

Tax Morale and Compliance Behavior: First Evidence on a Causal Link*

MARTIN HALLA

University of Linz & IZA

Forthcoming in

The B.E. Journals of Economic Analysis & Policy

(Last update: March 5, 2012)

Abstract

Recent literature on tax evasion emphasizes the importance of moral considerations to explain compliance behavior. As a consequence scholars aim to identify factors that shape this so-called tax morale. However, the causal link between tax morale and actual compliance behavior is not established yet. Exploiting exogenous variation in tax morale – given by the inherited part of tax morale of American-born from their ancestors country of origin – our Two-Stage Least Square Estimation provides first evidence on a causal effect of tax morale on the size of the underground production.

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

Keywords: Tax morale, tax evasion, tax compliance, underground production.

*Corresponding address: Martin Halla, Johannes Kepler University of Linz, Department of Economics, Altenbergerstr. 69, 4040 Linz, Austria, ph.: +43 70 2468 8706, fax: +43 70 2468 28706, email: martin.halla@jku.at. For helpful discussion and comments I would like to thank two anonymous referees, René Böheim, Johann K. Brunner, Gerald J. Pruckner, Johann Scharler, Mario Schnalzenberger and Rudolf Winter-Ebmer. This research was funded by the Austrian Science Fund (FWF): National Research Network S103, The Austrian Center for Labor Economics and the Analysis of the Welfare State. Data and codes to replicate the empirical analyses (as well as a Web Appendix with full estimation output) can be downloaded from the author's website

1 Introduction

It is a well-known fact that neo-classical models of tax compliance underpredict real-world compliance (Alm et al., 1992). Many scholars, therefore, concluded that the explanation for the tendency to comply must be that individuals are obeying a norm (Posner, 2000).¹ 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 try to quantify the intrinsic motivation to pay taxes (usually called *tax morale*) with survey data.

Typically, these papers try to identify factors (both on an individual- or on a more aggregated-level) that are correlated with a high level of tax morale.² By now, a substantial amount of empirical evidence on the association between tax morale and several socio-demographic characteristics from national and international samples is available. Further, a number of papers study different institutional arrangements that are correlated with a high level of tax morale. Clearly, this strand of literature presumes that tax morale affects actual compliance. But tax morale captured by survey data, does not measure individual *behavior* but an individual *attitude*. A high tax morale does not necessarily translate into a high level of tax compliance. However, the existence of such a causal link determines the significance of this literature. This causal link is not established yet. In fact, relatively little empirical evidence (to be discussed below) on the impact of tax morale on actual compliance behavior exists.

As a matter of fact it is extremely hard to identify this causal link. First, one has to obtain and quantify both dimensions. In the case of tax compliance, this is a non-trivial problem, since any form of non-compliance is difficult to observe. Naturally non-compliant tax payers will try to hide their deviant behavior in order to avoid punishment. That means, tax evasion cannot be observed and has to be estimated. Tax morale has to be obtained by adequate survey techniques. Given that one can observe and link both dimensions, a credible research design to establish a causal effect is needed.

On an individual-level a number of papers use survey data to contrast self-reported tax evasion with different measurements of an intrinsic motivation to comply with the tax code. One obvious critique of this research design is the questionable accuracy of self-reported tax evasion information. In fact, Elffers et al. (1987) show that the correlation between actual and self-reported tax

¹Kleven et al. (2011) demonstrate that the ability to cheat (i.e. self-reporting versus third-part reporting of income) is another important dimension of real world compliance, which has been ignored in the standard model so far.

²For a survey of this literature, see, for instance, Torgler (2007).

evasion is essentially zero.³ In order to solve this problem, scholars (e. g. Bosco and Mittone, 1997; Torgler et al., 2007) combine tax evasion data obtained in laboratory experiments with tax morale information from post-experiment questionnaires. As usual, the shortcoming of these 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. Even more importantly, the answers in the questionnaire may not be independent from the behavior in the experiment. The correlation between these two variables can be driven by reversed causality. Individuals may simply justify or confirm their own self-interested behavior (Wenzel, 2005). Put differently, in this case, actual behavior causally affects tax morale, but not *vice versa*.

On a more aggregate-level a number of papers present negative correlations between the level of tax morale and the size of the shadow economy: Weck (1983), Torgler (2005*b*) for Latin America, Alm and Torgler (2006) for the U.S. and Europe, Alm et al. (2006) for several transition countries, and Barone and Mocetti (2009) for Italy. This descriptive evidence allows several interpretations and does not necessarily imply a causal effect of tax morale on the size of the shadow economy. A confounding factor, such as administrative inefficiencies in the tax system, could be the driving force. There is also concern for potential reversed causality; the size of the shadow economy may affect tax morale.

Most recently, Torgler and Schneider (2007, 2009); Torgler et al. (2007) address this identification problem and suggest instrumental variable approaches to disentangle a causal effect. Each paper has a different focus and the suggested instrumental variables vary.⁴ All papers use a Two-Stage Least Squares estimation and find a statistically significant negative effect of tax morale on the size of the shadow economy. However, as discussed by Halla (2011), in each case there are plausible arguments for the invalidity of the instruments.

