The backfiring effect of auditing on tax compliance

Juan P. Mendoza [a,*], Jacco L. Wielhouwer [a], Erich Kirchler [b]

[a] Department of Accounting, VU Amsterdam, The Netherlands
[b] Faculty of Psychology, University of Vienna, Austria

ARTICLE INFO

ABSTRACT

Using country-level data from 2003–2014, we examine the association between auditing level (measured as number of verification actions taken by tax authorities per 100 taxpayers in each country) and tax compliance (measured as business executives' perception of tax evasion). Our hypothesis is that compliance increases until a certain auditing level is reached, and decreases beyond that level (i.e., an elevated auditing level backfires). In line with our expectation, the results of a series of tests indicate that there is a U-shaped association between auditing and tax evasion. We discuss how a potential backfiring effect may depend on the extent to which compliance is voluntary.

© 2017 Elsevier B.V. All rights reserved.

1. Introduction

Early theories of tax enforcement suggest that higher auditing levels lead to higher levels of tax compliance (Allingham & Sandmo, 1972; Srinivasan, 1973). During the last decades, however, empirical research has provided mixed evidence for this association (Beer, Kasper, Kirchler, & Erard, 2015; Gangl, Torgler, Kirchler, & Hofmann, 2014; Iyer, Reckers, & Sanders, 2010; Slemrod, Blumenthal, & Christian, 2001; Spicer & Thomas, 1982). As a result, subsequent theoretical work has examined why auditing can have positive, negative, or null effects on tax compliance (Alm, 2012; Devos, 2014; Feld & Frey, 2007; Kirchler, Hoelzl, & Wahl, 2008; Lederman, 2016; Torgler, 2002).

Despite these analyses and several cross-country examinations (e.g., Frey & Torgler, 2007; Fuest & Riedel, 2009; Richardson, 2006; Tsakumis, Curatola, & Porcano, 2007), little is known about the association between auditing and tax compliance at the country-level of analysis – the level at which tax variables of great significance are predominantly set (e.g., central government tax rates, auditing procedures, penalties for noncompliance). This study aims at filling this gap, and responds to recent calls from multilateral organizations for research on the effectiveness of audit strategies employed in practice (IMF, 2015; OECD, 2010).

* Corresponding author.
  E-mail addresses: j.mendozarodriguez@vu.nl (J.P. Mendoza), j.l.wielhouwer@vu.nl (J.L. Wielhouwer), erich.kirchler@univie.ac.at (E. Kirchler).

Building on behavioral theories of tax compliance, we hypothesize that the association between auditing and tax compliance is non-linear. Our expectation is that compliance increases until a certain auditing level is reached, and decreases beyond that level (i.e., an elevated auditing level backfires). Put differently, we expect a U-shaped association between the auditing level and noncompliance. The rationale is that, while holding everything else constant, an elevated auditing level may signal distrust in taxpayers and lead to the perception that the tax authority and its enforcement actions are excessive and unfair (Feld & Frey, 2002; Kirchler et al., 2008; Lawsky, 2008; Mooijman, van Dijk, van Dijk, & Ellemers, in press; Wahl, Kastlunger, & Kirchler, 2010; Wenzel, 2003). Perceived unfairness and feelings of distrust may in turn lead to reduced compliance (Hofmann, Hoelzl, & Kirchler, 2008; Kirchler, 2007; Murphy, 2005; Torgler, 2007). This goes in line with prior research suggesting that control in general, and crime deterrence in particular, can backfire when these are indicative of distrust and unfairness (Falk & Kosfeld, 2006; Kessler & Leider, 2016; Schildberg-Hörisch & Strassmair, 2012).

Using publicly available country-level data from 2003–2014, we examine the association between the auditing level, which is measured as the number of verification actions taken by tax authorities per 100 taxpayers in each country (OECD, 2017a), and business executives' perception of tax evasion (IMD, 2017a; World Bank, 2017b). Following prior research in this area, as control variables we include the penalty for tax offenses (as a percentage of unpaid taxes), short-term interest rate (which is typically imposed on unpaid taxes; Andreoni, Erard, & Feinstein, 1998), personal and corporate income tax rates, GDP per capita (Fuest & Riedel, 2009), and trust in government (as indicated by perceptions of effectiveness and transparency of government policy; IMD, 2017a; Kirchler et al., 2008). The empirical analysis employs a stringent test for U-shaped associations, includes a series of robustness tests, and takes into consideration possible endogeneity and reverse causality.

In line with our expectation, the results indicate that there is a U-shaped association between auditing and noncompliance. Tax evasion decreases before a certain auditing level is reached, and increases beyond that level. This supports the notion that control can backfire (Falk & Kosfeld, 2006), and may shed light on why researchers have at times observed contradictory or null effects of auditing on compliance. This backfiring effect has important implications for the design and implementation of tax enforcement strategies. The core implication is that more auditing may not be necessarily better. We elaborate more on theoretical and practical implications in the final section, and discuss the feasibility of alternative explanations based on data availability and variable measurement limitations.

2. Theory and hypothesis

Early theories of tax compliance build on the economic theory of crime deterrence (Becker, 1968), which predicts a positive association between auditing and compliance (Allingham & Sandmo, 1972; Sandmo, 2005; Srinivasan, 1973). The rationale is that auditing increases expected costs associated with getting caught not complying, and thus reduces the incentive to evade taxes.

Behavioral theories distinguish between enforced and voluntary compliance (e.g., Braithwaite, 2007; Kirchler et al., 2008; Slemrod, 1998). Enforced compliance is primarily explained by deterrence and enforcement actions, whereas voluntary compliance is primarily explained by other factors besides deterrence (e.g., trust in the authority, perceptions of procedural fairness, social norms).

Interestingly, auditing can simultaneously influence both enforced and voluntary compliance (Kirchler et al., 2008; Muehlbacher, Kirchler, & Schwarzenberger, 2011; Wahl et al., 2010). By definition, economic deterrence increases with auditing, so auditing has a positive effect on enforced compliance (for a discussion on the magnitude of this effect, see Alm, 2012; Andreoni et al., 1998). In contrast, the expected overall effect of auditing on voluntary compliance is less straightforward. As the number of audits increases from zero, auditing signals efforts to detect and control evasion, leads to perceptions of trustworthiness and fairness about the authority and its enforcement actions, and ultimately enhances voluntary compliance (Kirchler et al., 2008). As the number of audits becomes *exceedingly* larger, however, auditing can have an unintended (backfiring) effect, as we explain next.

