

IZA DP No. 4843

**The Link between the Intrinsic Motivation to
Comply and Compliance Behavior:
A Critical Appraisal of Existing Evidence**

Martin Halla

March 2010

The Link between the Intrinsic Motivation to Comply and Compliance Behavior: A Critical Appraisal of Existing Evidence

Martin Halla

*University of Linz
and IZA*

Discussion Paper No. 4843
March 2010

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ABSTRACT

The Link between the Intrinsic Motivation to Comply and Compliance Behavior: A Critical Appraisal of Existing Evidence^{*}

Recent economic literature emphasizes the importance of moral considerations to explain compliance behavior with respect to underground activities such as tax evasion. A considerable amount of research aims to identify factors that affect the intrinsic motivation to comply. However, the causal link between the intrinsic motivation to comply and actual compliance behavior is not established yet. We provide a discussion of the underlying identification problem and suggest (potentially) feasible empirical strategies to uncover a causal effect.

JEL Classification: O17, H26, A13, Z13, C81

Keywords: shadow economy, tax evasion, tax compliance, benefit fraud, tax morale, benefit morale, social norms

Corresponding author:

Martin Halla
Department of Economics
Johannes Kepler University of Linz
Altenbergerstr. 69
4040 Linz
Austria
E-mail: martin.halla@jku.at

^{*} For helpful discussion the author would like to give thanks to Friedrich G. Schneider and Benno Torgler. The usual disclaimer applies.

1. Introduction

In recent years, the economic literature on the shadow economy and tax evasion emphasized the importance of moral considerations (or social norms) to explain compliance behavior.² Likewise, research on public enforcement of law increasingly considers social norms because of their role to substitute or to complement formal laws, and because of the potential impact of laws on social norms (Polinsky and Shavell, 2000). This trend most likely results from the fact that neo-classical models of compliance – in the spirit of the economics-of-crime approach – over-predict real-world compliance. Many scholars therefore conclude that the explanation for the tendency to comply must be that individuals are obeying a norm (Posner, 2000).

As a response, theoretical papers incorporated individuals with an intrinsic motivation to comply (e.g. Gordon, 1989; Erard and Feinstein, 1994; Traxler, 2010). More recently, an increasing number of empirical papers (to be discussed below) try to quantify this intrinsic motivation with survey data. In most of the cases scholars study the case of tax evasion, and analyze the intrinsic motivation to pay taxes, which is known as *tax morale*.³ The increasing popularity of these studies can be shown by the number of papers indexed in *Google Scholar* over time.⁴ Figure 1 shows that the number of published papers on tax morale was below ten per year throughout the 1990s. However, thereafter the number sharply increased, and since 2006 we are observing more than 100 papers per year.⁵

[Figure1 somewhere around here]

These papers typically aim to identify factors (both on an individual- and a country-level) that affect the level of tax morale. A substantial amount of empirical evidence on the association between tax morale and several socio-demographic characteristics from national and international samples is available.⁶ Further, a number of papers identified different institutional arrangements that are correlated with a high level of tax morale (e.g. Torgler, 2005a).

² For a general discussion of moral considerations and social norms within economics see, for instance, Elster (1989), Posner (1997), and Posner (2002).

³ A smaller number of papers deal with *benefit morale*, i.e. with the individual reluctance to exploit the state via benefit fraud. See, for instance, Halla and Schneider (2008), Heinemann (2008), and Halla, Lackner and Schneider (2010).

⁴ Google Scholar, a freely-accessible Web search engine, indexes the full text of scholarly literature across an array of publishing formats and disciplines. It includes most peer-reviewed journals of large scholarly publishers. For more information see <http://scholar.google.com/intl/en/scholar/about.html>.

⁵ To be precise, German scholars around Günter Schmolders known as the ‘Cologne school of tax psychology’ already tried in the 1950s and 1960s to build a bridge between economics and social psychology. They emphasized the importance of tax morale to explain tax compliance behavior. See, for instance, Schmolders (1951/52, 1960, 1969).

⁶ For a survey of this literature see Torgler (2007).

These papers all – explicitly or implicitly – presume that tax morale affects actual compliance. However, it is important to bear in mind that an intrinsic motivation to comply captured by survey data, does not measure individual *behavior* but an individual *attitude*. That means a high intrinsic motivation to comply does not necessarily translate into a high level of compliance. In fact, relatively little empirical evidence on the impact of an intrinsic motivation to comply on actual compliance behavior exists.⁷ Most likely, this is due to the fact that is extremely hard to identify this causal link. First, one has to obtain and quantify both dimensions. In the case of compliance behavior, this is a non-trivial problem, since any form of non-compliance is difficult to observe. In general, non-compliant behavior cannot be measured and has to be *estimated*. The intrinsic motivation to comply has to be obtained by adequate survey techniques. Given that one can observe and link both dimensions, one needs a credible research design to establish a causal effect.

With respect to tax evasion almost all existing evidence is based on survey data. Obviously, this research design is problematic, since the accuracy of self-reported tax evasion information is questionable. A small number of papers try to solve this limitation by combining tax evasion data observed in laboratory experiments with information from post-experiment questionnaires. With respect to the shadow economy a number of papers present simple correlations between the level of tax morale and the size of the shadow economy. However, this descriptive evidence allows several interpretations. It is unclear, whether a causal effect of tax morale on the size of the shadow economy exists. Most recently, a small number of papers (Torgler and Schneider, 2007, 2009; Torgler, Schaffner and Macintyre, 2007) address this identification problem and suggest an instrumental variable approach to disentangle a causal effect.

We think a good understanding of the relation between the intrinsic motivation to comply and actual compliance behavior is very important. Future research in this area should pursue empirical strategies that are able to establish a causal link between these two dimensions. The existence of this causal link determines the significance of the whole strand of literature that analyses the determinants of the intrinsic motivation to comply.

The remainder of the paper is organized as follows. In Section 2 we discuss the necessary steps to establish this causal link, which includes a definition of the shadow economy that fits the question under consideration. Following that we summarize the literature that aims to quantify the related phenomena of tax evasion. Thereby, we will evaluate which methods produce data points that can be used to answer the question under consideration.

⁷ We will discuss the existing evidence in detail below.

Subsequently we will provide an overview of all publicly available survey data that are useful to measure the intrinsic motivation to abstain from underground activities, tax evasion, and benefit fraud. Under the assumption that we can observe and link both, the intrinsic motivation to comply and compliance behavior, we will discuss the econometric identification problem and highlight the necessary assumptions in order to establish a causal effect.

Based on these findings, Section 3 critically reviews the existing literature. At this stage we fully acknowledge that a clear identification of this causal effect is extremely difficult. However, we believe that only a thorough discussion of the identifying assumptions and their credibility allows steps forward. We do not provide a solution to solve the identification problem in this paper. However, we try to contribute to this literature by providing a clear discussion of the identification problem.

Before we conclude in Section 5, we briefly highlight in Section 4 the importance of unobserved heterogeneity for the case of tax morale and the shadow economy.

2. Establishing a Link between Attitude and Behavior

In order to identify a causal link between the intrinsic motivation to comply and actual compliance behavior one has to overcome a number of obstacles. First, one has to observe and link both dimensions. In a second step, a research design, most likely based on an instrumental variable approach, is needed.

2.1 Measuring Compliance Behavior

To start with, one has to decide which specific form of compliance should be studied. As usual, this choice (should be guided by the relevance of the question and) is limited by the availability of data. In the case of compliance, this problem is non-trivial since any form of non-compliance is difficult to observe. In general, non-compliant agents will try to hide their behavior in order to avoid punishment. That means, non-compliant behavior cannot be observed and has to be *estimated*.

This complicates an *a priori* assessment of the significance of the issue and may obscure an evaluation of the (quality of the) available data. The choice is further complicated by the fact that different forms of non-compliance that are usually studied by economists – such as underground economic activities and tax evasion – may overlap and are not mutually

exclusive. Therefore, the choice over the specific form of compliance will inevitably be related with the selection of its estimation method.⁸

2.1.1 *Defining the Shadow Economy*

In the economic literature there is no agreement on the definition of the shadow economy or on the method to estimate its size.⁹ For instance, the OECD¹⁰ uses a very broad definition that specifies five groups of activities that are collectively said to comprise the shadow economy¹¹:

- (i) *underground production*, defined as those activities that are productive and legal but are deliberately concealed from the public authorities to avoid payment of taxes or complying with regulations;
- (ii) *illegal production*, defined as those productive activities that generate goods and services forbidden by law or that are unlawful when carried out by unauthorized producers;
- (iii) *informal sector production*, defined as those productive activities conducted by unincorporated enterprises in the household sector that are unregistered and/or are less than a specified size in terms of employment, and that have some market production;
- (iv) *production of households for own final use*, defined as those productive activities that result in goods or services consumed or capitalized by the households that produced them; and
- (v) *production missed due to deficiencies in data collection program*, defined as all the productive activities that should be accounted for by the basic data collection program but are missed due to statistical deficiencies. It is sometimes referred to as the statistical underground.

For the purpose of the question under consideration a narrower definition of the shadow economy is reasonable. We suggest restricting the definition of the shadow economy to activities captured by the *underground production*. All other categories are either not (unambiguously) related to non-compliance behavior, or are not within the usual domain of this strand of literature. The *production of households for own final use* is clearly not connected to non-compliance behavior. In the case of the *informal sector production* it is less

⁸ A possible way to avoid this problem is laboratory experiments.

⁹ See, for instance, Frey and Schneider (2001).

¹⁰ See OECD (2002); this is a joint publication of the *Organisation for Economic Co-operation and Development*, the *International Monetary Fund*, the *International Labour Organisation*, and the *Statistical Committee of the Commonwealth of Independent States*.

¹¹ To be precise, OECD (2002) uses the term *non-observed economy*.

clear, since there can be some overlap. For instance, informal sector enterprises may prefer to remain unregistered in order to avoid compliance with regulations and to minimize production costs (see, OECD 2002, page 39). However, since these activities are not necessarily performed with the deliberate intention of non-complying (such as tax evasion, infringing labor legislation or other regulations) we argue to exclude this category. Finally *illegal production*, such as sale of drugs or trafficking stolen goods, is by definition non-compliant behavior. However, these forms of non-compliance are usually not studied (by economists) with a reference to an intrinsic motivation to comply.