³The authors managed to link tax audit data for approximately 700 Dutch taxpayers with survey responses.

⁴Torgler 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 instrumental variables, such as legal origins of commercial laws. Torgler et al. (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 instrument. Torgler and Schneider (2007) employ a panel data analysis to 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 *European and World Values Surveys* is used.

For instance, it is hard to rule out that the instrumental variable ‘weather conditions’ does not have a direct effect on the size of the shadow economy. The construction industry, a sector which is difficult to tax and known for high underground activities, is affected by weather conditions. In general, bad weather slows down construction activity, and may also reduce the size of the shadow economy. Moreover, the link between weather conditions and tax morale (i. e. the first stage) has no theoretical grounds and seems rather *ad hoc*.

In this paper we also study the link between tax morale and the shadow economy (or more precisely the underground production). To deal with potential endogeneity problems we suggest to exploit exogenous variation in tax morale given by the inherited part of tax morale. In particular, we follow an approach inspired by Rice and Feldman (1997) and estimate the tax morale inherited by the American-born from their ancestors country of origin. We show that tax morale of second-generation Americans are mainly and significantly influenced by the country of origin of their ancestors. This phenomenon can be explained by a causal effect of inherited moral values. We argue then that this inherited part is not affected by the current economic and institutional environment. Consequently, we use the *inherited part of tax morale* as an instrument for current tax morale in the home country.⁵

2 The link between tax morale and compliance behavior

To measure tax morale we use data from the *European and World Values Surveys* (WVS). In particular, we use responses to the following survey question: ‘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 this statement on an ordered scale from ‘never justifiable’ (1) to ‘always justifiable’ (10). As discussed by Halla (2011) this survey question is the best available source to measure tax morale. Compared to other survey questions, its formulation is quite general, and a reasonable number of respondents from a large set of countries over time is available.

⁵A similar approach is used by Algan and Cahuc (2009) to study the effect of civic virtue on the design of unemployment benefits and employer protection in member states of the *Organisation for Economic Co-operation and Development* (OECD). See, also Algan and Cahuc (2010).

To measure (non-)compliance behavior we suggest to use an estimate of the so-called *underground production*. The OECD defines the underground production as ‘[...] 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: (i) to avoid payment of income, value added or other taxes; (ii) to avoid the payment of social security contributions; (iii) to avoid having to meet certain legal standards such as minimum wages, maximum hours, safety or health standards, etc.; (iv) to avoid complying with certain administrative procedures, such as completing statistical questionnaires or other administrative forms’ (OECD, 2002, page 37). The underground production, therefore, comprises important methods of tax evasion, such as under-reporting of income. However, it excludes others that are non-productive, such as over-claiming deductions. Further, it covers other non-compliant behavior such as fraudulent receipt of unemployment benefits or infringement of employment regulations.

While in principal it would be preferable to contrast tax morale with an explicit estimate of tax evasion, this is not possible due to data limitations. We are not aware of consistent estimates of tax evasion for a set of countries over time. Whereas estimates of the underground production are available. Our primary data source is Schneider et al. (2010). This *World Bank* report provides estimates on the size of the underground production for a large set of countries over the period from 1999 to 2007 based on a Multiple Indicators Multiple Causes (MIMIC) model.⁶

2.1 Descriptive evidence

In a first step we use a very simple measure of tax morale and calculate country-averages based on individual-level data from the WVS. We combine all available country-years from the WVS from the years 1999 through 2004 with the estimated size of the underground production (UP) from Schneider et al. (2010). This gives us a sample of 64 countries.⁷ In line with the existing literature we find a negative correlation between the level of tax morale and the estimated size of the UP.

⁶For a comprehensive review of the available methods to estimate the size (of parts) of the shadow economy, see, Schneider and Enste (2000). As any other method, the MIMIC model is not without critique (see, e.g. Breusch, 2005; Dell’Anno and Schneider, 2006). However, estimates based on this method have been successfully used in recent empirical applications such as Johnson et al. (1998); Friedman et al. (2000).

⁷The average size of the UP as percentage of the official GDP is in the 27 OECD member countries 19.30 and in the 37 non-member states 35.92.

Table 1: OLS- & fixed effects estimations of the impact of tax morale on the UP^a

| | (OLS-I) Sample 1999-2003 | (OLS-II) Sample 1999-2003 | (OLS-III) Sample 1989-2003 | (FE-I) Sample 1989-2003 |
|--------------------------------|---------------------------------|---------------------------------|----------------------------------|----------------------------------|
| Tax morale ^b | -5.052** (2.320) [-0.274] | -4.172** (1.968) [-0.227] | -5.096*** (1.362) [-0.280] | -6.383*** (2.309) [-0.340] |
| OECD member state ^c | | -14.921*** (3.028) | -11.085*** (1.963) | |
| Constant | Yes | Yes | Yes | Yes |
| Year fixed effects | Yes | Yes | Yes | Yes |
| Country fixed effects | No | No | No | Yes |
| Mean (s.d.) of UP | 28.91 (13.14) | 28.91 (13.14) | 23.65 (12.72) | 23.65 (12.72) |
| Mean (s.d.) of tax morale | 8.74 (0.71) | 8.74 (0.71) | 8.66 (0.69) | 8.66 (0.69) |
| No. of observations/R-squared | 64/0.21 | 64/0.45 | 110/0.57 | 110/0.95 |