2.1. The backfiring effect of an elevated auditing level

2.1.1. Distrust

Auditing may signal distrust when audits "occur very frequently or in an inquisitorial style" (Kirchler et al., 2008, p. 213), are not well-targeted at tax evaders, and appear to disregard the efforts and actions of honest taxpayers (Feld & Frey, 2007; Kirchler, Muehlbacher, Kastlunger, & Wahl, 2010). Indeed, "if a revenue body's approach is perceived as very controlling, it can cause taxpayers to feel distrusted" (OECD, 2010, p. 6).

Stronger signs of distrust increase the social and psychological distance between tax authorities and taxpayers, and make their relationship more antagonistic (Braithwaite, 2003a; Braithwaite, Murphy, & Reinhart, 2007). An antagonistic tax climate reflects a "game of cops vs. robbers" in which taxpayers are seen as evaders who require constant and intensive supervision, and tax authorities are seen as distrustful enforcement officers or prosecutors (Feld & Frey, 2002; Kirchler et al., 2008). Assuming that taxpayers reciprocate the tax authority's attitudes and behaviors (Bazart & Bonein, 2014; Smith & Stalans, 1991), feelings of distrust may trigger reciprocal distrust in the tax authority (Feld & Frey, 2002; Kirchler et al., 2008; Levi, 1998; Torgler, 2003; Torgler, 2004; Wahl et al., 2010). A tax climate characterized by antagonism and mutual distrust can

foster resistance (a posture characterized by doubts about the authority' intentions to be cooperative and benign; Hollander & Einwohner, 2004). Resistant taxpayers act in their own interest and make efforts to object, oppose, and fight the authority, its rules and decisions (Braithwaite, 2003b; Murphy, 2004).

Moreover, auditing can be interpreted as external control that *forces* firms and individuals to comply (Feld & Frey, 2007; Frey & Jegen, 2001). Under this view, an elevated auditing level makes internal control redundant or unnecessary, and consequently "crowds out" taxpayers' internal motivation to pay taxes on a voluntary basis. In addition, auditing can be seen as a restriction to taxpayers' freedom, and consequently cause reactance (a reaction that is contrary to what the authority intends; Braithwaite, 2003b; Kirchler, 1999; Murphy, 2008).

2.1.2. Perceptions of unfairness

An elevated auditing level can also make taxpayers perceive that the tax authority and its enforcement actions are in different aspects unfair. Indeed, an elevated auditing level can be perceived as excessively or unnecessarily costly, intrusive, inquisitive, or time-consuming (Kirchler et al., 2008; Lawsky, 2008; Stalans & Lind, 1997; Wenzel, 2003). These negative perceptions may relate to distributive fairness (e.g., fairness in the assignment and exchange of audit-related efforts, costs, and benefits), procedural fairness (e.g., fairness in the application of audit rules and procedures, provision of supportive and respectful treatment, and acknowledgment of taxpayers' voice and opinions), or retributive fairness (e.g., appropriateness of investigation procedures; Hofmann et al., 2008; Wenzel, 2003). Overall, taxpayers may perceive that the audit process is unfair when it is "too rigid, ignorant of possible harm to the person or their business, or a mere harassment" (Wenzel, 2003, p. 58). Perceptions of unfairness may in turn trigger defiance (an attitude in which taxpayers openly refuse to obey the authority), and as in the above-mentioned case of distrust, these perceptions may reduce intrinsic motivations to pay taxes on a voluntary basis (Braithwaite, 2012; Hartner, Rechberger, Kirchler, & Schabmann, 2008; Sherman, 1993; Torgler, 2003).

2.1.3. Noncompliance as a social norm

By experiencing, observing, or reading about an elevated auditing level, taxpayers may infer that noncompliance is pervasive, socially accepted, and therefore justifiable behavior (Sheffrin & Triest, 1991; Wenzel, 2004, 2005). In this sense, an elevated auditing level may have the unintended effect of reinforcing noncompliance as a social norm. The rationale is that taxpayers are less likely to comply when they think that others are not complying (Bobek, Roberts, & Sweeney, 2007; Cullis & Lewis, 1997; Posner, 2000).

2.2. Hypothesis

In sum, an elevated auditing level can lead to perceptions of distrust, unfairness, or widespread noncompliance, which enhance negative tax attitudes and behaviors (e.g., defiance, resistance, reactance; Braithwaite, 2003b), and reduce voluntary compliance (Frey, 1994; Hofmann et al., 2008; Kirchler et al., 2008; Levi, 1998; Murphy, 2005; Torgler, 2007). Taking these distinct yet complementary lines of argument into consideration, we hypothesize that the effect of auditing on compliance is positive until a certain auditing level is reached, and negative beyond that level (i.e., an elevated auditing level backfires). Empirically, this hypothesis implies that there is a U-shaped association between the auditing level and noncompliance.

It is important to note, however, that this hypothesis applies to auditing levels observed in practice, which are significantly lower than the theoretical maximum (i.e., a level in which all taxpayers are audited; Torgler, 2002). In reality, tax authorities face various constraints that prevent them from introducing the maximum auditing level (constraints related to, e.g., budget, time, information, technology, or personnel; Andreoni et al., 1998; Graetz, Reinganum, & Wilde, 1986). In theory, however, an authority without constraints may introduce levels close to the maximum and ultimately push taxpayers towards absolute compliance.

3. Method

3.1. Variable measurement

In order to examine the association between the auditing level and tax compliance, we gather country-level data from independent sources, including the OECD, World Bank, and IMF. The data are publicly available, and cover the period 2003–2014. Instructions to access the dataset and conduct the analysis are presented in the Appendix.

The dependent variable is tax evasion, and the independent variable is the auditing level. The control variables are the penalty for tax offenses, short-term interest rate, personal and corporate income tax rates (i.e., PIT and CIT), GDP per capita, and trust in government.

3.1.1. Tax evasion

This measure corresponds to business executives' perception of the extent to which tax evasion is common practice in their country (0 = it is common practice, 10 = it is not common practice). This measure is gathered on a yearly basis by the Swiss Institute for Management Development (IMD, 2017a), and is made publicly available by the World Bank

(World Bank, 2017b). Respondents are top and middle management executives. We reverse score this measure so that larger values represent higher levels of tax evasion.