Our recommended definition of the shadow economy comprises therefore as stated by OECD (2002) on page 37 “*only activities that may be both productive in an economic sense and also quite legal (provided certain standards or regulations are complied with) but deliberately concealed from public authorities for the following kinds of reasons*”:

- to avoid payment of income, value added or other taxes;
- to avoid the payment of social security contributions;
- to avoid having to meet certain legal standards such as minimum wages, maximum hours, safety or health standards, etc.;
- to avoid complying with certain administrative procedures, such as completing statistical questionnaires or other administrative forms (see OECD, 2002, page 38).

Therefore, this definition of the shadow economy covers well-known practices such as under-reporting of income in order to avoid taxation, fraudulent receipt of unemployment benefits, unofficially operating enterprises who want avoid long and costly bureaucratic procedures, or infringement of employment regulations or immigration laws by firms who hire labor “off the books”. Clearly, this quantity is related to tax evasion. But it does not cover all forms of tax evasion. It includes some methods of tax evasion (e.g. under-reporting of income), and excludes others that are non-productive, such as over-claiming deductions or exemptions.¹²

The economic literature offers various so-called *direct* and *indirect approaches* to estimate the size of the shadow economy, and a full review of the existing methods is well beyond the scope of this paper.¹³ In general, the methods that are used to estimate the size of the shadow economy and the extent of tax evasion largely overlap. In the next section, we will

¹² Tanzi (1999) provides a detailed discussion of the nexus between the shadow economy and tax evasion.

¹³ For a comprehensive review of the literature on the shadow economy, see, Schneider and Enste (2000). Some recent applications are Lemieux, Fortin and Fr chet te, (1994); Lyssiotou, Pashardes and Stengos, (2004) and Breusch (2005).

briefly discuss these methods, and evaluate which methods produce data points that can be used to answer the question under consideration.

2.1.2 *Measuring Tax Evasion*

Compared to the shadow economy, it is probably easier to find a consensus among economic scholars on the definition of tax evasion.¹⁴ In reality, however, there are many grey areas where the distinction between tax evasion and tax avoidance is not so clear. Both tax evasion and tax avoidance are an attempt to reduce one's own tax liabilities. They only differ in legal respects. Tax evasion is an illegal activity, whereas tax avoidance is consistent with existing law. However, even tax authorities may often inappropriately characterise particular cases (Slemrod and Yitzhaki, 2002).¹⁵ Several approaches to estimate the extent of tax evasion – on an individual- and a more aggregate-level – have been suggested in the literature.¹⁶

One category is called *direct approaches*. Traditionally, researchers tried to collect individual-level data on tax evasion with survey techniques. This has the advantage that questions on tax morale can be added easily. However, the accuracy of self-reported tax evasion information is highly questionable. Elffers, Weigel, and Helsing, (1987) managed to link tax audit data for approximately 700 Dutch taxpayers with survey responses. They show that the correlation between assessed and self-reported tax evasion is essentially zero.

In principal, the most reliable information about tax evasion should be obtained by tax audits. However, it is widely recognized that even intensive audits are not able to reveal all kinds of non-compliance (Slemrod and Yitzhaki, 2002). Moreover, regular tax audits do not constitute a representative sample. In general, tax authorities do not randomly select taxpayers to audit, but use properties of submitted returns which indicate the likelihood of noncompliance.¹⁷ Consequently, the best available source is data from randomly assigned tax audits, such as the *U.S. Taxpayer Compliance Measurement Program (TCMP)*.¹⁸ However, this data has the disadvantage that it does not include information on tax morale. In general, it is challenging to obtain survey responses that can be linked with tax audit data. It is likely

¹⁴ The development of the literature through the 1980s is surveyed by Cowell (1990). More recent literature surveys are provided by Andreoni, Erard and Feinstein (1998); Slemrod and Yitzhaki (2002); and Slemrod (2007).

¹⁵ Denis Healey, a former UK Chancellor of the Exchequer, phrased it strikingly: “*The difference between tax avoidance and tax evasion is the thickness of a prison wall*”.

¹⁶ As mentioned before, most of these approaches are also to infer on the size of the shadow economy.

¹⁷ Similarly, data from tax amnesties have a sample selection problem (Andreoni, Erard and Feinstein, 1998).

¹⁸ Under the TCMP, the Examination Branch of the *Internal Revenues Service (IRS)* periodically conducted (until the late 1980s) random in-depth audits to estimate compliance and revenue lost from non-compliance. The resulting data consisted of detailed information about what the taxpayer reported, and what the examiner concluded was correct.

ineffective if the IRS sends a questionnaire to U.S. taxpayers asking about their tax morale. Moreover, survey data collected by other institutions may often not be linkable to tax audit data due to privacy law concerns.¹⁹

An alternative direct approach to obtain individual-level data taken in the literature is based on laboratory experiments. For an early application, see, Friedland, Maital and Rutenberg (1978). Clearly, the shortcoming of this data is the artificial setting in which it is generated. For instance, it is unclear whether individuals would behave differently when they deal with real tax authorities instead of experimenters. However, it has the advantage that it can be augmented by tax morale information from pre- or post-experimental questionnaires (see, e.g. Bosco and Mittone, 1997; Torgler, Schaffner and Macintyre, 2007). Certainly, it has to be carefully checked, whether pre-experimental questionnaires affect compliance behavior and/or compliance behavior affects answers in post-experimental questionnaires.

A second category discussed in the literature is *indirect approaches*. These methods usually provide more aggregated estimates of tax evasion. Typically, these approaches rely on inferring the levels or trends in tax evasion from observable quantities, such as currency demand or national income and product accounts. Clearly, these approaches are able to produce tax evasion estimates for a large set of countries and years. These can be matched with country-averages of tax morale from international surveys. However, one has to note that these approaches have been heavily criticized in the literature on tax evasion. For instance, Slemrod and Yitzhaki (2002) conclude that “*none of these approaches is likely to be reliable [...] as their accuracy depends either on unverifiable assumptions or on how well the demand for currency is estimated*”.

Most recently, a small number of papers (e.g. Gorodnichenko, Martinez-Vazquez and Sabirianova Peter, 2009) combine economic theory and natural experiments to obtain estimates of tax evasion. While this seems to be in general a very promising direction of research, their results cannot be used for our question under consideration. Typically, these methods provide only a few data points, which are hard to link with tax morale data.

2.2 Measuring the Intrinsic Motivation to Comply

The selection of the method to quantify the intrinsic motivation to comply seems to be rather straightforward. Since one tries to measure an attitude, the only available choice is survey

¹⁹ The aforementioned paper by Elffers, Weigel, and Hessing (1987) provides a notable exception.

techniques. The more challenging part is the design of the survey questions. Among others, the formulation of the survey question has to target the form of compliance under consideration.

Table A1 and A2 (in the Appendix) provide an overview on all publicly available international survey data that is useful to measure the intrinsic motivation to abstain from underground activities, tax evasion, and benefit fraud. Only a few survey programs include questions on the intrinsic motivation to comply. Fortunately, these are all very well organized ongoing academic projects that cover a reasonable number of respondents from a large set of countries.

2.2.1 Measuring Tax Morale

The most extensive data is available to measure tax morale. To our best knowledge, there are four international survey data-sets available: (i) the *European and World Values Surveys* (WVS), (ii) the *European Social Survey* (ESS), (iii) the *International Social Survey Programme* (ISSP), and (iv) the *Latinobarometro*.²⁰ Each survey is a pooled cross-sectional data.

As Table A1 shows, each survey differs in the exact formulation of the question, the possible answer categories, and the available country-years. The WVS and the ESS offer both two different questions on tax morale. The first question in the WVS (TM_1^{WVS}) refers to tax morale in a very general way: “Please tell me for each of the following statements whether you think it can always be justified, never be justified, or something in between: Cheating on taxes if you have a chance”. Respondents are asked to evaluate on an ordered scale from “never justifiable” (1) to “always justifiable” (10). The second question (TM_2^{WVS}) is more specific, “[...] Paying cash for services to avoid taxes”, and offers the same scale to answer. Similar, the ESS asks first (TM_1^{ESS}) “How much you agree or disagree with each of these statements: Citizens should not cheat on their taxes”, and then (TM_2^{ESS}) “How wrong, if at all, do you consider the following ways of behaving to be? How wrong is someone paying cash with no receipt so as to avoid paying VAT or other taxes?”. In the first case respondents can answer on a five-point scale, in the second case on a four-point scale. The only available question in the ISSP (TM^{ISSP}) is more specific, and refers to tax cheating via income under-reporting: “Consider the following situations below. Do you feel it is wrong or not wrong if a taxpayer does not report all of his or her income in order to pay less income tax?”. The scale of answers ranges from “not wrong” (1) to “seriously wrong” (4).

²⁰ Due to the rather specific geographic restriction we will not cover the *Latinobarometro* in this paper. For further information, see <http://www.latinobarometro.org/>.

TM_1^{WVS} offers by far the most observations and is, therefore, the most widely used variable to study tax morale. It was included in each of the four survey waves (covering the time period from 1981 to 2003). Information on respondents from 80 countries was collected and in sum data from 184 country-years is available. TM_2^{WVS} was only asked in the fourth wave and provides data from 33 country-years. TM_1^{ESS} and TM_2^{ESS} were both included in the second wave of the ESS, where fieldwork was conducted between April, 2004 and December, 2006. In each case information from respondents from 25 countries is available. TM^{ISSP} has been collected in 1991 and 1998. In sum, this provides data from 48 country-years.

It is *a priori* not clear how the ideal survey question to capture tax morale should be formulated. However, one could argue that a more general formulation, (i.e. that does not only refer to one method of tax evasion) is preferable. Given this criteria, TM_1^{WVS} and TM_1^{ESS} are superior to TM_2^{WVS} , TM_2^{ESS} and TM^{ISSP} . Nevertheless, it would be reassuring if the correlation between responses from different questions *within* one survey is high. This can be checked for the WVS and ESS. The Spearman's rank correlation coefficient between TM_1^{WVS} and TM_2^{WVS} is in the full sample on an individual-level (38,560 observations from 33 countries) equal to 0.52 and the hypothesis of a zero rank correlation can be rejected with at a significance level below 0.001. For the ESS (44,802 observations from 25 countries) we observe a considerably lower correlation of 0.28 between the two alternative measures of tax morale. Nevertheless, the hypothesis of a zero rank correlation can again be rejected at a significance level below 0.001. Notably, on a country-level the correlation coefficients are more pronounced (WVS: 0.70 and ESS: 0.58).