^a The dependent variable is equal to the estimated size of the underground production (UP) as percentage of the official GDP. The sample 1999-2003 uses estimates of Schneider et al. (2010) for the following country-years (country abbreviations follow the ISO 3166-1-alpha-2 code elements): AL02, AR 99, AT 99, BA01, BD02, BE99, BG99, BY00, CA00, CL00, CN01, CZ99, DE99, DK99, DZ02, EG00, ES00, ES99, FI00, FR99, GB99, GR99, HR99, HU99, ID01, IE99, IN01, IR00, IS99, IT99, JO01, JP00, KG03, KR01, KV99, LT99, LU99, MA01, MD02, MK01, MT99, MX00, NG00, NL99, PE01, PH01, PK01, PL99, PT99, RO99, RU99, SE99, SI99, SK99, SV99, TR01, TZ01, UA99, UG01, US99, VE00, VN01, ZA01, ZW01. The sample 1989-2003 includes in addition estimates of Johnson et al. (1997) for BG90, BY90, CZ90, EE90, HU91, LT90, LV90, PL89, PL90, RO93, RU90, RU95, SI92, SI95, SK90, SK91; of Loayza (1996) for AR90/93, BR90/93, CL90/93, MX90/93; of Bajada and Schneider (2005) for CN94/95, IN89/90, IN94/95, KR89/90, TW94/95; and of Schneider (2005) for AT89/90, AU94/95, BE89/90, CH89/90, DE89/90, DE97/98, DK89/90, ES89/90, ES94/95, FI89/90, FR89/90, GB89/90, IE89/90, IT89/90, JP94/95, NL89/90, NO89/90, NO89/90, NZ97/98, PT89/90, SE89/90, US89/90, US94/95. Method of estimation is ordinary least squares. The table shows estimated coefficients with standard error in round parenthesis below. In squared brackets beta coefficients are reported. *, ** and *** indicate statistical significance at the 10-percent level, 5-percent level, and 1-percent level, respectively. ^bThis variable is equal to country-averages based on individual-level data from the *European and World Values Surveys*. ^cThis is a binary variable equal to one if the country is a OECD member state, and zero otherwise.

The correlation coefficient of minus 0.13 is quite small. However, the relation increases considerably in an OLS regression framework. Specification OLS-I in Table 1 controls for year fixed effects. The beta coefficient is equal to minus 0.274.⁸ That means, an increase in the level of tax morale by one standard deviation (0.71 points) is associated with a decrease of the UP by 0.27 standard deviations. This is equivalent to an reduction of 3.6 percentage points. This effect is statistically significant at the 3.4 percent level. When we in addition control for OECD membership (see OLS-II), we find a somewhat lower coefficient. This simple estimation model explains about 45 percent of the variation in the UP.

These OLS estimates suggest that tax morale may have a positive effect on compliance behavior. However, the estimated coefficients are not very reliable, since a correlation between tax morale and relevant unobserved factors (such as the quality of institutions) has to be expected. One way to mitigate the problem of unobserved heterogeneity is given by country fixed effects. In order to obtain longitudinal information we amend our data set with information from four other sources.⁹ This gives us an unbalanced panel data set covering 70 countries between the years 1989 and 2000.¹⁰ A cross-section analysis of this bigger sample (see specification OLS-III) shows again a strong correlation between tax morale and the size of the UP. The beta coefficient is equal to minus 0.280. Notably, the size of the coefficient did almost not change due to the amendment of the UP data by different sources. Specification FE-I includes country fixed effects. The effect of tax morale on the size of the UP persists. After controlling for unobserved country-specific time-invariant heterogeneity, the estimated effect even increases in size and statistical significance. We obtain a beta coefficient of minus 0.340.

However, before we can conclude that tax morale does reduce the size of the UP, we have to evaluate the credibility of the fixed effects estimates. The fixed effects model gives an unbiased estimate of the causal effect of tax morale, as long as it is not correlated with time-varying unobservables that affects the UP, and reversed causality can be ruled out. However, these assumptions seem rather strong. For instance, both tax morale and compliance behavior

⁸In order to allow a comparison of the quantitative importance of tax morale across different specifications we report in all estimation tables beta coefficients on tax morale in squared brackets below the standard errors.

⁹Johnson et al. (1997) provides estimates for Central and Eastern European countries, Loayza (1996) for South American countries, Bajada and Schneider (2005) for Asian countries, and Schneider (2005) for other OECD member states.

¹⁰For almost half of the countries at least two observations in time are available. Details on the country-years included are provided in the notes to Table 1.

may be influenced by varying factors, such as quality of institutions. Or, as argued above it is also plausible that compliance behavior has an impact on individuals’ attitude. That means, individuals justify or confirm their own (self-interested) behavior. If one of these hypotheses holds, the fixed effects estimate is inconsistent. In order to account for these potential endogeneity problems, we suggest in the next section an instrumental variable approach. This allows us to estimate the causal effect of tax morale consistently and free from asymptotic bias from unobserved time-varying heterogeneity.