3.1.2. Auditing level

The auditing level is calculated based on two measures published by the OECD in the Tax Administration Comparative Reports (2017a). The first measure is the number of verification actions, which corresponds to "all of the activities typically undertaken by revenue bodies to check whether taxpayers have properly reported their tax liabilities in the returns filed by them" (OECD, 2013, p. 212). These actions may also be referred to as audits, controls, examinations, investigations, or inquiries; and may include field, desk, or correspondence actions. The second measure that we employ is taxpayer population, which corresponds to the total number of registered taxpayers for personal and corporate income taxes.

The auditing level is equal to the number of verification actions divided by taxpayer population (and then multiplied by 100 so that it can be interpreted as number of verification actions per 100 taxpayers). This measure is therefore similar to other well-known ratio measures, such as crime rates or per capita indicators, which are commonly used to make cross-country comparisons.

3.1.3. Penalty and interest rate

The punishment of tax offenses typically involves a penalty equivalent to a percentage of the amount of unpaid taxes plus interests for delayed payments (Andreoni et al., 1998; OECD, 2013). The penalty measure is taken from the OECD Tax Administration Comparative Reports (2017a). To guarantee comparability between countries, we focus on penalties that are expressed as percentages, and exclude observations based on alternative types of penalties (e.g., fixed monetary sums). The penalty measure is equal to the maximum percentage that tax offenders must pay for tax offenses. The interest rate measure corresponds to the short-term lending interest rate, as all countries in the sample impose interests once the payment date has prescribed, and the rates they employ are generally influenced by bank lending rates (Cummings, Martinez-Vazquez, McKee, & Torgler, 2009; OECD, 2007). Interest rates of OECD and non-OECD countries are taken from the OECD Database (2017b) and from the World Bank (2017a), respectively.

3.1.4. PIT and CIT

The PIT measure corresponds to the top marginal personal income tax rate. The CIT measure corresponds to the combined corporate income tax rate, which is equal to the central government rate plus the sub-central rate. We control for tax rates as they may have a direct influence on tax compliance, although their influence can be ambiguous (Allingham & Sandmo, 1972; Andreoni et al., 1998; Sandmo, 2005). Tax rates of OECD and non-OECD countries are taken from the OECD Tax Database (2017c) and from the KPMG Tax Database (2017), respectively.

3.1.5. GDP per capita

This measure is taken from the IMF World Economic Outlook Database (2017). It corresponds to the gross domestic product (GDP) per capita, based on purchasing-power-parity (in current international US dollars). We control for GDP per capita as it may capture unobserved factors that may influence compliance (factors associated with, e.g., wealth, economic development, the government's enforcement capacity). This measure has been employed in prior country-level analyses of tax compliance (Fuest & Riedel, 2009; Richardson, 2006).

3.1.6. Trust in government

Trust in government is based on two measures gathered by the IMD (2017a), and which are made publicly available by the World Bank (World Bank, 2017b). These measures correspond to business executives' perception of the extent to which (i) government decisions are effectively implemented, and (ii) government policy is transparent (both item scales range from 0 to 10). Trust in government is equal to the average score of these two measures (Cronbach's $\alpha = 0.94$).

3.2. Variable transformations

Auditing level, penalty, interest rate, and GDP per capita have significantly skewed distributions, so these variables are transformed using their natural logarithm. To examine the expected U-shaped association, the transformed auditing level measure is mean-centered and its square is incorporated into the analysis.

3.3. Sample

Since the publication of the first OECD Tax Administration Comparative Report (OECD, 2017a), the OECD has covered an increasing number of countries. The 2004 report includes 30 OECD and zero non-OECD countries, whereas the 2015 report includes 34 OECD and 22 non-OECD countries. Notably, not all observations are available for all the countries and years that are included in these reports.

We take two concrete actions to make most use of the available observations. First, we impute missing values into variables that are relatively stable through time (i.e., variables for which within-country variance is relatively small). These variables are taxpayer population, penalty, interest rate, PIT, and CIT. Each imputed value corresponds to the closest available

observation. Notably, we do not impute missing values into tax evasion or verification actions, which are the key variables of the analysis.

To rule out the possibility that imputing values into taxpayer population affects the expected association between the auditing level and tax evasion, we also test our hypothesis without imputing values into taxpayer population. Table 1 contains the number of country-year observations available for each variable, with and without imputed values.

The second action we take is to examine whether there is a control variable that significantly reduces the sample, and assess whether it can be safely excluded from the analysis. As shown in Table 1, penalty is a control variable with a particularly low number of available observations. This is because there are only four years and a limited number of countries for which penalties for tax offenses are expressed in percentages.

To determine whether penalty can be safely removed from the analysis, we employ criteria that minimize the risk of omitted variable bias: (i) the variable under consideration should be significantly stable through time (so that its potential influence is likely captured by country fixed-effects), (ii) in the regression analysis, this variable should not be significantly associated with the dependent variable, and (iii) including or excluding this variable should not affect the main association of interest (between auditing level and tax evasion). This is indeed the case for the penalty for tax offenses. For robustness, we test our hypothesis with and without this variable.

The model that makes most use of available observations is taken as the baseline model, which in this case includes imputed values and excludes the penalty for noncompliance. The list of countries included in the baseline model is presented in Table 2.

3.4. Descriptive statistics

Table 1 contains descriptive statistics based on untransformed data (i.e., without applying the natural logarithm to auditing level, penalty, interest rate, and GDP per capita), which allows for a more intuitive interpretation of means. On average, tax authorities take 7.4 verification actions per 100 taxpayers, and impose penalties equivalent to 1.4 times the amount of unpaid taxes. There is considerable variation in the tax evasion and auditing level measures, which indicates that these are suitable for empirical analysis. There are no significant differences between variables with and without imputed values, which indicates that the employed value imputation technique is appropriate.

It is important to note that 85% of the untransformed auditing level observations lie between 0.2 and 20 verification actions per 100 taxpayers. For reference, the audit rate of the IRS in the U.S. ranged between 0.8% and 4.75% between the 1960s and 1990s (Andreoni et al., 1998), and the estimated average audit rate for U.S. public firms was 29% between 1992 and 2008 (Hoopes, Mescall, & Pittman, 2012). This provides support for the validity of the auditing level measure, and suggests that observations between countries and years are comparable.