[Table 1 and Table 2 somewhere around here]

Table 1 and Table 2 provide descriptive statistics on the tax morale variables per country. In the case of the WVS (see column 5 in Table 1) the correlation coefficients vary between 0.29 (Romania) and 0.67 (Germany and Portugal). In each country the hypothesis of a zero rank correlation can be rejected. As expected, the ESS (see column 4 in Table 2) shows lower within country correlations, but still, independence can be rejected in each case at a significance level below 0.001. Interestingly, in both surveys the Spearman's rank correlation coefficients between the two alternative measures of tax morale are not significantly correlated with the number of observations, or with the level of the country-mean of either tax morale

variable.²¹ In the case of the WVS both variables are measured on the same scale and their means can be compared. As column 3 shows, in most of the cases the level of tax morale is higher for the more general question,

Ideally, different measures of tax morale are not only highly correlated within surveys, but also *across* surveys. Unfortunately, there is little overlap between the available country-years of the different surveys. In fact, it is impossible to compare the variables from the ESS with those from the WVS and the ISSP. However, there is some overlap between the WVS and the ISSP. In order to gain some more country-years for this comparison, we extrapolate each country average of TM_1^{WVS} , TM_2^{WVS} and TM^{ISSP} by one year forward. That means we impute all missing country-averages of tax morale in year t with the value of the preceding year ($t-1$). As Table 3 shows, there is a positive, but only modest, correlation (0.24, $N=41$) between TM_1^{WVS} and TM^{ISSP} .²² However, if we drop two observations on Austria – which are certainly outliers in the ISSP-sample – the correlation increases to 0.43 ($N=39$), and independence can be rejected at a significance level below 0.01. The correlation between TM_2^{WVS} and TM^{ISSP} is practically zero (0.08, $N=20$). This result remains unchanged, even if the one observation on Austria is excluded.²³

[Table 3 somewhere around here]

In a final step we check whether the interrelation between alternative measures of tax morale and individual characteristics is similar *within* and *across* surveys. Therefore, we regress TM_1^{WVS} , TM_2^{WVS} , TM_1^{ESS} and TM_2^{ESS} on basic socio-economic characteristics. We manage to measure the dimensions of age, sex, marital status, children, educational attainment, household income and labor market status (base category is employed) in both surveys on an almost equal scale.²⁴ After cleaning the data-sets, about 30,700 observations from the WVS and about 16,600 from the ESS remain²⁵. Estimation results are presented in Table 4. For all

²¹ Results are not shown in paper, but available upon request.

²² Without imputation the correlation between TM_1^{WVS} and TM^{ISSP} is equal to 0.34 ($N=5$).

²³ Without imputation there are no observations for a comparison of TM_2^{WVS} and TM^{ISSP} available.

²⁴ Details on the definition of the variables are provided in the notes to Table 4.

²⁵ In the case of the WVS (compare Table 1) we had to exclude all 954 observations from Portugal from the regression analysis, since household income is only available on a 6-point scale. Another 6,833 observations (from various countries) are excluded due to missing information on one or more covariates. In the case of the EES (compare Table 2) we had to exclude all observations from France (1,784), Estonia (1,772), Hungary (1,420) and Ukraine (1,612). For France and Hungary information on self-employment is missing. For Estonia and Ukraine no information on the household income is provided. Another 9,436 observations (from various countries) are excluded due to missing information on one or more covariates. In most of these cases non-response on the household income question is responsible. The Spearman's rank correlation coefficients between the two

four measurements of tax morale we find consistent results with respect to the estimated coefficients' signs. (The only exception is the size of the place of residence). In most of the cases, even the statistical significance coincides across estimations. As previously found in the literature (see, e.g. Halla and Schneider, 2008) a higher age, being female, being married, a higher educational attainment, and being out of labor force (compared to being employed) are associated with a higher level of tax morale. Whereas, a high household income, self-employment, and unemployment are negatively correlated with the intrinsic motivation to pay taxes. The only notable difference is that the share of explained variation is comparably low in the case of TM_1^{ESS} .

[Table 4 somewhere around here]

2.2.2 *Measuring Benefit Morale*

We are aware of two international surveys, the WVS and the ISSP, that include a question to study the phenomenon of benefit morale, i.e. the intrinsic motivation to abstain from cheating on the state via benefit fraud. Table 2A shows the exact formulation of the question, the possible answer categories, and the available country-years.

The question on benefit morale (BM^{WVS}) in the WVS questionnaire is very similar to TM_1^{WVS} , and reads as follows: “Please tell me for each of the following statements whether you think it can always be justified, never be justified, or something in between: Claiming governments benefits to which you are not entitled”. Again, respondents are asked to evaluate on an ordered scale from “never justifiable” (1) to “always justifiable” (10). The available question in the ISSP (BM^{ISSP}) is a little bit more specific and refers to providing incorrect information: “Consider the following situations below. Do you feel it is wrong or not wrong if a person gives the government incorrect information about himself to get government benefits that he is not entitled to?”. The scale of answers ranges from “not wrong” (1) to 4 “seriously wrong” (4).²⁶

BM^{WVS} was included in each of the four survey waves. Information on respondents from 80 countries covering data from 186 country-years is available. BM^{ISSP} has been collected in 1991 and 1998. In sum, this provides data only 48 country-years.

alternative tax morale variables in the reduced samples (WVS: 0.52, ESS: 0.30) are very similar to those from the full samples discussed above.

²⁶ It should be noted that the ESS asks, “Suppose you planned to get benefits or services you were not entitled to. How many of your friends or relatives do you think you could ask for support?”. This question is not perfectly suited to capture benefit morale, but rather measures the (perceived) benefit morale among the respondent's circle of friends and acquaintances.

In order to check the correlation between the two available measurements of benefit morale, we apply the same imputation procedure as in the case of tax morale. The correlation between BM^{WVS} and BM^{ISSP} is positive (0.20, $N=41$), however, not statistically significant from zero at conventional significance levels (p-value is equal to 0.20). If we restrict our sample to observations without imputations only 5 observations are left, however, the correlation increases to 0.85.

Finally, one may also be interested in the correlation between tax morale and benefit morale.²⁷ Table 3 shows that tax morale and benefit morale are highly correlated within a survey. The correlation between tax morale in one survey, and benefit morale in another survey is, however, practically zero.

To sum up this section, we have shown that available measurements of tax morale and benefit morale are quite consistent *within* a given survey and to a lesser extent *across* surveys. While we think that all discussed variables are suitable to study the question under consideration, we have a slight preference for variables based on more general formulated survey questions. Clearly, in terms of available data points the WVS is superior to ISSP and the ESS. Unfortunately, no survey question that explicitly (or comprehensively) refers to the shadow economy is available. A major shortcoming is that no data source offers individual-level panel data on the intrinsic motivation to comply.

2.3 The Identification Problem

In the ideal case, the researcher has access to information on both variables for a random sample of people on an individual-level, combined with a large set of covariates \mathbf{X}_{it} , over time t . While it is practically impossible to obtain a data-set that fulfils all these criteria, we will for a moment assume that it exists.²⁸ This data-set would allow us to estimate an equation of the following form:

$$behaviour_{it} = \alpha + \beta \cdot attitude_{it} + \Gamma \mathbf{X}_{it} + \varepsilon_{it}. \quad (1)$$

Assuming that $attitude_{it}$ and ε_{it} are uncorrelated, an ordinary least squares regression (OLS) would give us a β^{OLS} that is equal to an unbiased estimate of the causal impact of the intrinsic motivation to comply on actual compliance behavior. Whether this necessary assumption is fulfilled, depends crucially on the set of covariates \mathbf{X}_{it} . Since OLS is a control strategy, we can

²⁷ Halla and Schneider (2008) provide a more elaborate discussion.

²⁸ Alternatively, one may think of repeated observations of countries c over time t .

increase the likelihood to observe an unbiased estimate by controlling for a large set of covariates.²⁹ In particular, \mathbf{X}_{it} has to include all variables that affect $behaviour_{it}$ and that are correlated with $attitude_{it}$. However, many determinants of $behaviour_{it}$ (such as an innate disposition to comply or aspects of socialization) are most likely unobservable. If these factors are correlated with $attitude_{it}$, the OLS-estimate is biased.

One way to mitigate this problem is controlling for individual fixed effects δ_i . These account for unobserved time-invariant individual heterogeneity. Adding in addition year fixed effects ζ_t we get the following equation:

$$behaviour_{it} = \alpha + \beta \cdot attitude_{it} + \Gamma \mathbf{X}_{it} + \delta_i + \zeta_t + \varepsilon_{it}. \quad (2)$$

The fixed effects model in (2) gives an unbiased estimate of the causal effect, as long as the relevant $attitude_{it}$ is not correlated with time-varying unobservables that affect $behaviour_{it}$, and reversed causality can be ruled out. Are these reasonable assumptions? It is hard to evaluate the case of time-varying omitted variables on a general basis. However, reversed causality (or simultaneity) cannot be ruled out, or seems almost highly likely.³⁰ While it is plausible to assume that the intrinsic motivation to comply ($attitude_{it}$) affects actual compliance behaviour ($behaviour_{it}$), it is also reasonable that actual behaviour has an impact on individuals' attitude. That means, individuals justify or confirm their own (self-interested) behaviour.³¹ If this hypothesis is true, then the fixed effects estimate from (2) is inconsistent.³²

A potential estimation strategy to solve endogeneity problems, such as this simultaneity bias, is an instrumental variable (IV) approach.³³ An IV approach can give a consistent estimate when OLS cannot. Therefore, a valid IV, let's call it z_{it} , has to be available. The IV has to be correlated with $attitude_{it}$, but uncorrelated with any other determinant of $behaviour_{it}$. The second requirement can be stated as follows: $Cov(z_{it}, \varepsilon_{it}) = 0$. Since ε_{it} is unobserved, there is no way to prove that an IV is actually valid. The researcher must rely on theoretical

²⁹ It should be noted that more control variables is not necessary better. If variables are themselves outcome variables, then they should not be included in the regression. For a detailed discussion see chapter 3 in Angrist and Pischke (2009). Therefore, in our case, we should not control for factors that are determined by the intrinsic motivation to comply.

³⁰ Another source of contemporaneous correlation between $attitude_{it}$ and ε_{it} is measurement error in the intrinsic motivation to comply.

³¹ We will discuss this hypothesis in more detail in the next section.

³² See, for instance, Wooldridge (2002) for a formal discussion.

³³ An IV approach will be especially important, if one cannot control for time-invariant unobserved heterogeneity.

justifications in order to persuade critical readers.³⁴ As usual, the hardest part is to find such a credible IV. In fact, specific advice does not exist. Angrist and Krueger (2001) recommend that “*progress comes from detailed institutional knowledge and the careful investigation of the forces at work*”.