2.2 Evidence from an instrumental variable approach

Both tax morale, and the UP are clearly affected by the current economic and institutional environment of the country in which people live and pay taxes. Our instrumental variable (IV) approach is based on the idea that tax morale, as any other moral value or social norm, is in addition also partly inherited over generations. This inherited part in tax morale should not be, or at least not instantaneously, affected by the economic and institutional environment. Therefore, a quantification of the inherited part in tax morale could serve as an IV for current tax morale.

In order to obtain an estimate of the inherited part of tax morale (that is not confounded by the current economic and institutional environment), we use the inherited tax morale by American-born individuals from their ancestors country of origin. In particular, we use responses to the following tax morale question from the American *General Social Survey* (GSS): ‘*Consider the situations listed below. Do you feel it is wrong or not wrong if a taxpayer does not report all of his income in order to pay less income taxes*’. The possible answers categories are ‘*not wrong*’ (1), ‘*a bit wrong*’ (2), ‘*wrong*’ (3) and ‘*seriously wrong*’ (4). We create a binary variable equal to one if the respondent answered ‘*seriously wrong*’ or ‘*wrong*’, and zero otherwise.¹¹ We explain the variation in this measure of tax morale with the following linear probability model,

$$TM_{itc}^{GSS} = \alpha^{GSS} + \beta^{GSS} \cdot \mathbf{X}_{it}^{GSS} + \gamma_t^{GSS} + \delta_c + \varepsilon_{itc}^{GSS}, \quad (1)$$

where TM_{itc}^{GSS} represents the tax morale of American-born respondent i in year t whose ancestors came from country c . We control for basic socio-demographic characteristics \mathbf{X}_{it} (comprising information on sex, age, marital status and religious denomination) and the year of the survey γ_t . The question on tax morale was included in the GSS in the years 1991 and 1998. Our main variables of interest are the binary variables δ_c representing the countries of origin

¹¹The upper panel of Figure 1 shows the distribution of the initial answer categories.

Figure 1: Distribution of answer categories (tax morale questions)