3.5. Model specification

The regression analysis is based on a fixed-effects model with year dummies and standard errors clustered by country, which are robust to heteroskedasticity and autocorrelation. The baseline model is labeled as Model 1, and is specified as follows:

$$tax\ evasion_{i,t+1} = \alpha_i + \beta_1\ auditing\ level_{i,t} + \beta_2\ auditing\ level_{i,t}^2 + controls + u_{i,t},$$

where i is the country identifier, t is the year, α is the intercept for each country, *controls* are control variables with their respective regression coefficients (including year dummies), and u is the error term.

Table 1
Descriptive statistics.

	imputed values	N	M	SD	min	max
Tax evasion		381	5.43	1.66	1.00	9.30
Verification actions		381	0.82	1.58	0.00	12.00
Taxpayer population		211	20.23	45.39	0.15	305.90
Taxpayer population	✔	381	21.12	44.93	0.15	305.90
Auditing level		381	7.42	17.59	0.01	238.97
Penalty		107	1.44	1.57	0.10	10.00
Penalty	✔	315	1.47	1.53	0.10	10.00
Interest rate		372	5.01	6.40	0.02	67.10
Interest rate	✔	381	5.06	6.33	0.02	67.08
CIT		365	26.66	7.25	10.00	40.90
CIT	✔	381	26.74	7.29	10.00	40.87
PIT		365	37.08	11.51	0.10	59.00
PIT	✔	381	36.75	11.54	0.10	59.00
GDP per capita		381	28.02	14.18	2.91	81.10
Trust in government		381	4.59	1.60	0.75	8.50

Each imputed value corresponds to the closest available observation. Missing values are not imputed into *tax evasion* or *verification actions*, which are the key variables of the analysis. *Verification actions* and *taxpayer population* are expressed in millions.

Table 2
Countries included in the analysis.

Argentina	Estonia	Japan	Russia
Australia	Finland	Korea	Singapore
Austria	France	Latvia	Slovak Republic
Belgium	Germany	Lithuania	Slovenia
Brazil	Greece	Luxembourg	South Africa
Bulgaria	Hong Kong	Malaysia	Spain
Canada	Hungary	Mexico	Sweden
Chile	Iceland	Netherlands	Switzerland
China	India	New Zealand	Thailand
Colombia	Indonesia	Norway	United Kingdom
Croatia	Ireland	Poland	United States
Czech Republic	Israel	Portugal	
Denmark	Italy	Romania	

3.6. Robustness strategy

To assess whether the results of Model 1 (the baseline model) are robust to alternative method choices, we conduct a series of additional tests. In Model 2, we do not impute values into taxpayer population (to rule out the possibility that this affects the auditing level measure and its potential association with tax evasion). Model 3 excludes outliers, defined as auditing level values below 0.1 and above 100 verifications per 100 taxpayers. Model 4 includes the penalty for tax offenses (to rule out the possibility that omitting this variable influences our results).

Models 5–9 take into consideration possible issues associated with endogeneity and reverse causality. Our hypothesis focuses on the contemporaneous association between the auditing level and tax evasion. Examining non-contemporary associations requires the review of different theory (related to, e.g., lagged effects; Mittone, 2006), and lies beyond the scope of the present study. The examination of contemporaneous associations can nevertheless raise questions about possible endogeneity or reverse causality. We take three concrete actions in response to this.

First, we examine the association between auditing level observations from year t (which are year-end observations that cover the entirety of year t) and tax evasion observations from year $t + 1$ (which are observations that cover a large fraction of year t and a small fraction of year $t + 1$).[1] Second, in addition to estimating standard errors that are robust to autocorrelation, we also test our hypothesis using (i) a model in which the disturbance term is assumed to be first-order autoregressive (Model 5; Baltagi & Wu, 1999), and (ii) a model based on the Arellano-Bond System Generalized Method-of-Moments (GMM) estimator, which uses lags to generate instruments for possibly endogenous independent variables (Model 6; Roodman, 2006).[2] In Model 6, CIT, PIT, and year dummies are treated as exogenous; interest rate and GDP per capita as predetermined; and trust in government and auditing level as endogenous variables (with a lag limit equal to 6). Third, given that causality requires the cause to antecede the effect, we test whether the examined association violates the expected temporal order. Model 7 tests whether auditing explains prior evasion, while Models 8 and 9 test whether prior evasion explains auditing.

4. Results

Our hypothesis is that there is a U-shaped association between the auditing level and tax evasion. To test this expectation, we employ a stringent test for U-shaped associations (we use the Stata command -*utest*-; Lind & Mehlum, 2010). This test requires the sign and magnitude of the linear and quadratic terms of the auditing level to be indicative of a concave curve, the slope at the minimum value to be negative and significant (i.e., the left-side of the association must be decreasing), and the slope at the maximum value to be positive and significant (i.e., the right-side of the association must be increasing).

The regression results of Models 1–4 are reported in Table 3. The last row of the table corresponds to the test of the U-shaped association. In these four models, the regression coefficient of *auditing level*^{2} is positive and significant, and the result of the U-shape test provides further support for the expected association. Models 2, 3, and 4 follow the same pattern of results as the baseline model (Model 1). The results hold when we do not impute missing values into taxpayer population (Model 2), exclude outliers (Model 3), and include the penalty for tax offenses (Model 4).

These findings indicate that tax evasion decreases until a certain auditing level is reached, and increases beyond that level (i.e., an elevated auditing level backfires). Fig. 1 depicts this association as predicted by Model 1.

---

[1] Perceptions about tax evasion and trust in government are gathered each year between January and April (IMD, 2017b). This implies that these perceptions capture a small fraction of the year in which they are gathered, and a large fraction of year before. For this reason, when we analyze data from year t, we employ observations of tax evasion and trust in government from year $t + 1$. In this way we rule out the possibility that year-end auditing level explains tax evasion of several months of the year before.

[2] For Model 5, we use the Stata command -*xtregar*-. This model is also referred to as a regression with AR(1) disturbances. For Model 6, we use the Stata command -*xtabond2*-, follow the "standard choices" of instruments (i.e., two-step system GMM, corrected standard errors, small-sample adjustments, and orthogonal deviations; Roodman, 2006, p. 38), collapse instruments to reduce the instrument count, and in general follow recommendations on how to conduct and report this type of estimations (Roodman, 2009).

Table 3
Results of Models 1–4.