To sum up, we argue in this section that it seems impossible to identify the causal effect of the intrinsic motivation to comply on actual compliance behavior, without an IV approach.

3. Existing Evidence: A Short but Critical Review

3.1. Tax Morale and Tax Evasion

With respect to tax evasion almost all evidence is based on survey data. A number of papers contrast self-reported tax evasion with different measurements of an intrinsic motivation to comply.³⁵ For instance, Torgler, Demir, Macintyre and Schaffner (2008) examine the relation between self-reported tax evasion and tax morale based on survey data from the U.S. and Turkey. Their regression analysis shows that a high level of tax morale is associated with a low level of tax evasion. One obvious critique of such a research design is the questionable accuracy of self-reported tax evasion data.

In order to solve this problem, some papers (e.g. Bosco and Mittone, 1997; Torgler, Schaffner and Macintyre, 2007) combine tax evasion data observed in laboratory experiments with tax morale information from post-experiment questionnaires, and confirm the findings obtained with survey data. However, it is not clear whether the answers in the questionnaire are independent from the behaviour in the experiment. Therefore, it is not clear whether the intrinsic motivation to comply with the tax law *causally* affects compliance behaviour. The correlation between these two variables can be explained by simultaneity or reversed causality. If individuals justify or confirm their own self-interested behaviour, then actual behaviour has an impact on individuals’ moral considerations (Wenzel, 2005). In line with this argument, Halla and Schneider (2008) point out that tax morale deteriorates with rising income, while benefit morale improves with rising income. The authors conclude that individuals who have comparably more opportunities and low cost to commit a certain offense develop the attitude that it is a minor offence. Rich people have comparably more opportunities to commit tax evasion; they self-servingly adjust their attitude that cheating on taxes is more or less

³⁴ For a further discussion and aspects of estimation, see, for instance, Angrist and Pischke (2009).

³⁵ Other papers, such as Kaplan and Reckers (1985); Webley, Cole and Eidjar (2001), study the relation between the perceived prevalence of tax evasion among others, and the respondent’s own self-reported compliance behavior.

justifiable. For poor people it is easier to fraudulently collect benefits; they report that benefit fraud is not a big deal.

In sum, we are not aware of any convincing empirical evidence that tax morale has a causal impact on tax compliance.

3.2. Tax Morale and the Shadow Economy

With respect to the shadow economy some papers report a negative correlation between average levels of tax morale and the size of the shadow economy: Weck (1983), Torgler (2005b) for Latin America, Alm and Torgler (2006) for the U.S. and Europe, Alm, Martinez-Vazquez, Torgler (2006) for several transition countries, and Barone and Mocetti (2009) for Italy. As in the case of tax evasion, these descriptive results allow different interpretations. First, a low level of tax morale may causally impact peoples' behaviour, which results in a bigger shadow economy. Second, a pronounced shadow economy may undermine peoples' tax morale. Or third, the correlation may just be driven by an unobserved factor, such as complex tax legislation.

Most recently, a small number of papers (Torgler and Schneider, 2007, 2009; Torgler, Schaffner and Macintyre, 2007) tries to disentangle the causal effect of tax morale on the size of the shadow economy based on an IV approach. In each case the authors use a definition of the shadow economy as suggest in Section 2.1.1. The estimates of the size of the shadow economy are based on a combination of the DYMIMIC-method and the currency demand method.³⁶ Tax morale is captured as country-averages based on a re-scaled variable from the WVS (and the *Latinobarometro*). Since each paper has a different focus, the exact sample, the set of control variables, and the suggested IV's vary. In essence, Torgler and Schneider, (2009) present a cross-sectional analysis of the effect of tax morale and institutional quality on the size of the shadow economy, where the authors try to account for the endogeneity of tax morale and institutional quality with a set of IV's, such as legal origins of commercial laws. Torgler, Schaffner and Macintyre, 2007 include a panel data analysis of the impact of tax morale on the size of the shadow economy, where weather conditions (a measure for cloudiness) serves as an IV for tax morale. Finally Torgler and Schneider (2007) employ a panel data analysis of study the effect of tax morale, institutional quality, and governance on the size of the shadow economy. To instrument for tax morale a measure of cloudiness and an index for moral values based on data from the WVS is used. All papers use a Two-Stage Least Squares estimation

³⁶ For further details all papers refer to Schneider (2005a,b).

(2SLS) and find a statistically significant negative effect of tax morale on the size of the shadow economy.³⁷ In turn we will discuss each IV approach in more detail.

In Torgler and Schneider (2009) tax morale and institutional quality is instrumented by the following variables: legal origin (English, German and French), latitude, fractionalization (language), religion (protestant, catholic), and the legal system (political rights). These IV's (or subsets of them) are widely used in the literature as a source of exogenous variation in institutions and their quality (La Porta, Lopez-de-Silanes, Shleifer and Vishny, 1999). The origin of single IV's is different. For instance, the idea to use legal origin as an IV has been suggested by La Porta, Lopez-de-Silanes, Shleifer and Vishny (1997, 1998). The authors use legal origins of commercial laws to uncover the causal effect of legal rules on financial development. It is argued that legal origin is largely exogenous, since it was typically introduced into various countries through conquest and colonization. Subsequently, a large body of research has shown that the influence of legal origins on legal rules is not restricted to finance.³⁸ However, the usage of these variables as IV's is not without critique.³⁹ Moreover, one disadvantage of these IV's is that most of them do not vary over time, and cannot be combined with a fixed effects model.

In any case, the application of these IV's for tax morale requires further discussion and should not be used *ad hoc*. A discussion of the expected and actual sign of each IV in the first stage would be informative. Further, the assumption that none of these IV's influences the size of the shadow economy through channels other than tax morale needs some support. For instance, it is unclear whether some legal origins are correlated with higher penalties for non-compliance (or higher enforcement effort), which would have an independent effect on the size of the shadow economy.

Torgler, Schaffner and Macintyre (2007) instrument tax morale with a measure for cloudiness. The authors cite literature showing that cloudiness has a negative impact on individual's well-being and they find that it has also a statistically significant negative impact on tax morale (in their first stage regression). However, it is hard to rule out that weather conditions do not have a separate effect on the size of the shadow economy. For instance, the

³⁷ The estimated quantitative impact of tax morale on the size of the shadow economy from OLS and 2SLS can unfortunately not be directly compared. In the former case standardized coefficients are listed, while in the latter case only conventional (or unstandardized) coefficients are presented.

³⁸ For a survey of this literature see, La Porta, Lopez-de-Silanes and Shleifer (2008).

³⁹ Respectively, La Porta et al. (2008) comment on page 326 “[that their] interpretation of the meaning of legal origins has evolved considerably over time [...] and that is the idea that legal origins – broadly interpreted as highly persistent systems of social control of economic life – have significant consequences for the legal and regulatory framework of the society, as well as for economic outcomes”. On page 291 they admit “[that] legal origins influence many spheres of law making and regulation, which makes it dangerous to use them as instruments”.

construction industry, a sector which is difficult to tax and known for high underground activity, is affected by weather conditions. In general, bad weather (such as rain, snow, or wind) slows down construction activity, and may also reduce the size of the shadow economy.

Torgler and Schneider (2007) suggest a (in addition to a measure of cloudiness as discussed above) an index for moral values based on data from the WVS as an IV for tax morale. This index is based on benefit morale and on the justifiability of avoiding a fare on public transport. As expected, this index enters statistically significant in the first stage regression and the first requirement of the IV is clearly fulfilled. However, under the assumption that the intrinsic motivation to comply has an impact on compliance behaviour, the second assumption could possibly fail. For instance, if a low level of benefit morale translates into higher benefit fraud, then the IV is correlated with the error term in the second stage.

We think that the papers using an IV approach to disentangle the causal effect of tax morale on the size of the shadow economy constitute a very promising direction of research. In fact, these papers are some of the rare exceptions that aim to address the causality issue between compliance attitude and compliance behaviour. We believe that a waterproof identification of this causal effect is almost impossible, since it is hard to think of a natural experiment that provides an “as-if” randomly assigned high intrinsic motivation to comply. Nevertheless, future research in this area should pursue along the lines of existing empirical strategies. However, provide a thorough discussion on the validity of the used IV’s. Murray (2006) states strikingly “*Indeed, all instruments arrive on the scene with a dark cloud of invalidity hanging overhead. This cloud never goes entirely away, but researchers should chase away as much of the cloud as they can.*”

4. The Importance of Unobserved Heterogeneity

In this section we want to evaluate the importance of time-invariant unobserved heterogeneity in the relation between the intrinsic motivation to comply and actual compliance behavior. Therefore, we study the case of tax morale and the shadow economy, or more precisely the *underground production*.

We use estimates of the size of the shadow economy from Schneider, Buehn and Montenegro (2010). This paper uses a narrow definition of the shadow economy (as suggested in Section 2.1.1) that coincides with the *underground production*. Based on a multiple indicators multiple causes (MIMIC) model estimates on the size of the underground production

for 162 countries over the period 1999 to 2007 are available.⁴⁰ If we match this data with country-averages of TM_1^{WVS} we get a sample of 52 country-years.⁴¹ In order to gain more observations (in particular observations for countries over time), we amend our data-set with estimates of the underground production for the available country-years in the WVS before 1999.⁴² In sum, we have an unbalanced panel data-set with 75 observations from 53 countries. In line with the literature we find a negative correlation between tax morale and underground production in our sample, see Figure 2. However, the correlation coefficient of minus 0.07 is not statistically significant different from zero (p-value=0.55).⁴³

[Figure 2 somewhere around here]

In Table 5 we analyze this relationship based on a series of regressions. Column 1 shows that when controlling for year fixed effects, tax morale and the size of the underground production are statistically significantly correlated. This association is lower in OECD-member states (see column 3) compared to non-member states (see column 4).⁴⁴ Most importantly, column 5 shows that once we control for country fixed effects, the estimated coefficient on tax morale remains statistically significant, but switches sign.⁴⁵ If we control for a small set of control variables (see column 5) the coefficient turns statistically insignificant, however, it stays positive.

[Table 5 somewhere around here]

It seems that tax morale is correlated with unobserved country-specific time-invariant heterogeneity in a way that disregarding country fixed effects can diametrically reverse results. The suggested positive relationship between tax morale and the size of the underground production is counter-intuitive. That means, either our measurements of tax morale and/or the

⁴⁰ Details on the estimation method are provided in Section 3.1 of Schneider, Buehn and Montenegro (2010).

⁴¹ The notes to Table 5 list these country-years.