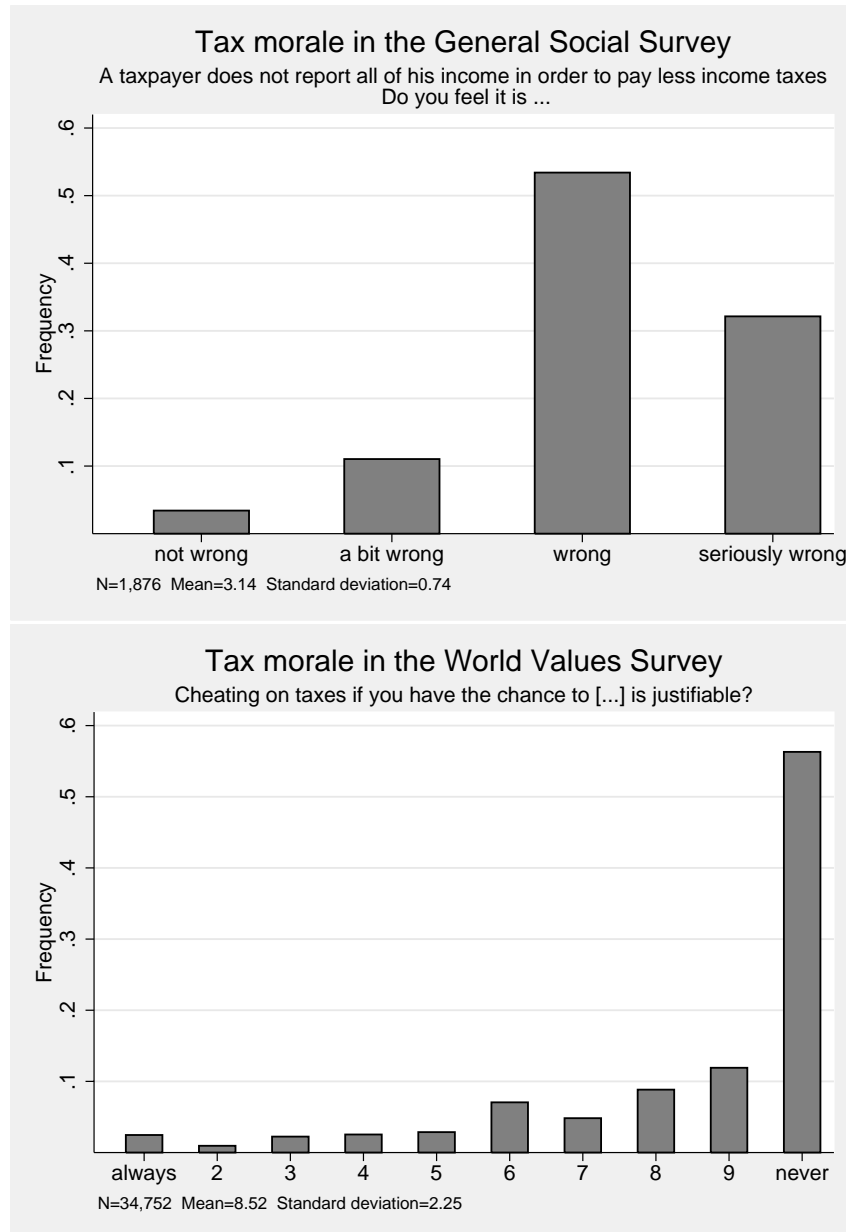


Table 2: Descriptive statistics on the tax morale measurements

| Country c | TM^{GSS^a} | | TM^{WVS^d} | |
|----------------|--------------|-------------------|--------------|-------------------|
| | N^b | Mean ^c | N^e | Mean ^f |
| Austria | 18 | 0.80 | 1,475 | 0.86 |
| Belgium | 6 | 0.61 | 1,874 | 0.69 |
| Canada | 59 | 0.85 | 1,890 | 0.85 |
| China | 13 | 0.95 | 973 | 0.92 |
| Czech Republic | 31 | 0.86 | 1,867 | 0.88 |
| Denmark | 17 | 0.80 | 1,017 | 0.87 |
| Finland | 95 | 0.79 | 988 | 0.82 |
| France | 78 | 0.74 | 1,543 | 0.76 |
| Germany | 452 | 0.78 | 1,986 | 0.83 |
| Greece | 8 | 0.70 | 1,052 | 0.78 |
| Hungary | 20 | 0.88 | 971 | 0.85 |
| Ireland | 262 | 0.76 | 968 | 0.84 |
| Italy | 107 | 0.72 | 1,956 | 0.83 |
| Japan | 7 | 0.97 | 1,213 | 0.91 |
| Mexico | 76 | 0.80 | 1,440 | 0.84 |
| Netherlands | 39 | 0.92 | 996 | 0.81 |
| Norway | 44 | 0.80 | 1,118 | 0.81 |
| Philippines | 14 | 0.85 | 1,177 | 0.76 |
| Poland | 78 | 0.69 | 1,068 | 0.84 |
| Portugal | 7 | 0.52 | 974 | 0.81 |
| Romania | 4 | 0.88 | 1,080 | 0.77 |
| Russia | 34 | 0.79 | 2,372 | 0.78 |
| Spain | 25 | 0.74 | 1,162 | 0.84 |
| Sweden | 26 | 0.81 | 1,009 | 0.85 |
| Switzerland | 17 | 0.76 | 1,070 | 0.80 |
| United Kingdom | 339 | 0.82 | 937 | 0.83 |
| | 72 | 0.79 | 1,314 | 0.82 |

^a This measure of tax morale is based on individual responses to the following question from the American *General Social Survey* of the years 1991 and 1998: ‘Consider the situations listed below. Do you feel it is wrong or not wrong if a taxpayer does not report all of his income in order to pay less income taxes’. The possible answers categories are ‘not wrong’ (1), ‘a bit wrong’ (2), ‘wrong’ (3) and ‘seriously wrong’ (4). ^b Number of available observations per country of origin. ^c These figures give the share of American-born respondents with ancestors from country c who answered ‘seriously wrong’ or ‘wrong’. ^d This measure of tax morale is based on individual responses to the following question from the *European and World Values Surveys* of the years 1999/2000: ‘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 this statement on an ordered scale from ‘never justifiable’ (1) to ‘always justifiable’ (10). For the following countries no survey data for the year 1999 or 2000 was available and we used information from the years in brackets: China (2001), Norway (1996), Philippines (2001) and Switzerland (1996). ^e Number of available observations per home country. ^f These figures give the share of respondents from country c who answered between (6) and (10).

c of the American-born respondent i . We argue that these binary variables capture the inherited part of tax morale transmitted from country of origin, which is passed down through generations. The information on the country of origin is based on the following question: ‘*From what countries or part of the world did your ancestors come?*’.¹² Our sample of 1,876 American-born respondents comprises 26 different countries of origin.¹³ As Table 2 shows we have on average 72 observations per country of origin available; the median is equal to 29. The number of available respondents varies substantially across countries of origin from only 4, as in the case of Romania, to 452 for Germany. Nevertheless, we will see that all estimated country fixed effects are highly statistically significant. In Section 2.2.1 we will show that excluding the five countries (Belgium, Greece, Japan, Portugal and Romania) with the least number of observations does not change our qualitative results.

To get comparable estimates of tax morale in home countries we create a binary variable based on the tax morale question from the WVS (mainly from the years 1999/2000).¹⁴ Table 2 provides descriptive statistics. In line with our classification for the GSS, we divide the scale of possible answers in half. Hence, we create a binary variable equal to one if the respondent answered between (6) and (10), and zero otherwise.¹⁵ We then use an equivalent linear probability model:

$$TM_{itc}^{WVS} = \alpha^{WVS} + \beta^{WVS} \cdot \mathbf{X}_{it}^{WVS} + \zeta_c + \varepsilon_{itc}^{WVS}, \quad (2)$$

where TM_{itc}^{WVS} stands for the tax morale of respondent i in year t from home country c . Here, the variables of main interest are the binary variables ζ_c , representing the home country c of respondent i . The rest of the estimation model is equivalent to (1).

Table 3 summarizes the estimation results from (1) and (2). Irrespective of the tax morale variable used, we find (as documented in the literature) that being female, older, married, and belonging to any religious denomination is related to a higher tax morale. The binary variables capturing the country of origin and the home country are all highly statistically significant.