	(1) baseline model	(2) without imputed values	(3) without outliers	(4) including *penalty*
Auditing level	0.059	0.108*	0.034	0.046
	(0.048)	(0.059)	(0.051)	(0.055)
Auditing level^2	0.053***	0.052**	0.055**	0.063***
	(0.019)	(0.022)	(0.022)	(0.019)
Trust in government	−0.471***	−0.574***	−0.453***	−0.518***
	(0.078)	(0.092)	(0.081)	(0.088)
Interest rate	−0.133	−0.131	−0.146	−0.230*
	(0.111)	(0.135)	(0.113)	(0.134)
CIT	−0.023	−0.022	−0.017	−0.007
	(0.030)	(0.037)	(0.030)	(0.033)
PIT	0.009	−0.000	0.009	0.010
	(0.021)	(0.019)	(0.021)	(0.022)
GDP per capita	−0.863	−0.864	−0.632	−0.376
	(1.295)	(1.566)	(1.316)	(1.367)
Penalty				0.192
				(0.190)
Constant	10.332**	11.152*	9.329*	8.407
	(5.049)	(6.027)	(5.130)	(5.572)
N	381	211	371	315
Number of countries	50	50	49	38
R^2 (within)	0.484	0.559	0.456	0.524
F	32.06***	49.14***	36.54***	36.08***
U-shape test	2.30**	1.75**	2.18**	2.80***

Standard errors clustered by country are in parenthesis, and are robust to heteroskedasticity and serial correlation. Year dummies (time effects) are included in all models, but not reported in this table. Model 1 is the baseline model, which has imputed values, and excludes the penalty for tax offenses. In Models 2, 3, and 4, we do not impute missing values into taxpayer population, exclude outliers, and include the penalty for tax offenses, respectively.
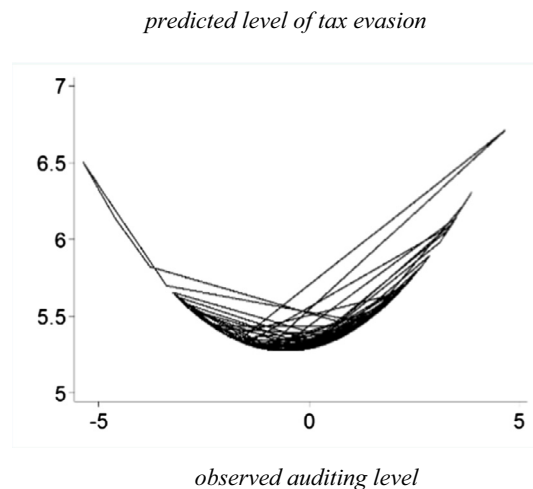* $p < 0.10$
** $p < 0.05$
*** $p < 0.01$

predicted level of tax evasion

observed auditing level

Fig. 1. Association predicted by Model 1. *Auditing level* is mean-centered.

In line with prior studies of tax morale and voluntary compliance, the association between tax evasion and trust in government is negative and significant (Kirchler et al., 2008; Wahl et al., 2010). This provides support for the idea that favorable views about the authority motivate taxpayers to pay their taxes. Tax evasion is not associated with the penalty or the other control variables – except for interest rate in Model 4, which has a negative sign and is indeed expected to act as a deterrent against tax evasion.

4.1. Endogeneity and reverse causality

The regression results of Models 5–9 are reported in Table 4. Although the Wooldridge test indicates that there is no serial correlation in the baseline model ($F = 2.22$, $p = 0.14$), we nevertheless test our hypothesis assuming first-order

Table 4
Results of Models 5 – 9.

	(5) assuming auto-correlation	(6) with possibly endogenous regressors	(7) using prior *tax evasion* as DV	(8) using *auditing level* as DV	(9) using *auditing level*^{2} as DV
Auditing level	0.035	−0.020	−0.071		0.302
	(0.059)	(0.122)	(0.046)		(0.470)
Auditing level^2	0.059***	0.058**	0.005		
	(0.022)	(0.025)	(0.020)		
Prior tax evasion				−0.138	0.067
				(0.095)	(0.243)
N	331	381	333	333	333
Number of countries	47	50	47	47	47
R^2 (within)	0.582		0.415	0.072	0.061
F	23.295***	25.700***	15.599***	1.820*	1.482
U-shape test	2.42***	1.77**			
Number of instruments		47			
AB test for AR(2)		0.87			
Hansen test		0.42			
Difference Hansen tests					
GMM instruments		0.45			
Endogenous variables only		0.23			

Prior tax evasion refers to observations from year *t*. Model 5 assumes the disturbance term is first-order autoregressive. Model 6 is based on the Arellano–Bond (AB) System GMM estimator (Roodman, 2006, 2009), and treats *CIT*, *PIT*, and year dummies as exogenous; *interest rate* and *GDP per capita* as predetermined; and *trust in government* and *auditing level* as endogenous (with a lag limit equal to 6). AR(2) refers to second-order autocorrelation in first differences. Model 7 takes *prior tax evasion* as the dependent variable (DV). Models 8 and 9 take *auditing level* and *auditing level*^{2} as the DV, respectively. Year dummies (time effects) and all control variables of the baseline model are included in all models, but not reported in this table. Standard errors are in parenthesis. For Models 7–9, standard errors are clustered by country and are robust to heteroskedasticity and serial correlation.
* $p < 0.10$
** $p < 0.05$
*** $p < 0.01$

autocorrelation (Model 5), and obtain the same pattern of results. In favor of the baseline model specification, the implication is that the observations of each year can be treated as independent.

Model 6 is based on a two-step system GMM estimator (Roodman, 2006), which uses lags as instruments of possibly endogenous regressors. It can be seen that, when auditing level and trust in government are treated as endogenous, the model complies with the necessary conditions for this type of estimation, and yields results that are consistent with the results of the baseline model. In this case, the number of instruments does not exceed the number of countries (i.e., 47 < 50), the ArellanoBond test indicates that there is no second-order autocorrelation in first differences ($p = 0.87$), and the Hansen and difference-in-Hansen tests indicate that the employed instruments can be treated as exogenous. The implication is that the results hold when we take measures to control for possible endogeneity.

Models 7, 8, and 9 are used to test whether the association between auditing level and tax evasion violates the expected temporal order. In line with our expectation of a contemporaneous association, there is no evidence indicating that auditing explains prior evasion (Model 7), or that the auditing level – or the square of auditing level – are explained by prior evasion (Models 8 and 9).