⁴² Prof. Schneider kindly provided data on the estimated size of the underground production based on a MIMIC model for the following country-years: Australia (1995), Austria (1990), Belgium (1990), Canada (1990), Denmark (1990), Finland (1990), Finland (1996), France 1990), Germany (1990, 1997), Ireland (1990), Italy (1990), Japan (1990, 1995), Spain (1990, 1995), Sweden (1990, 1996), Switzerland (1989, 1996), United Kingdom (1990) and United States (1990, 1995).

⁴³ Notably, as Figures A1 to A4 (in the Appendix) show, the correlation between the other available tax morale variables and the estimated underground production is in each case positive. For TM_2^{ESS} and TM_2^{ISSP} the correlation coefficient is even statistically significant at conventional levels.

⁴⁴ Column 2 shows, as expected, that the underground production is on average lower in OECD-member states (about minus 14 percent of the GDP).

⁴⁵ We do not have observations over time for non-OECD-member states.

underground production have fundamental problems, or unobserved time-varying unobservables are crucial. In the latter case an IV-strategy (as described in Section 2.3) can be used to fix the problem. An ideal IV would vary over time, such that the IV strategy complements the fixed effects model.

5. Conclusions

Why should economists be interested in the (determinants of the) intrinsic motivation to comply, if this variable has no causal impact on actual compliance behavior? In this paper we argued that a good understanding of this relationship is very important for this strand of literature and further empirical evidence is needed. We think future research should pursue empirical strategies that are able to establish a causal link between these two dimensions.

We see three (potentially) feasible endeavors that could help to uncover a causal effect. First, we suggest trying to create a link between randomized in depth-audits (such as the TCMP) and panel survey data. Second, we recommend laboratory experiments augmented with well designed survey techniques. The effect of survey participation on compliance behavior could be checked with random assignment to different groups with a pre-experimental survey, a post-experimental survey, and no survey. Third, scholars should pursue along the lines of existing IV approaches, however, provide a thorough discussion of the validity of the used IV's.

References

- Alm, James and Benno Torgler (2006), 'Culture Differences and Tax Morale in the United States and Europe', *Journal of Economic Psychology* 27(2), 224–246.
- Alm, James, Jorge Martinez-Vazque and Benno Torgler (2006), 'Russian Attitudes toward Paying Taxes – before, during, and after the Transition', *International Journal of Social Economics* 33(12), 832–857.
- Andreoni, James, Brian Erard and Jonathan Feinstein (1998), 'Tax Compliance', *Journal of Economic Literature* 36(2), 818–860.
- Angrist, Joshua D. and Alan B. Krueger (2001), 'Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments', *Journal of Economic Perspectives* 15(4), 69–87.
- Angrist, Joshua D. and Jörn-Steffen Pischke (2009), *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton University Press, Princeton, NJ.
- Barone, Guglielmo and Sauro Mocetti (2009), *Tax morale and Public Spending Inefficiency*, Economic Working Paper 732, Economic Research Department, Bank of Italy.
- Bosco, Luigi and Luigi Mittone (1997), 'Tax Evasion and Moral Constraints: Some Experimental Evidence', *Kyklos* 50(3), 297–324.
- Breusch, Trevor (2005), *Estimating the Underground Economy using MIMIC Models*, Unpublished manuscript, Australian National University.
- Elffers, Henk, Russell H. Weigel and Dick J. Hessing (1987), 'The Consequences of Different Strategies for Measuring Tax Evasion Behavior', *Journal of Economic Psychology* 8(3), 311–337.
- Elster, Jon (1989), 'Social Norms and Economic Theory', *Journal of Economic Perspectives* 3(4), 99–117.
- Cowell, Frank A. (1990), *Cheating the Government*, MIT Press, Cambridge, MA.
- Erard, Brian and Jonathan S. Feinstein (1994), 'Honesty and Evasion in the Tax Compliance Game', *RAND Journal of Economics* 25(1), 1–19.
- Frey, Bruno S. and Friedrich Schneider (2001), *Informal and Underground Economy*, in N. J. Smelser and P. B. Baltes, eds, 'International Encyclopedia of Social and Behavioral Science', Elsevier Ltd., Oxford, pp. 7441–7446.
- Friedland, Nehemiah, Shlomo Maital and Aryeh Rutenberg (1978), 'A Simulation Study of Income Tax Evasion', *Journal of Public Economics* 10(1), 107–116.
- Gordon, James P. (1989), 'Individual Morality and Reputation Costs as Deterrents to Tax Evasion', *European Economic Review* 33(4), 797–805.

Gorodnichenko, Yuriy, Jorge Martinez-Vazquez and Klara Sabirianova Peter (2009), 'Myth and Reality of Flat Tax Reform: Micro Estimates of Tax Evasion Response and Welfare Effects in Russia', *Journal of Political Economy* 117(3), 504–554.

Halla, Martin and Friedrich G. Schneider (2008), *Taxes and Benefits: Two Distinct Options to Cheat on the State?*, IZA Discussion Paper 3536, Institute for the Study of Labor (IZA), Bonn.

Halla, Martin, Mario Lackner and Friedrich G. Schneider (2010), 'An Empirical Analysis of the Dynamics of the Welfare State: The Case of Benefit Morale', *Kyklos* 62(1), 55–74.

Heinemann, Friedrich (2008), 'Is the Welfare State Self-Destructive? A Study of Government Benefit Morale', *Kyklos* 61(2), 237–257.

Kaplan, Steven E. and Philip M. J. Reckers (1985), 'A Study of Tax Evasion Judgments', *National Tax Journal* 38(1), 97–102.

La Porta, Rafael, Florencio Lopez-de-Silanes and Andrei Shleifer (2008), 'The Economic Consequences of Legal Origins', *Journal of Economic Literature* 46(2), 285–332.

La Porta, Rafael, Florencio Lopez-de-Silanes, Andrei Shleifer and Robert W. Vishny (1997), 'Legal Determinants of External Finance', *Journal of Finance* 52(3), 1131–1150.

La Porta, Rafael, Florencio Lopez-de-Silanes, Andrei Shleifer and Robert W. Vishny (1998), 'Law and Finance', *Journal of Political Economy* 106(6), 1113–1155.

La Porta, Rafael, Florencio Lopez-de-Silanes, Andrei Shleifer and Robert W. Vishny (1999), 'The Quality of Government', *Journal of Law, Economics and Organization* 15(1), 222–279.

Lemieux, Thomas, Bernard Fortin and Pierre Fréchette (1994), 'The Effect of Taxes on Labor Supply in the Underground Economy', *American Economic Review* 84(1), 231–254.

Lyssioutou, Panayiota, Panos Pashardes and Thanasis Stengos (2004), 'Estimates of the Black Economy Based on Consumer Demand Approaches', *Economic Journal* 114(497), 622–640.

Murray, Michael P. (2006), 'Avoiding Invalid Instruments and Coping with Weak Instruments', *Journal of Economic Perspectives* 20(4), 111–132.

OECD (2002), *Measuring the Non-Observed Economy – A Handbook*, Technical report, Organisation for Economic Co-operation and Development, Paris.

Polinsky, Mitchell A. and Steven Shavell (2000), 'The Economic Theory of Public Enforcement of Law.', *Journal of Economic Literature* 38(1), 45–76.

Posner, Eric A. (2000), 'Law and Social Norms: The Case of Tax Compliance', *Virginia Law Review* 86(8), 1781–1819.

Posner, Eric A. (2002), *Law and Social Norms*, Harvard University Press.

Posner, Richard A. (1997), 'Social Norms and the Law: An Economic Approach', *American Economic Review* 87(2), 365–369.

Schmölders, Günter (1951/52), 'Finanzpsychologie', *Finanzarchiv* 13(1), 1–36.

Schmölders, Günter (1960), *Das Irrationale in der öffentlichen Finanzwirtschaft*, Rowohlt, Hamburg.

Schmölders, Günter (1969), Steuermentalität und Steuermoral in Großbritannien, Frankreich, Italien und Spanien, in B.Beichelt, B.Biervert, J.Daviter, G.Schmölders and B.Strümpel, eds, 'Steuernorm und Steuerwirklichkeit Band II', Westdeutscher Verlag, Köln.

Schneider, Friedrich (2005a), 'Shadow Economies Around the World: What Do We Really Know?', *European Journal of Political Economy* 21(3), 598–642.

Schneider, Friedrich (2005b), *Shadow Economies of 145 Countries all over the World: What Do We Really Know?*, CREMA Working Paper Series 13, Center for Research in Economics, Management and the Arts (CREMA).

Schneider, Friedrich, Andreas Buehn and Claudio E. Montenegro (2010), *Shadow Economies all over the World: New Estimates for 162 Countries from 1999 to 2007*, Background paper for the World Bank study of the informal sector in Central, Southern Europe and the Baltic countries (Task number P112988).

Schneider, Friedrich and Dominik Enste (2000), 'Shadow Economies: Sizes, Causes, and Consequences', *Journal of Economic Literature* 38(1), 77–114.

Slemrod, Joel and Shlomo Yitzhaki (2002), Tax avoidance, Evasion, and Administration, in A. J.Auerbach and M.Feldstein, eds, 'Handbook of Public Economics', Vol. 3, Elsevier, chapter 22, pp. 1423–1470.

Slemrod, Joel (2007), 'Cheating Ourselves: The Economics of Tax Evasion', *Journal of Economic Perspectives* 21(1), 25–48.

Tanzi, Vito (1999), 'Uses and Abuses of Estimates of the Underground Economy', *Economic Journal* 109(456), F338–F347.

Torgler, Benno (2005a), 'Tax Morale and Direct Democracy', *European Journal of Political Economy* 21(2), 521–531.

Torgler, Benno (2005b), 'Tax Morale in Latin America', *Public Choice* 122(1-2), 133–157.

Torgler, Benno (2007), *Tax Compliance and Tax Morale: A Theoretical and Empirical Analysis*, Edward Elgar, Cheltenham, UK.

Torgler, Benno and Friedrich Schneider (2007), *Shadow Economy, Tax Morale, Governance and Institutional Quality: A Panel Analysis*, IZA Discussion Paper 2563, Institute for the Study of Labor (IZA).

Torgler, Benno and Friedrich Schneider (2009), 'The Impact of Tax Morale and Institutional Quality on the Shadow Economy', *Journal of Economic Psychology* 30(2), 228–245.

Torgler, Benno, Ihsan C. Demir, Alison Macintyre and Markus Schaffner (2008), 'Causes and Consequences of Tax Morale: An Empirical Investigation', *Economic Analysis and Policy (EAP)* 38(2), 313–339.