¹²In the case respondents named more than one country, we selected the first mentioned country other than the USA.

¹³Broad categories such as ‘Africa’ or ‘Asia’ have been excluded. The remaining 26 countries restrict our sample for the subsequent analysis. Compared to the estimation analysis in Table 1 we have to exclude a number of countries, since there is no information available to estimate the inherited part of tax morale.

¹⁴If in a country no survey has taken place in the year 1999 or 2000, we use data from the year 2001 or 1996. Details are provided in the notes to Table 2.

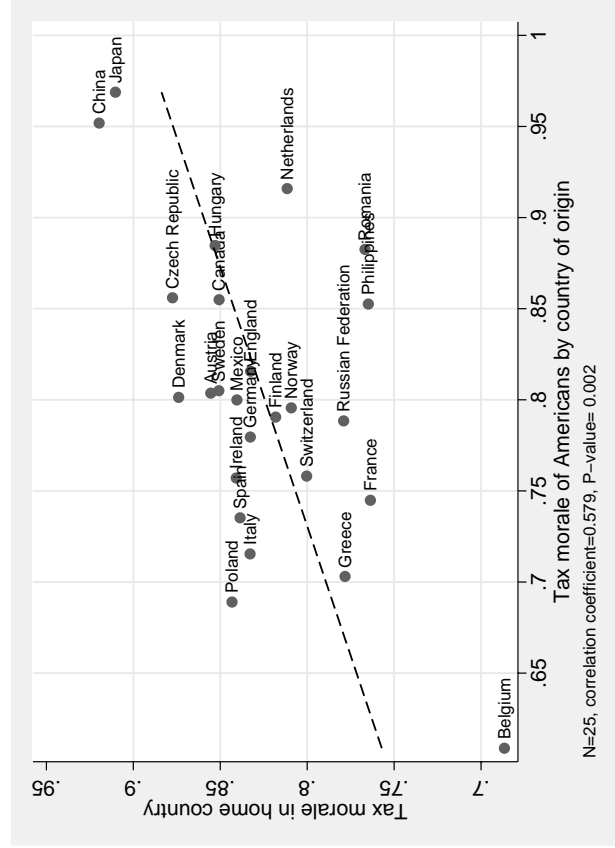
¹⁵The lower panel of Figure 1 shows the distribution of the initial answer categories.

Table 3: Tax morale by country of origin and by the home country^a

| | (I) ^b TM^{GSS} | | (II) ^b TM^{WVS} | |
|------------------------|-----------------------------|---------|------------------------------|---------|
| Male | -0.036*** | (0.012) | -0.034*** | (0.004) |
| Age | 0.002*** | (0.000) | 0.002*** | (0.000) |
| Married | 0.042** | (0.016) | 0.020*** | (0.006) |
| Religious ^c | 0.082** | (0.032) | 0.021*** | (0.006) |
| Year 1998 | -0.002 | (0.010) | | |
| δ_c/ζ_c : | | | | |
| Austria | 0.804*** | (0.027) | 0.855*** | (0.012) |
| Belgium | 0.609*** | (0.024) | 0.687*** | (0.012) |
| Canada | 0.855*** | (0.023) | 0.851*** | (0.012) |
| China | 0.952*** | (0.024) | 0.920*** | (0.011) |
| Czech Republic | 0.856*** | (0.023) | 0.877*** | (0.012) |
| Denmark | 0.801*** | (0.025) | 0.874*** | (0.012) |
| Finland | 0.790*** | (0.025) | 0.818*** | (0.011) |
| France | 0.745*** | (0.024) | 0.764*** | (0.012) |
| Germany | 0.780*** | (0.023) | 0.833*** | (0.012) |
| Greece | 0.703*** | (0.020) | 0.778*** | (0.009) |
| Hungary | 0.885*** | (0.026) | 0.853*** | (0.012) |
| Ireland | 0.757*** | (0.022) | 0.841*** | (0.012) |
| Italy | 0.715*** | (0.022) | 0.833*** | (0.011) |
| Japan | 0.969*** | (0.019) | 0.910*** | (0.012) |
| Mexico | 0.800*** | (0.019) | 0.840*** | (0.009) |
| Netherlands | 0.916*** | (0.025) | 0.811*** | (0.012) |
| Norway | 0.796*** | (0.025) | 0.809*** | (0.011) |
| Philippines | 0.853*** | (0.019) | 0.765*** | (0.010) |
| Poland | 0.689*** | (0.021) | 0.843*** | (0.012) |
| Portugal | 0.517*** | (0.024) | 0.805*** | (0.012) |
| Romania | 0.883*** | (0.028) | 0.767*** | (0.012) |
| Russia | 0.788*** | (0.023) | 0.779*** | (0.012) |
| Spain | 0.735*** | (0.020) | 0.839*** | (0.011) |
| Sweden | 0.805*** | (0.026) | 0.851*** | (0.011) |
| Switzerland | 0.758*** | (0.023) | 0.800*** | (0.012) |
| United Kingdom | 0.816*** | (0.025) | 0.833*** | (0.011) |
| No. of obs./R-squared | 1,876/0.86 | | 34,176/0.89 | |

^a Method of estimation is ordinary least squares. The table shows estimated coefficients with robust standard errors (allowing for clustering by country and heteroskedasticity of unknown form) in round parenthesis below. *, ** and *** indicate statistical significance at the 10-percent level, 5-percent level, and 1-percent level, respectively. ^b For a definition of the dependent variables TM^{GSS} and TM^{WVS} see the notes to Table 2. ^c This binary variable is equal to one if the respondent is a member of any religious denomination, and zero otherwise.