5. Discussion

The present analysis responds to the long-lasting call to examine "whether there is any specific deterrent effect of an audit and to uncover the reasons for the presence or absence of such an effect" (Andreoni et al., 1998, p. 844). It also responds to calls from multilateral organizations to use standardized country-level data in order to evaluate the effectiveness of enforcement strategies employed in practice (IMF, 2015; OECD, 2010). Building on behavioral theories of tax compliance, we hypothesize that compliance increases until a certain auditing level is reached, and decreases beyond that level. Using country-level data from 2003–2014, we obtain strong support for a U-shaped association between the auditing level and tax evasion.

In lay terms, this implies that more auditing is not necessarily better. Seen from a different angle, tax authorities may benefit from *trusting* taxpayers to some extent, as high levels of distrust may be undesirable (Mendoza & Wielhouwer, 2015). In particular, tax authorities may have incentives to investigate how taxpayers perceive different auditing levels. Over two decades ago, researchers were already finding evidence in favor of a positive – yet decreasing – association between auditing and compliance, which raises the question of whether there are limits on what auditing can achieve (Alm, McClelland, & Schulze, 1992). To the best of our knowledge, however, this is the first study to examine a possible U-shaped association between auditing and tax evasion.

An important point of discussion is whether the measures employed in this study have an adequate degree of validity and can be considered as "credible" (Alm, 2012; Slemrod & Weber, 2012). One advantage of the auditing level measure is that it is based on hard (externally verifiable) data. The disadvantage, however, is that with the available data it is not possible to distinguish between different types of verification actions. It would be ideal to control for the fraction of verification actions that are, for example, random vs. targeted, assigned to firms vs. individuals, uniquely vs. repeatedly targeted at the same taxpayer, or intrusive and inquisitive vs. educational and supportive. Controlling for these attributes would shed light on why or when a possible backfiring effect may take place.

The advantage of the tax evasion measure is that it is based on perceptions of business executives, who represent a relatively homogeneous group, may have on average relatively more knowledge of tax evasion practices than the general population, and do not report their own level of compliance but rather their perception of tax evasion in general. This measure is therefore different from other self-report measures which are relatively more subject to untruthful reporting or desirability bias. The disadvantage, however, is that people may under-estimate or over-estimate the general level of tax evasion. This problem is largely reduced if misperceptions are systematic and somehow captured by year- and country-effects. Alternatively, heterogeneous misperceptions about tax evasion would generate noise and act against rather than in favor of finding statistical support for an expected and well-defined pattern of results. In sum, we acknowledge that these two measures are imperfect, but there are various reasons why they may have an adequate degree of validity.

Another point of discussion relates to possible endogeneity and reverse causality. In practice, tax authorities may adjust auditing efforts based on the level of tax evasion that they observe, or audits may be designed to identify repeat offenders. This raises questions about the direction and dynamic nature of the association between auditing and tax evasion. A crucial point is that prior theory provides several reasons why the *square* of the auditing level might be associated with tax evasion, but it is difficult to find theory in favor of an association in the opposite direction. Reverse causality would imply that tax authorities follow a possibly contradictory strategy: as tax evasion increases, the authority would respond with more auditing if the auditing level is already high, and with less auditing if the auditing level is already low. In addition to these theoretical arguments, we find no evidence indicating that the observed association violates the expected temporal order (i.e., auditing is not explained by prior evasion or vice versa).

We also include stringent tests to rule out potential issues associated with autocorrelation, and employ instruments to deal with the possible endogenous nature of the auditing level. Although these tests provide support for our hypothesis, endogeneity cannot be fully ruled out with the available data, and future studies may further explore the expected backfiring effect using alternative measures, instruments, and procedures (e.g., with the use of experimental methods).

We include relevant control variables based on prior research on country-level compliance. We acknowledge that there can be other variables that we do not account for and which could influence tax evasion (e.g., the complexity of the tax system). If these variables are relatively stable through time, it is possible that they are largely captured by the fixed effects. Future research on the backfiring effect hypothesis may explore the influence of other relevant variables that vary across years and countries and are not included in our model.

Finally, we observe that tax evasion decreases with trust in government, which prior research identifies as a strong predictor of voluntary compliance. This suggests that observing a backfiring effect may depend on the extent to which compliance is voluntary. When compliance is mainly motivated by enforcement (i.e., when compliance is mainly enforced and not voluntary), observing a backfiring effect might be less likely. In line with this view, prior research suggests that backfiring effects of control and deterrence are tightly connected with perceptions of distrust and unfairness (Balliet & Van Lange, 2013a, 2013b; Kessler & Leider, 2016; Mooijman et al., in press; Schildberg-Hörisch & Strassmair, 2012). The expected backfiring effect may thus relate to how excessive auditing hinders voluntary compliance.

Acknowledgments

References

Allingham, M. G., & Sandmo, A. (1972). Income tax evasion: A theoretical analysis. *Journal of Public Economics, 1*(3), 323–338.
Alm, J. (2012). Measuring, explaining, and controlling tax evasion: Lessons from theory, experiments, and field studies. *International Tax and Public Finance, 19*(1), 54–77.
Alm, J., McClelland, G. H., & Schulze, W. D. (1992). Why do people pay taxes? *Journal of Public Economics, 48*(1), 21–38.
Andreoni, J., Erard, B., & Feinstein, J. (1998). Tax compliance. *Journal of Economic Literature, 36*(2), 818–860.
Balliet, D., & Van Lange, P. A. (2013a). Trust, conflict, and cooperation: A meta-analysis. *Psychological Bulletin, 139*(5), 1090–1112.
Balliet, D., & Van Lange, P. A. (2013b). Trust, punishment, and cooperation across 18 societies: A meta-analysis. *Perspectives on Psychological Science, 8*(4), 363–379.
Baltagi, B. H., & Wu, P. X. (1999). Unequally spaced panel data regressions with AR(1) disturbances. *Econometric Theory, 15*(06), 814–823.
Bazart, C., & Bonein, A. (2014). Reciprocal relationships in tax compliance decisions. *Journal of Economic Psychology, 40*, 83–102.
Becker, G. S. (1968). Crime and punishment: An economic approach. *Journal of Political Economy, 76*(2), 169–217.
Beer, S., Kasper, M., Kirchler, & Erard, B. (2015). Audit impact study. TAS Research and Related Studies, Volume 2. Online available at <taxpayeradvocate.irs.gov>.