Torgler, Benno, Markus Schaffner and Alison Macintyre (2007), *Tax Compliance, Tax Morale, and Governance Quality*, International Studies Program Working Paper, Andrew Young School of Policy Studies, Georgia State University.

Traxler, Christian (2010), 'Social Norms and Conditional Cooperative Taxpayers', *European Journal of Political Economy* 26(1), 89–103.

Webley, Paul, Michaela Cole and Ole-Petter Eidjar (2001), 'The Prediction of Self-reported and Hypothetical Tax-Evasion: Evidence from England, France and Norway', *Journal of Economic Psychology* 22(2), 141–155.

Weck, Hannelore (1983), *Schattenwirtschaft: Eine Möglichkeit zur Einschränkung der öffentlichen Verwaltung? Eine ökonomische Analyse*, in 'Finanzwissenschaftliche Schriften Nummer 22', Lang, Bern-Frankfurt.

Wenzel, Michael (2005), 'Motivation or Rationalisation? Causal relations between Ethics, Norms and Tax compliance', *Journal of Economic Psychology* 26(4), 491–508.

Wooldridge, Jeffrey M. (2002), *Econometric Analysis on Cross Section and Panel Data*, MIT Press, Cambridge, MA.

Appendix

Figure 1: Hits for “Tax Morale” on Google Scholar over Time

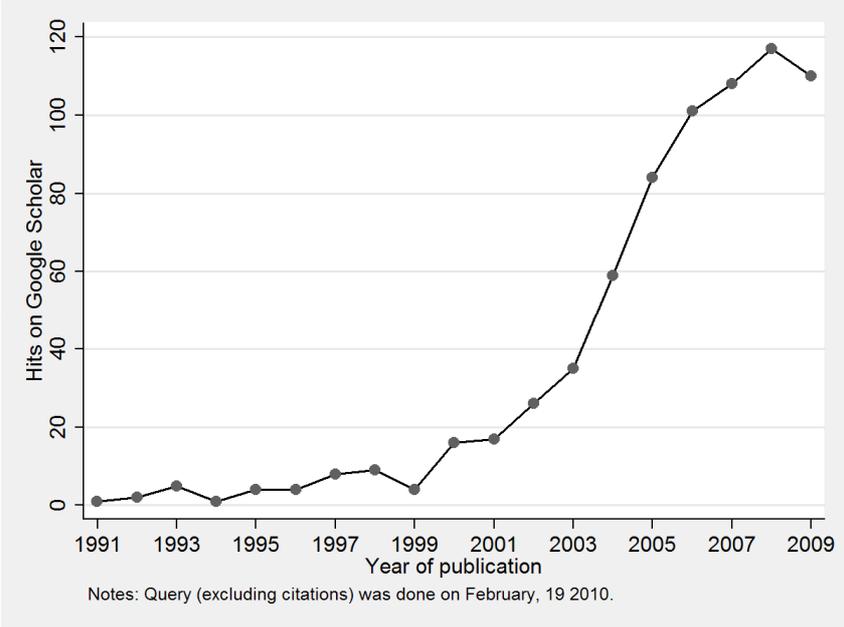
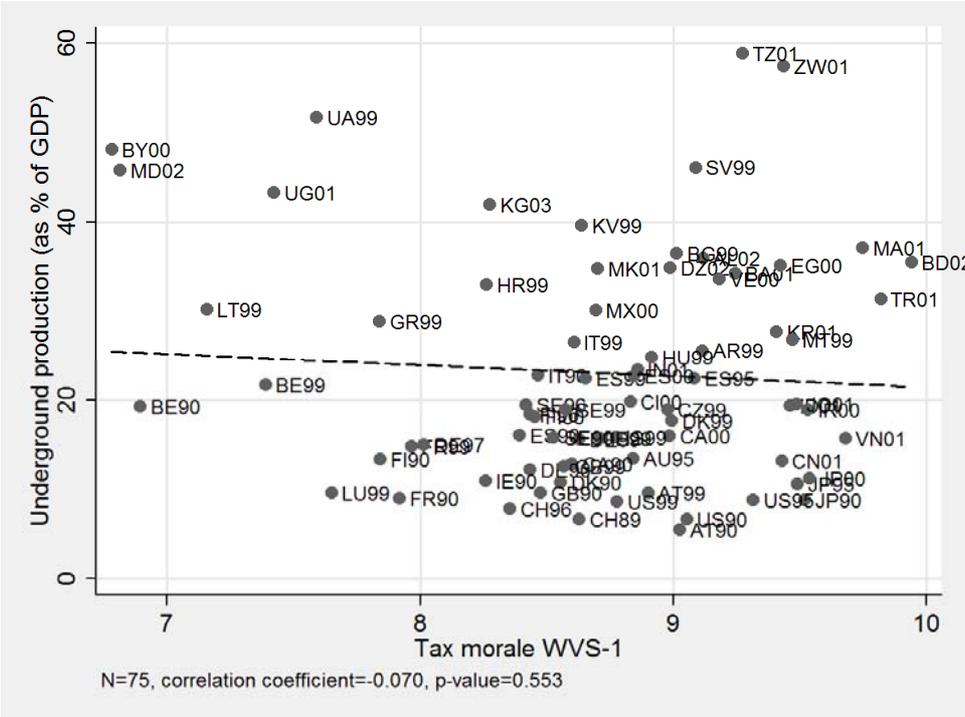
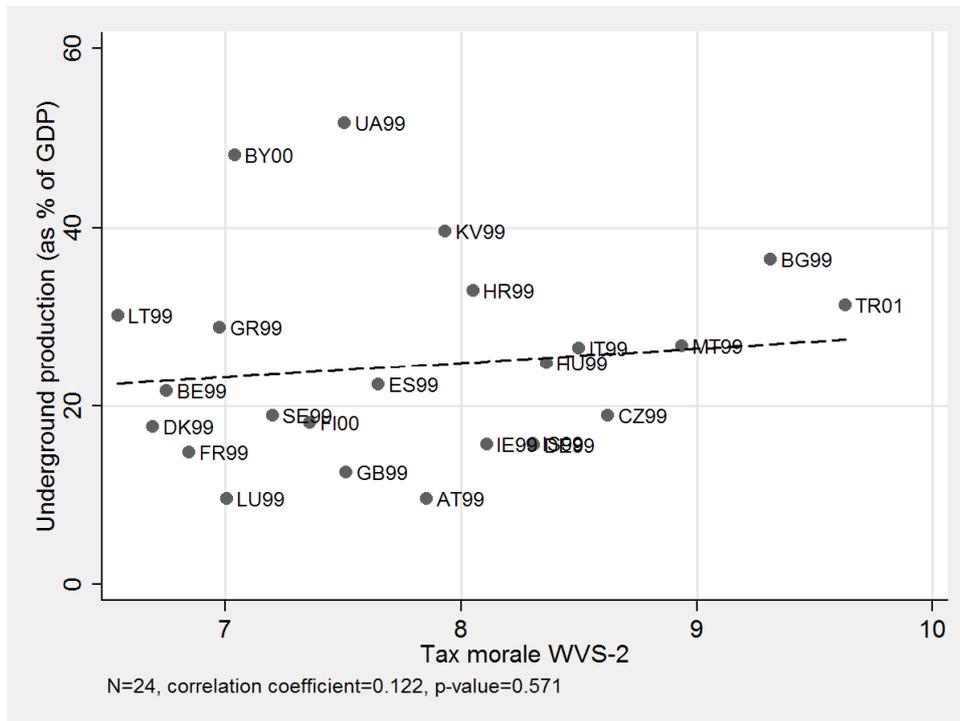


Figure 2: Relation between Underground Production and Tax Morale WVS-1



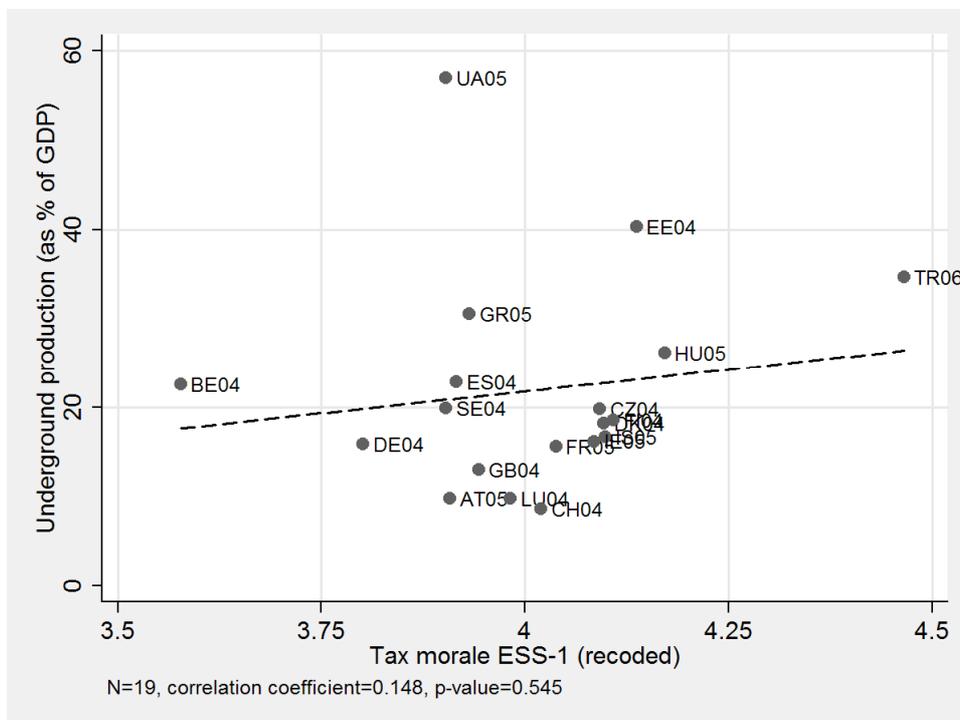
Notes: AL: Albania, AR: Argentina, AT: Austria, AU: Australia, BA: Bosnia and Herzegovina, BD: Bangladesh, BE: Belgium, BG: Bulgaria, BY: Belarus, CA: Canada, CH: Switzerland, CN: China, CZ: Czech Republic, DE: Germany, DK: Denmark, DZ: Algeria, EG: Egypt, ES: Spain, FI: Finland, FR: France, GB: United Kingdom, GR: Greece, HR: Croatia, HU: Hungary, ID: Indonesia, IE: Ireland, IN: India, IR: Iran, IS: Iceland, IT: Italy, JO: Jordan, JP: Japan, KG: Kyrgyzstan, KR: Republic of Korea, KV: Latvia, LT: Lithuania, LU: Luxembourg, MA: Morocco, MD: Republic of Moldova, MK: Republic of Macedonia, MT: Malta, MX: Mexico, SE: Sweden, SV: El Salvador, TR: Turkey, TZ: United Republic of Tanzania, UA: Ukraine, UG: Uganda, US: United States, VE: Venezuela, VN: Vietnam, ZW: Zimbabwe.

