Atkeson L (2003) Not all cues are created equal: The conditional impact of female candidates on political engagement. *The Journal of Politics* 65(4): 1041–61.

Barnes T and Beaulieu E (2014) Gender stereotypes and corruption: How candidates affect perceptions of election fraud. *Politics and Gender* 10(3): 365–391.

Brockmann H, Philipp G and Seelkopf L (2016) Happy taxation: Increasing tax compliance through positive rewards? *Journal of Public Policy* 36(3): 1–26.

Brown-Kruse J and Hummels D (1993) Gender effects in laboratory public goods contribution: Do individuals put their money where their mouth is? *Journal of Economic Behavior and Organization* 22: 255–267.

Bruner D, D'Attoma J and Steinmo S (2017) Going Dutch? *The Role of Gender in the Provision of Public Goods through Tax Compliance*. Working Paper.

Burns N, Schlozman K and Verba S (2001) *The Private Roots of Public Action: Gender, Equality, and Political Participation*. Cambridge, MA: Harvard University Press.

Cadsby CB, Maynes E and Trivedi VU (2006) Tax compliance and obedience to authority at home and in the lab: A new experimental approach. *Experimental Economics* 9(4): 343–359.

Center for American Women and Politics (CAWP) (2015) *Gender Differences in Voter Turnout*. Available at: http://www.cawp.rutgers.edu/sites/default/files/resources/genderdiff.pdf (accessed 25 October 2016).

Chermak JM and Krause K (2002) Individual response, information, and intergenerational common pool problems. *Journal of Environmental Economics and Management* 43(1): 47–70.

Chung J and Trivedi VU (2003) The effect of friendly persuasion and gender on tax compliance behavior. *Journal of Business Ethics* 47(2): 133–145.

Dollar D, Fisman R and Gatti R (2000) Are women really the 'fairer' sex? Corruption and women in government. *Journal of Economic Behavior and Organization* 46(4): 423–429.

Eckel CC and Grossman PJ (1998) Are women less selfish than men?: Evidence from dictator experiments. *The Economic Journal* 108(448): 726–735.

Eckel CC and Grossman PJ (2001) Chivalry and solidarity in ultimatum games. *Economic Inquiry* 39(2): 171–188.

Eckel CC and Grossman PJ (2008) Men, women and risk aversion: Experimental evidence. *Handbook of Experimental Economics Results* 1: 1061–1073.

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

Fox RL and Lawless JL (2011) Gendered perceptions and political candidacies: A central barrier to women's equality in electoral politics. *American Journal of Political Science* 55(1): 59–73.

Gërxhani K (2007) Explaining gender differences in tax evasion: The case of Tirana, Albania. *Feminist Economics* 13(2): 119–155.

Greiner B (2004) The online recruitment system orsee 2.0. - a guide for the organization of experiments in economics. *University of Cologne, Working Paper Series in Economics* 10(23): 63–104.

Hasseldine J and Hite PA (2002) Framing, gender and tax compliance. *Journal of Economic Psychology* 24(4): 517–533.

Hausmann R, Zahidi S, Tyson L, et al. (2009) The global gender gap report 2009. *World Economic Forum*.

Inglehart R and Norris P (2003) *Rising tide: Gender equality and cultural change around the world*. Cambridge, UK: Cambridge University Press.

Kastlunger B, Dressler SG, Kirchler E, et al. (2010) Sex differences in tax compliance: Differentiating between demographic sex, gender-role orientation, and prenatal masculinization (2D:4D). *Journal of Economic Psychology* 31(4): 542–552.

Kinsey KA (1991) *Deterrence and alienation effects of IRS enforcement: An analysis of survey data Issue 9102*. American Bar Foundation.

Lawless JL and Fox RL (2010) *It Still Takes a Candidate: Why Women Don't Run for Office*. New York, NY: Cambridge University Press.

Lohse T and Qari S (2014) Gender differences in deception behaviour – The role of the counterpart. *Applied Economics Letters* 21(10): 702–705.

Morgan-Collins M (2013) Universal suffrage and the rise of socialist parties: Evidence from eleven European countries. In: *American Political Science Association 2013 annual meeting paper*. Available at: https://www.researchgate.net/publication/256053534_Universal_Suffrage_and_the_Rise_of_Socialist_Parties_Evidence_from_Eleven_European_Countries

Murphy RO, Ackermann KA and Handgraaf M (2011) Measuring social value orientation. *Judgment and Decision Making* 6(8): 771–781.

Pampel F, Andrighetto G and Steinmo S (forthcoming) Culture, institutions, and tax compliance: A cross-national experiment and survey, Working Paper.

Powell M and Ansic D (1996) Gender differences in risk behaviour in financial decision-making: An experimental analysis. *Journal of Economic Psychology* 18(6): 605–628.

Preece J and Stoddard O (2015) Why women don't run: Experimental evidence on gender differences in political competition aversion. *Journal of Economic Behavior & Organization* 117: 296–308.

Rothstein B (2011) *The Quality of Government: Corruption, Social Trust, and Inequality in International Perspective*. Chicago, IL: University of Chicago Press.

Schlozman KL, Burns N and Verba S (1994) Gender and the pathways to participation: The role of resources. *The Journal of Politics* 56(4): 963–90.

Seguino S, Stevens T and Lutz M (1996) Gender and cooperative behavior: Economic man rides alone. *Feminist Economics* 2(1): 1–21.

Sell J and Wilson RK (1991) Levels of information and contributions to public goods. *Social Forces* 70(1): 107.

Sell J, Griffith WI and Wilson RK (1993) Are women more cooperative than men in social dilemmas? *Social Psychology Quarterly* 56(3): 211–222.

Smith K and Stalans L (1991) Encouraging tax compliance with positive incentives: a conceptual framework and research directions. *Law and Policy* 13: 35–53.

Smith M (2012) Social regulation of the gender pay gap in the EU. *European Journal of Industrial Relations* 18(4): 365–380.

Solow JL and Kirkwood N (2002) Group identity and gender in public goods experiments. *Journal of Economic Behavior and Organization* 48(91): 403–412.

Spicer MW and Hero RE (1985) Tax evasion and heuristics: A research note. *Journal of Public Economics* 26(2): 263–267.

Steinmo S and Thelen K (1992) *Structuring Politics: Historical Institutionalism in Comparative Analysis*. New York, NY: Cambridge University Press.

Torgler B and Valev NT (2010) Gender and public attitudes toward corruption and tax evasion. *Contemporary Economic Policy* 28(4): 554–568.

United Nations Development Programme (2016) *The Human Development Report*. New York: Human Development Report.

Verba S, Burns N and Schlozman KL (1997) Knowing and caring about politics: Gender and political engagement. *The Journal of Politics* 59(04): 1051–1072.

Zhang N, Andrighetto G, Ottone S, et al. (2016) Willing to pay? Tax compliance in Britain and Italy: An experimental analysis. *PLoS One* 11(2): e0150277.