Bobek, D. D., Roberts, R. W., & Sweeney, J. T. (2007). The social norms of tax compliance: Evidence from Australia, Singapore, and the United States. *Journal of Business Ethics, 74*(1), 49–64.

Braithwaite, V. (2003a). A new approach to tax compliance. In V. Braithwaite (Ed.), *Taxing democracy, understanding tax avoidance and evasion* (pp. 1–11). Aldershot: Ashgate Publishing Limited.

Braithwaite, V. (2003b). Taxing democracy. In V. Braithwaite (Ed.), *Taxing democracy, understanding tax avoidance and evasion* (pp. 15–39). Aldershot: Ashgate Publishing Limited.

Braithwaite, V. (2007). Responsive regulation and taxation: Introduction. *Law & Policy, 29*(1), 3–10.

Braithwaite, V. (2012). Resistant and dismissive defiance towards tax authorities. In A. Crawford & A. Hucklesby (Eds.), *Legitimacy and compliance in criminal justice* (pp. 91–115). New York and London: Routledge.

Braithwaite, V., Murphy, K., & Reinhart, M. (2007). Taxation threat, motivational postures, and responsive regulation. *Law & Policy, 29*(1), 137–158.

Cullis, J. G., & Lewis, A. (1997). Why people pay taxes: From a conventional economic model to a model of social convention. *Journal of Economic Psychology, 18*(2), 305–321.

Cummings, R. G., Martinez-Vazquez, J., McKee, M., & Torgler, B. (2009). Tax morale affects tax compliance: Evidence from surveys and an artefactual field experiment. *Journal of Economic Behavior & Organization, 70*(3), 447–457.

Devos, K. (2014). Tax compliance theory and the literature. In *Factors influencing individual taxpayer compliance behaviour* (pp. 13–65). Australia: Springer.

Falk, A., & Kosfeld, M. (2006). The hidden costs of control. *The American Economic Review, 96*(5), 1611–1630.

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

Feld, L. P., & Frey, B. S. (2007). Tax compliance as the result of a psychological tax contract: The role of incentives and responsive regulation. *Law & Policy, 29*(1), 102–120.

Frey, B. S. (1994). How intrinsic motivation is crowded out and in. *Rationality and Society, 6*(3), 334–352.

Frey, B. S., & Jegen, R. (2001). Motivation crowding theory. *Journal of Economic Surveys, 15*(5), 589–611.

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

Fuest, C. & Riedel N. (2009). Tax evasion, tax avoidance and tax expenditures in developing countries: A review of the literature. Report prepared for the UK Department for International Development.

Gangl, K., Torgler, B., Kirchler, E., & Hofmann, E. (2014). Effects of supervision on tax compliance: Evidence from a field experiment in Austria. *Economics Letters, 123*(3), 378–382.

Graetz, M. J., Reinganum, J. F., & Wilde, L. L. (1986). The tax compliance game: Toward an interactive theory of law enforcement. *Journal of Law, Economics, & Organization*, 1–32.

Hartner, M., Rechberger, S., Kirchler, E., & Schabmann, A. (2008). Procedural fairness and tax compliance. *Economic Analysis and Policy, 38*(1), 137–152.

Hofmann, E., Hoelzl, E., & Kirchler, E. (2008). Preconditions of voluntary tax compliance: Knowledge and evaluation of taxation, norms, fairness, and motivation to cooperate. *Journal of Psychology, 216*(4), 209–217.

Hollander, J. A., & Einwohner, R. L. (2004). Conceptualizing resistance. *Sociological Forum, 19*(4), 533–554.

Hoopes, J. L., Mescall, D., & Pittman, J. A. (2012). Do irs audits deter corporate tax avoidance? *The Accounting Review, 87*(5), 1603–1639.

IMD. (2017a). IMD – World Competitiveness Yearbook. Online available at <http://www.imd.org/wcc/wcy-world-competitiveness-yearbook/>.

IMD. (2017b). Methodology and principles of analysis. Online available at <http://www.imd.org/wcc/world-competitiveness-center-mission/methodology/>.

IMF. (2015). Current challenges in revenue mobilizaton: Improving tax compliance. Staff Report. Online available at <https://www.imf.org/external/np/pp/eng/2015/020215a.pdf>.

IMF. (2017). World Economic Outlook Database. Online available at <https://www.imf.org/external/pubs/ft/weo/2015/02/weodata/index.aspx>.

Iyer, G. S., Reckers, P. M., & Sanders, D. L. (2010). Increasing tax compliance in Washington State: A field experiment. *National Tax Journal, 63*(1), 7–32.

Kessler, J. B., & Leider, S. (2016). Procedural fairness and the cost of control. *Journal of Law, Economics, & Organization, 32*(4), 685–718. http://dx.doi.org/10.1093/jleo/eww009.

Kirchler, E. (1999). Reactance to taxation: Employers attitudes towards taxes. *The Journal of Socio-Economics, 28*(2), 131–138.

Kirchler, E. (2007). *The economic psychology of tax behaviour*. Cambridge University Press.

Kirchler, E., Muehlbacher, S., Kastlunger, B., & Wahl, I. (2010). Why pay taxes? A review of tax compliance decisions. In: Alm, J., Martinez-Vazquez, J., Torgler, B. (Eds.), *Developing alternative frameworks for explaining tax compliance*. Routledge International Studies in Money and Banking (pp. 15–32).

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

KPMG. 2017. Tax tools & resources. Online available at <https://home.kpmg.com/xx/en/home/services/tax/tax-tools-and-resources.html>.

Lawsky, S. B. (2008). Fairly random: On compensating audited taxpayers. *Connecticut Law Review, 41*, 161–200.

Lederman, L. (2016). *To what extent does enforcement crowd out voluntary tax compliance? Unpublished manuscript*. Bloomington, U.S.: Indiana University Maurer School of Law.

Levi, M. (1998). A state of trust. In V. Braithwaite & M. Levi (Eds.). *Trust and governance* (Vol. 1, pp. 77–101). New York: Russel Sage Foundation.

Lind, J. T., & Mehlum, H. (2010). With or without U? The appropriate test for a U-shaped relationship. *Oxford Bulletin of Economics and Statistics, 72*(1), 109–118.