Figure 2: Correlation between tax morale of Americans by country of origin and tax morale in the home country^a



^a This figure depicts the correlation between tax morale of Americans by country of origin (i.e. the estimated country fixed effects from equation 1, $\hat{\delta}_c$, horizontal axis) and the tax morale in the home country (i.e. the estimated country fixed effects from equation 2, $\hat{\zeta}_c$, vertical axis).

The country fixed effects associated with the home country ζ_c are in general quantitatively more important than those of the country of origin δ_c . However, in the latter case the standard deviation of the coefficients is twice as high as in former case ($sd_{\zeta_c} = 0.049$, $sd_{\delta_c} = 0.099$). Both facts indicate that living in the USA has a homogenizing effect on tax morale.

Most importantly, Figure 2 shows the inherited part of tax morale, captured by $\hat{\delta}_c$, is highly correlated with the tax morale in the home country $\hat{\zeta}_c$. The correlation coefficient is equal to 0.58 and highly statistically significant (p-value= 0.002).¹⁶ This indicates that an intergenerational transmission of tax morale (that persists across space) takes place.¹⁷ We now explore the effect of tax morale on the size of the UP based on a Two-Stage Least Square (2SLS) estimation. In particular, we estimate

$$UP_c = \theta_0 + \theta_1 \cdot \hat{\zeta}_c + \theta_2 \cdot \text{Historical } UP_c + \epsilon_c, \quad (3)$$

where UP_c denotes the size of the UP of country c in the year 2000, and $\hat{\zeta}_c$ is derived from the first stage regression of tax morale in the home country on tax morale of Americans by their country of origin:

$$\hat{\zeta}_c = \pi_0 + \pi_1 \cdot \hat{\delta}_c + \pi_2 \cdot \text{Historical } UP_c + v_c. \quad (4)$$

Our 2SLS estimation strategy is valid if $\hat{\delta}_c$ is correlated with $\hat{\zeta}_c$, but uncorrelated with any other unobserved determinant of UP_c . The first requirement can be tested. It turns out that we have a very strong first stage. The second requirement, which can be stated as follows $Cov(\hat{\delta}_c, \epsilon_c) = 0$, is not testable, since ϵ_c can not be observed.

A potential concern is that if the level of the UP is correlated over time, and past levels of the UP affected the tax morale of people who came to the USA, and consequently still influence the tax morale passed on their descendants. In order to fully solve this problem it would be ideal to control for the level of the UP from the exact date before the ancestors left their home country. This approach can not be perfectly implemented, since the date of migration is not known and it varies across observations.

¹⁶If we use the original scaling of the tax morale variables a strong correlation pattern between the two alternative measures still holds. The correlation coefficient of 0.351 has a p-value of 0.086.

¹⁷Note, Figure 2 excludes Portugal. Portugal displays a very low inherited tax morale ($\delta_{PT} = 0.52$) and appears to be an outlier. Including Portugal we would observe a correlation coefficient of 0.52 (p-value= 0.006). We exclude Portugal from our main analysis, since this increases the strength of our first stage. However, we will show in Section 2.2.1 that including Portugal does not change our qualitative results in the second stage.

Table 4: 2SLS-estimations of the impact of tax morale on the UP in the year 2000^a

| | (2SLS-I) | | (2SLS-II) | | | |
|--------------------------------|--------------------------|---------------------------|----------------------------------|--------------------------|---------------------------|-----------------------------------|
| | First Stage ^g | Reduced Form ^h | Second Stage ^h | First Stage ^g | Reduced Form ^h | Second Stage ^h |
| $\hat{\zeta}_c^b$ | | | -56.799* (29.057) [-0.293] | | | -55.048** (24.781) [-0.284] |
| δ_c^c | 0.400*** (0.105) | -22.712** (10.241) | | 0.405*** (0.101) | -22.280** (9.942) | |
| Historical UP ^d | -0.001 (0.001) | 0.913*** (0.117) | 0.833*** (0.151) | -0.002* (0.001) | 0.834*** (0.125) | 0.708*** (0.159) |
| GDP p.c. ^e | | | | -0.001 (0.001) | -0.132 (0.087) | -0.215** (0.098) |
| OECD member-state ^f | 0.040 (0.026) | -8.662*** (2.532) | -6.412** (2.545) | 0.061** (0.028) | -6.728** (2.768) | -3.346 (2.646) |
| Constant | 0.493*** (0.106) | 32.395*** (10.344) | 60.405** (25.755) | 0.512*** (0.102) | 34.063*** (10.099) | 62.246*** (23.460) |
| Mean (s.d.) of dep. var. | 0.83 (0.05) | 21.30 (9.69) | 21.30 (9.69) | 0.83 (0.05) | 21.30 (9.69) | 21.30 (9.69) |
| R-squared | 0.52 | 0.88 | - | 0.58 | 0.90 | - |
| F-statistic on δ_c | 14.47 | - | - | 16.12 | - | - |