Figure A1: Relation between Underground Production and Tax Morale WVS-2



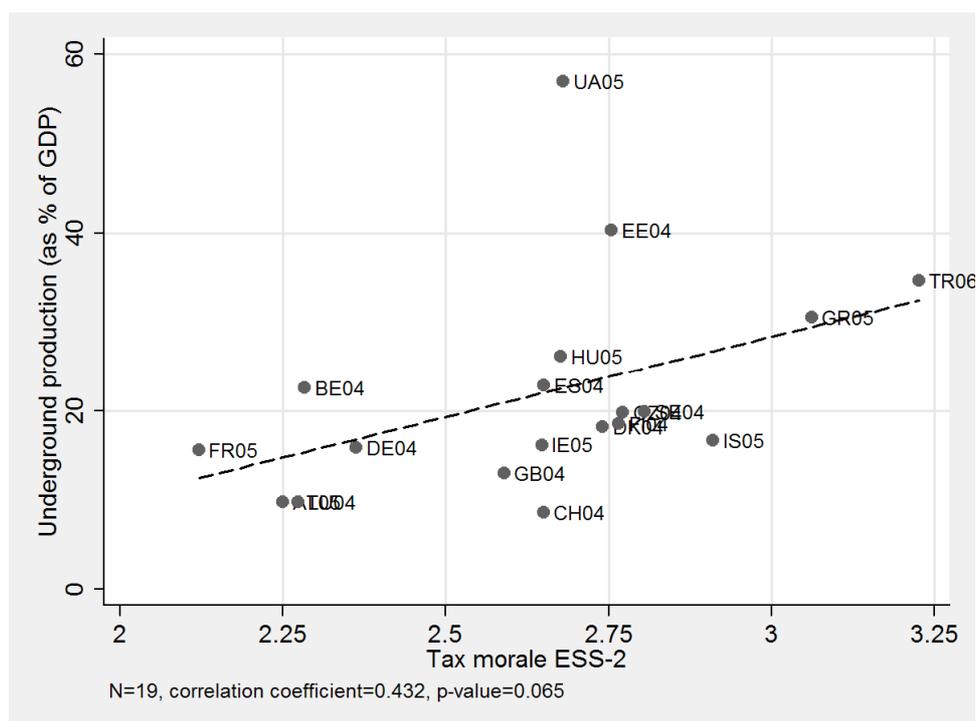
Notes: AT: Austria, BE: Belgium, BG: Bulgaria, BY: Belarus, CZ: Czech Republic, DE: Germany, DK: Denmark, ES: Spain, FI: Finland, FR: France, GB: United Kingdom, GR: Greece, HR: Croatia, HU: Hungary, IE: Ireland, IS: Iceland, IT: Italy, KV: Latvia, LT: Lithuania, LU: Luxembourg, MT: Malta, SE: Sweden, TR: Turkey, UA: Ukraine.

Figure A2: Relation between Underground Production and Tax Morale ESS-1



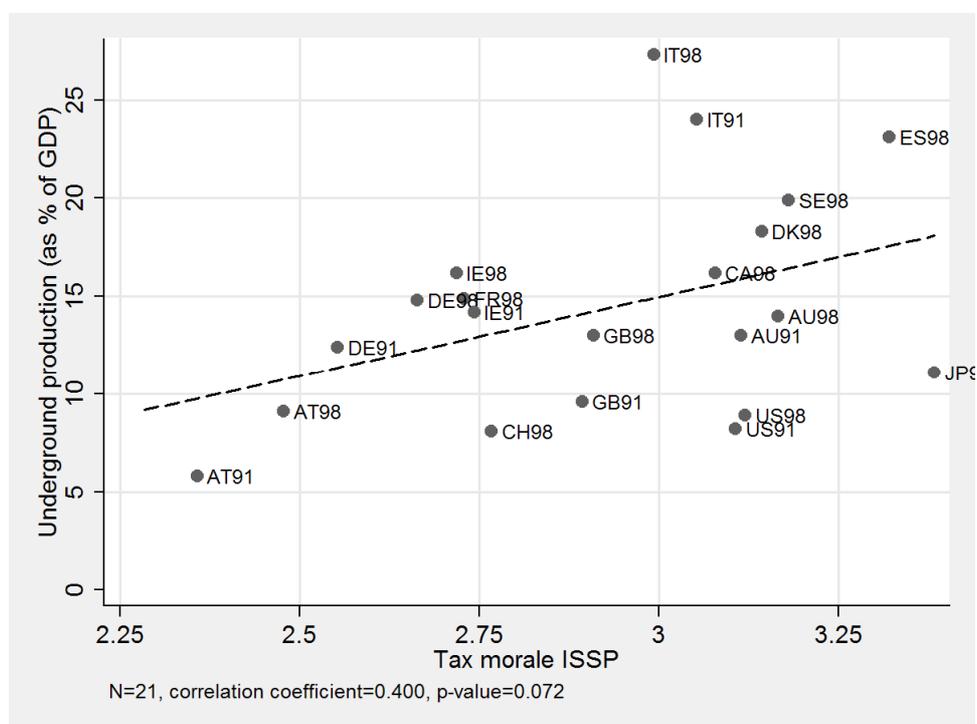
Notes: AT: Austria, BE: Belgium, CH: Switzerland, CZ: Czech Republic, DE: Germany, DK: Denmark, EE: Estonia, ES: Spain, FI: Finland, FR: France, GB: United Kingdom, GR: Greece, HU: Hungary, IE: Ireland, IS: Iceland, LU: Luxembourg, SE: Sweden, TR: Turkey, UA: Ukraine.

Figure A3: Relation between Underground Production and Tax Morale ESS-2



Notes: AT: Austria, BE: Belgium, CH: Switzerland, CZ: Czech Republic, DE: Germany, DK: Denmark, EE: Estonia, ES: Spain, FI: Finland, FR: France, GB: United Kingdom, GR: Greece, HU: Hungary, IE: Ireland, IS: Iceland, LU: Luxembourg, SE: Sweden, TR: Turkey, UA: Ukraine.

Figure A4: Relation between Underground Production and Tax Morale ISSP



Notes: AT: Austria, AU: Australia, CA: Canada, CH: Switzerland, DE: Germany, DK: Denmark, ES: Spain, FR: France, GB: United Kingdom, IE: Ireland, IT: Italy, JP: Japan SE: Sweden, US: United States.

Table A1: Available Survey Data on Tax Morale^a

Survey	Question	Answers	Years covered (no. of countries)	Total no. of countries	Total no. of country- years
WVS	TM_1^{WVS} : “Please tell me for each of the following statements whether you think it can always be justified, never be justified, or something in between: Cheating on taxes if you have a chance”.	Respondents are asked to evaluate on an ordered scale from “never justifiable” (1) to “always justifiable” (10).	1981 (12) 1982 (5) 1983 (1) 1984 (2) 1989 (2) 1990 (34) 1991 (6) 1992 (1) 1993 (1) 1994 (1) 1995 (11) 1996 (22) 1997 (8) 1998 (9) 1999 (33) 2000 (11) 2001 (19) 2002 (5) 2003 (1)	80	184
WVS	TM_2^{WVS} : “Please tell me for each of the following statements whether you think it can always be justified, never be justified, or something in between: Paying cash for services to avoid taxes”.	Respondents are asked to evaluate on an ordered scale from “never justifiable” (1) to “always justifiable” (10).	1999 (30) 2000 (2) 2001 (1)	33	33
ESS	TM_1^{ESS} : “How much you agree or disagree with each of these statements: Citizens should not cheat on their taxes”	Respondents can answer: “agree strongly” (1), “agree” (2), “neither agree nor disagree” (3), “disagree” (4), or “disagree strongly” (5).	2004 (16) 2005 (8) 2006 (1)	25	25
ESS	TM_2^{ESS} : “How wrong, if at all, do you consider the following ways of behaving to be? How wrong is someone paying cash with no receipt so as to avoid paying VAT or other taxes?”	Respondents can answer: “not wrong at all” (1), “a bit wrong” (2), “wrong” (3), or “seriously wrong” (4).	2004 (16) 2005 (8) 2006 (1)	25	25
ISSP	TM^{ISSP} : “Consider the following situations below. Do you feel it is wrong or not wrong if a taxpayer does not report all of his or her income in order to pay less income tax?”	Respondents can answer: “not wrong” (1), “a bit wrong” (2), “wrong” (3), or “seriously wrong” (4).	1991 (17) 1998 (31)	31	48

Table A2: Available Survey Data on Benefit Morale^a

Survey	Question	Answers	Years covered (no. of countries)	Total no. of countries	Total no. of country- years
WVS	<i>BM^{WVS} : “Please tell me for each of the following statements whether you think it can always be justified, never be justified, or something in between: Claiming governments benefits to which you are not entitled”.</i>	Respondents are asked to evaluate on an ordered scale from “never justifiable” (1) to “always justifiable” (10).	1981 (12) 1982 (6) 1983 (1) 1984 (2) 1989 (2) 1990 (36) 1991 (6) 1992 (1) 1993 (1) 1994 (1) 1995 (11) 1996 (21) 1997 (8) 1998 (9) 1999 (33) 2000 (11) 2001 (18) 2002 (5) 2003 (2)	81	186
ISSP	<i>BM^{ISSP} : “Consider the following situations below. Do you feel it is wrong or not wrong if a person gives the government incorrect information about himself to get government benefits that he is not entitled to”</i>	Respondents can answer: “not wrong” (1), “a bit wrong” (2), “wrong” (3), or “seriously wrong” (4).	1991 (17) 1998 (31)	31	48