Mendoza, J. P., & Wielhouwer, J. L. (2015). Only the carrot, not the stick: Incorporating trust into the enforcement of regulation. *PLoS ONE, 10*(2). http://dx.doi.org/10.1371/journal.pone.0117212.

Mittone, L. (2006). Dynamic behaviour in tax evasion: An experimental approach. *The Journal of Socio-Economics, 35*(5), 813–835.

Mooijman, M., van Dijk, W.W., van Dijk, E., & Ellemers, N. (in press). On sanction-goal justifications: How and why deterrence justifications undermine rule compliance. *Journal of Personality and Social Psychology*.

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.

Murphy, K. (2004). The role of trust in nurturing compliance: A study of accused tax avoiders. *Law and Human Behavior, 28*(2), 187–209.

Murphy, K. (2005). Regulating more effectively: The relationship between procedural justice, legitimacy, and tax non-compliance. *Journal of Law and Society, 32*(4), 562–589.

Murphy, K. (2008). Enforcing tax compliance: To punish or persuade? *Economic Analysis and Policy, 38*(1), 113–135.

OECD. (2007). Tax administration in OECD and selected non-OECD countries: Comparative information series 2006. OECD Publishing. Online available at <http://www.oecd.org/ctp/consumption/CIS-2006.pdf>.

OECD. (2010). Understanding and influencing taxpayers' compliance behaviour. OECD Publishing. Online available at <http://www.oecd.org/ctp/administration/46274793.pdf>.

OECD (2013). *Tax administration 2013: Comparative information on OECD and other advanced and emerging economies*. OECD Publishing.

OECD. (2017a). Forum on tax administration – Comparative information. Online available at <http://www.oecd.org/tax/forum-on-tax-administration>.

OECD. (2017b). OECD statistics – Short-term interest rates. Online available at <http://stats.oecd.org/index.aspx?queryid=26668>.

OECD. (2017c). Tax administration database. Online available at <http://www.oecd.org/site/ctpfta/taxadministrationdatabase.htm>.

Posner, E. A. (2000). Law and social norms: The case of tax compliance. *Virginia Law Review, 86*(8), 1781–1819.

Richardson, G. (2006). Determinants of tax evasion: A cross-country investigation. *Journal of International Accounting, Auditing and Taxation, 15*(2), 150–169.

Roodman, D. (2006). How to do Xtabond2: An introduction to difference and system GMM in Stata. Center for Global Development Working Paper No. 103. Online available at http://dx.doi.org/10.2139/ssrn.982943.

Roodman, D. (2009). A note on the theme of too many instruments. *Oxford Bulletin of Economics and Statistics, 71*(1), 135–158.

Sandmo, A. (2005). The theory of tax evasion: A retrospective view. *National Tax Journal, 58*, 643–663.

Schildberg-Hörisch, H., & Strassmair, C. (2012). An experimental test of the deterrence hypothesis. *Journal of Law, Economics, & Organization, 28*(3), 447–459. http://dx.doi.org/10.1093/jleo/ewq015.

Sheffrin, S. M., & Triest, R. K. (1991). Can brute deterrence backfire? Perceptions and attitudes in taxpayer compliance. In J. Slemrod (Ed.), *Why people pay taxes: Tax compliance and enforcement* (pp. 193–218). University of Michigan Press.

Sherman, L. W. (1993). Defiance, deterrence, and irrelevance: A theory of the criminal sanction. *Journal of Research in Crime and Delinquency, 30*(4), 445–473.

Slemrod, J. (1998). On voluntary compliance, voluntary taxes, and social capital. *National Tax Journal, 51*, 485–491.

Slemrod, J., Blumenthal, M., & Christian, C. (2001). Taxpayer response to an increased probability of audit: Evidence from a controlled experiment in Minnesota. *Journal of Public Economics, 79*(3), 455–483.

Slemrod, J., & Weber, C. (2012). Evidence of the invisible: Toward a credibility revolution in the empirical analysis of tax evasion and the informal economy. *International Tax and Public Finance, 19*(1), 25–53.

Smith, K. W., & Stalans, L. J. (1991). Encouraging tax compliance with positive incentives: A conceptual framework and research directions. *Law & Policy, 13* (1), 35–53.

Spicer, M. W., & Thomas, J. E. (1982). Audit probabilities and the tax evasion decision: An experimental approach. *Journal of Economic Psychology, 2*(3), 241–245.

Srinivasan, T. N. (1973). Tax evasion: A model. *Journal of Public Economics, 2*(4), 339–346.

Stalans, L., & Lind, E. A. (1997). The meaning of procedural fairness: A comparison of taxpayers and representatives views of their tax audits. *Social Justice Research, 10*(3), 311–331.

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

Torgler, B. (2003). Tax morale, rule-governed behaviour and trust. *Constitutional Political Economy, 14*(2), 119–140.

Torgler, B. (2004). Tax morale, trust and corruption: Empirical evidence from transition countries. Center for Research in Economics, Management and the Arts (CREMA), Working Paper (2004-05).

Torgler, B. (2007). *Tax compliance and tax morale: A theoretical and empirical analysis.* Edward Elgar Publishing.

Tsakumis, G. T., Curatola, A. P., & Porcano, T. M. (2007). The relation between national cultural dimensions and tax evasion. *Journal of International Accounting, Auditing and Taxation, 16*(2), 131–147.

Wahl, I., Kastlunger, B., & Kirchler, E. (2010). Trust in authorities and power to enforce tax compliance: An empirical analysis of the "Slippery Slope framework. *Law & Policy, 32*(4), 383–406.

Wenzel, M. (2003). Tax compliance and the psychology of justice: Mapping the field. In V. Braithwaite (Ed.), *Taxing democracy* (pp. 41–70). England: Ashgate Publishing Ltd.

Wenzel, M. (2004). An analysis of norm processes in tax compliance. *Journal of Economic Psychology, 25*(2), 213–228.

Wenzel, M. (2005). Motivation or rationalisation? Causal relations between ethics, norms and tax compliance. *Journal of Economic Psychology, 26*(4), 491–508.

World Bank. (2017a). Lending interest rate. Online available at <http://data.worldbank.org/indicator/FR.INR.LEND>.

World Bank. (2017b). Worldwide Governance Indicators. Online available at <http://info.worldbank.org/governance/wgi/index.aspx>.