Table 1: Comparison of the Survey Questions on Tax Morale in the WVS^a

	N	TM_1^{WVS}	TM_2^{WVS}	$TM_1^{WVS} - TM_2^{WVS}$	$Corr(TM_1^{WVS}, TM_2^{WVS})$
Austria	1,467	8.91	7.84	1.06	0.47
Belarus	824	6.76	7.05	-0.29	0.62
Belgium	1,833	7.37	6.74	0.63	0.45
Bulgaria	923	9.03	9.32	-0.29	0.52
Croatia	987	8.25	8.05	0.20	0.65
Czech Republic	1,849	8.98	8.62	0.35	0.55
Denmark	1,007	8.99	6.69	2.29	0.38
Estonia	901	7.83	7.73	0.11	0.60
Finland	1,017	8.45	7.37	1.08	0.55
France	1,551	7.97	6.86	1.11	0.49
Germany	1,935	8.65	8.31	0.34	0.67
Great Britain	977	8.56	7.52	1.04	0.53
Greece	1,061	7.81	6.98	0.83	0.35
Hungary	947	8.90	8.37	0.53	0.55
Iceland	957	8.77	8.30	0.46	0.58
Ireland	976	8.72	8.10	0.62	0.49
Italy	1,948	8.61	8.51	0.10	0.49
Latvia	970	8.63	7.94	0.69	0.54
Lithuania	860	7.08	6.56	0.52	0.53
Luxembourg	1,128	7.64	7.01	0.63	0.47
Malta	1,002	9.47	8.94	0.53	0.48
Netherlands	995	8.27	6.74	1.53	0.38
Northern Ireland	898	8.63	7.98	0.65	0.55
Poland	945	8.86	8.29	0.57	0.36
Portugal	954	8.60	8.72	-0.12	0.67
Romania	979	8.23	8.17	0.07	0.29
Russian	2,138	7.97	7.86	0.11	0.51
Slovakia	1,271	8.84	7.34	1.50	0.38
Slovenia	987	8.66	7.73	0.93	0.49
Spain	1,095	8.63	7.65	0.97	0.43
Sweden	1,002	8.57	7.20	1.37	0.42
Turkey	1,195	9.82	9.63	0.19	0.34
Ukraine	981	7.56	7.50	0.07	0.62
Mean	1,168	8.42	7.81	0.62	0.50

^aColumn 1 shows the available observations per country. Columns 2 to 4 list the respective country-means. Column 5 shows the correlation coefficient within a country. For the definition of TM_1^{WVS} and TM_2^{WVS} see Table A1.

Table 2: Comparison of the Survey Questions on Tax Morale in the ESS^a

	N	TM_1^{ESS}	TM_2^{ESS}	$Corr(TM_1^{ESS}, TM_2^{ESS})$
Austria	2,040	3.91	2.25	0.27
Belgium	1,728	3.58	2.28	0.31
Switzerland	2,059	4.02	2.65	0.21
Czech Republic	2,661	4.09	2.77	0.25
Germany	2,706	3.80	2.36	0.32
Denmark	1,450	4.10	2.74	0.40
Estonia	1,772	4.14	2.75	0.20
Spain	1,529	3.92	2.65	0.17
Finland	1,988	4.11	2.77	0.34
France	1,784	4.04	2.12	0.29
Great Britain	1,865	3.94	2.59	0.30
Greece	2,321	3.93	3.06	0.25
Hungary	1,420	4.17	2.68	0.25
Ireland	2,231	4.08	2.65	0.21
Iceland	558	4.10	2.91	0.37
Luxembourg	1,505	3.98	2.27	0.11
Netherlands	1,853	3.93	2.40	0.24
Norway	1,749	3.94	2.75	0.40
Poland	1,620	4.08	2.58	0.28
Portugal	1,924	4.17	3.00	0.17
Sweden	1,916	3.90	2.81	0.32
Slovenia	1,382	4.08	2.58	0.25
Slovakia	1,416	3.93	2.72	0.16
Turkey	1,713	4.47	3.23	0.20
Ukraine	1,612	3.90	2.68	0.28
Mean	1,792	4.01	2.65	0.26

^aColumn 1 shows the available observations per country. Columns 2 and 3 list the respective country-means. Column 4 shows the correlation coefficient within a country. For the definition of TM_1^{ESS} and TM_2^{ESS} see Table A1. Note, we recoded TM_1^{ESS} such that higher values indicate a higher tax morale.

Table 3: Correlation among Different Measurements of Tax Morale and Benefit Morale^a

	TM_1^{WVS}	TM_2^{WVS}	TM_1^{ESS}	TM_2^{ESS}	TM^{ISSP}	BM^{WVS}	BM^{ISSP}
TM_1^{WVS}	1						
TM_2^{WVS}	0.692*** (N=33)	1					
TM_1^{ESS}	-	-	1				
TM_2^{ESS}	-	-	0.584*** (N=25)	1			
TM^{ISSP}	0.238 (N=41) ^b	0.080 (N=20) ^b	-	-	1		
BM^{WVS}	0.403*** (N=196)	0.537*** (N=34)	-	-	-0.081 (N=41)	1	
BM^{ISSP}	0.029 (N=41)	-0.381* (N=20)	-	-	0.561*** (N=70)	0.200 (N=41) ^b	1

^aFor the definition of the variables see Table A1 and A2. *, ** and *** indicate statistical significance at the 10-percent level, 5-percent level, and 1-percent level.

Table 4: Relation of Tax Morale with Socio-economic Characteristics^a

	Dependent variable			
	TM_1^{WVS}	TM_2^{WVS}	TM_1^{ESS}	TM_2^{ESS}
Age	0.023*** (0.002)	0.025*** (0.002)	0.005*** (0.001)	0.006*** (0.001)
Female	0.338*** (0.037)	0.360*** (0.034)	0.104*** (0.014)	0.040*** (0.014)
Married ^b	0.198*** (0.048)	0.180*** (0.044)	0.066*** (0.014)	0.057*** (0.017)
Children ^c	-0.032 (0.048)	0.003 (0.065)	-0.044* (0.023)	-0.033** (0.013)
School leaving age ^d	0.012** (0.005)	0.005 (0.005)	0.008*** (0.003)	0.005* (0.003)
Household income ^e	-0.020** (0.009)	-0.044*** (0.011)	0.001 (0.003)	-0.012** (0.005)
Self-employed ^f	-0.304*** (0.084)	-0.150 (0.105)	-0.107*** (0.023)	-0.042 (0.030)
Unemployed ^f	-0.223*** (0.065)	-0.036 (0.052)	-0.021 (0.034)	-0.008 (0.029)
Out of labor force ^f	0.061 (0.048)	0.086* (0.046)	0.036** (0.016)	0.028** (0.012)
Size of place of residence ^g	-0.125*** (0.045)	-0.132** (0.051)	0.010 (0.013)	0.033** (0.012)
No. of observations	30,773	30,773	28,778	28,778
Adjusted R-squared	0.112	0.126	0.052	0.105

^aThe dependent variable is in each estimation a measure of tax morale (TM), where higher values indicate a higher level of TM. In columns 2 and 3 TM is measured on a ten point-scale; in columns 4 and 5 on a four-point scale (see Tab). Each estimation includes country and year fixed effects. Method of estimation is ordinary least squares. Standard errors (allowing for clustering by countries) are in parentheses below. *, ** and *** indicate statistical significance at the 10-percent level, 5-percent level, and 1-percent level. ^bThis is a binary variable equal to one if the individual is married, and zero otherwise. ^cThis is a binary variable equal to one if the individual is parent, and zero otherwise. ^dIn the case of the columns 2 and 3 (data from the WVS) some observations have been imputed base on the highest educational level. For detailed information please refer to the Data Appendix in Halla and Schneider (2008). ^eIn the case of columns 2 and 3 this ordinal variable is measured on a ten-point scale; in columns 4 and 5 on a twelve-point scale. ^fThe base group is equal to employed individuals. ^gThis ordinal variable is measured on a three-point scale.

Table 5: Tax Morale and the Size of the Underground Production^a

VARIABLES	(1) Full sample	(2) Full sample	(3) OECD sample	(4) Non-OECD sample	(5) OECD sample	(6) OECD sample
TM_1^{WVS}	-4.639** (1.870)	-4.225** (1.621)	-3.544* (1.779)	-5.183* (2.712)	2.829** (1.284)	2.152 (1.708)
GDP p. c. (in \$1,000) ^a						0.064 (0.329)
GDP deflator ^a						0.014 (0.221)
Population size (in mill.) ^a						-0.053 (0.052)
OECD ^b		-13.517*** (2.854)				
Constant	46.722** (19.064)	76.857*** (16.036)	43.402*** (15.299)	84.792*** (25.423)	2.863 (8.613)	-9.046 (21.243)
Year fixed effects	yes	yes	yes	yes	yes	yes
Country fixed effects	no	no	no	no	yes	yes
No. of observations	75	75	46	29	46	46
Adjusted R-squared	0.366	0.525	0.316	0.009	0.979	0.977

^aThe dependent variable is in each case an estimate of the underground production measured as percentage of GDP. Schneider, Buehn and Montenegro (2010) is the source of the following country-years: Albania (2002), Algeria (2002), Argentina (1999), Austria (1999), Bangladesh (2002), Belgium (1999), Bosnia and Herzegovina (2001), Bulgaria (1999), Belarus (2000), Canada (2000), Chile (2000), China (2001), Croatia (1999), Czech Republic (1999), Denmark (1999), El Salvador (1999), Finland (2000), France (1999), Germany (1999), Greece (1999), Hungary (1999), Iceland (1999), India (2001), Indonesia (2001), Iran (2000), Ireland (1999), Italy (1999), Japan (2000), Jordan (2001), Republic of Korea (2001), Kyrgyzstan (2003), Latvia (1999), Lithuania (1999), Luxembourg (1999), Malta (1999), Mexico (2000), Republic of Moldova (2002), Morocco (2001), Vietnam (2001), Zimbabwe (2001), Spain (1999), Spain (2000), Sweden (1999), Turkey (2001), Uganda (2001), Ukraine (1999), Republic of Macedonia (2001), Egypt (2000), United Kingdom (1999), United Republic of Tanzania (2001), United States (1999) and Venezuela (2000). The remaining country-years – Australia (1995), Austria (1990), Belgium (1990), Canada (1990), Denmark (1990), Finland (1990), Finland (1996), France 1990, Germany (1990, 1997), Ireland (1990), Italy (1990), Japan (1990, 1995), Spain (1990, 1995), Sweden (1990, 1996), Switzerland (1989, 1996), United Kingdom (1990) and United States (1990, 1995) – are imputed by estimates kindly provided by Prof. Schneider. Method of estimation is ordinary least squares. Standard are in parentheses below. *, ** and *** indicate statistical significance at the 10-percent level, 5-percent level, and 1-percent level. ^aThis variable is derived from the *OECD Factbook 2007*. ^bThis is a binary variable equal to one if the country is an OECD-member state, and zero otherwise.