^aThe number of observations is in each specification equal to 25. Method of estimation is least squares. The table shows estimated coefficients with standard error in round parenthesis below. In squared brackets beta coefficients are reported. *, ** and *** indicate statistical significance at the 10-percent level, 5-percent level, and 1-percent level, respectively. ^bThis variable is a measure of tax morale in country c , where $c = \{\text{Austria, Belgium, Canada, China, Czech Republic, Denmark, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Japan, Mexico, Netherlands, Norway, Philippines, Poland, Romania, Russia, Spain, Sweden, Switzerland, United Kingdom}\}$; see equation (2) and Table 3. ^cThis variable is a measure of the inherited part of tax morale of American-born individuals from their ancestors country of origin c ; see equation (1) and Table 3. ^dThis variable is equal to the estimated size of the underground production (UP) as percentage of the official GDP of country c in the year 1989/90. Sources are given in footnote 9. ^eThis variable (derived from the *World Development Indicators* provided by the *World Bank*) is measured in 1,000 USD. ^fThis binary variable is equal to one if the country is a OECD member state, and zero otherwise. ^gThe dependent variable is equal to $\hat{\zeta}_c$. ^hThe dependent variable is equal to the estimated size of the (UP) as percentage of the official GDP of country c in the year 2000 from Schneider et al. (2010).

However, we can use estimates on the size of the UP for all countries from the late eighties, denoted by Historical UP_c . This gives a lag of more than a decade compared to our outcome variable from the year 2000.¹⁸ We are confident that after controlling for the historical level of UP in country c , the inherited part of tax morale affects the size of the current UP only through the channel of current tax morale.

Table 4 shows two different specifications with varying control variables. In each case the first stage shows a highly statistically significant positive effect of inherited tax morale and tax morale in the home country. The F-statistic on the excluded instrument is in each case well above ten – the critical value suggest by Stock et al. (2002). We conclude that our instrument is sufficiently strong.

The estimated coefficients on the second stage show the expected signs and are reasonable in size. Considering specification 2SLS-I, we see that the UP is estimated to be about 6.4 percentage points lower among OECD member states. The sample average is 21.30 percent of the GDP. The historical level of UP is a strong predictor for the size of the UP in the year 2000. Most importantly, we find that a higher tax morale decreases the size of the UP. The estimated coefficient – given by the ratios of the reduced form and the first stage effect of TM^{GSS} – is about minus 56.80 and statistically significant at a 5.1 percent level.¹⁹ The beta coefficient of minus 0.293 suggests that an increase in tax morale by one standard deviation (equal to 0.05 points) decreases the size of the UP by 0.293 standard deviations or 2.84 percentage points. This effect is somewhat lower compared to the fixed effects model. To get a better idea of the quantitative importance of tax morale we consider an increase of tax morale in Belgium. Belgium has the lowest tax morale in our sample ($\hat{\zeta}_{BE} = 0.69$) and an estimated UP of 22.2 percent of GDP. If tax morale would increase to the sample average of $\bar{\zeta}_c = 0.83$ the UP is estimated to decrease to 14.33 percent.

2.2.1 Sensitivity analysis

It is *a priori* not clear which covariates we should include in our analysis. More control variables are not necessary better. Control variables which are themselves outcome variables (i. e. factors that are determined by tax morale)

¹⁸The estimates are based on the sources cited in footnote 9. Best to our knowledge, no consistent estimates on the size of UP for a large sample of countries are available before 1989.

¹⁹We will provide evidence below that this effect is very robust, and its significance generally increases if more covariates are included.

should not be included in the 2SLS estimation (Angrist and Pischke, 2009, Chapter 3). Papers studying the determinants of the UP usually control for the official GDP (e.g. Johnson et al., 1997). In fact, GDP is a problematic covariate in our case, since it may be affected by tax morale. However, in any case, the specification 2SLS-II in Table 4 shows that including GDP has almost no impact on the estimated size of the UP.

The most important dimension determining the size of the UP put forward by the literature is governance.²⁰ Scholars emphasize the significance of low regulatory ‘burden’, less corruption, and a better rule of law. To check the robustness of our results, we control for the key dimensions of governance: (i) voice and accountability, (ii) political stability and absence of violence, (iii) government effectiveness, (iv) regulatory quality, (v) rule of law, and (vi) control of corruption. The data is from the the *Worldwide Governance Indicators Project* provided by the *World Bank*. These six measures are based on the views of thousands of stakeholders worldwide, including respondents to household and firm surveys, and experts from nongovernmental organizations, and public sector agencies.²¹ In each case a higher scores indicates a better environment. Since the correlation among the different measurements is very high, we include in Table 5 one variable in turn. All measurements of the quality of governance (except voice and accountability) are highly statistically significant and enter with a negative sign. Therefore, we can support the hypothesis that better governance is associated with a lower UP. Most importantly, we observe that the significant effect of tax morale persists. Compared to the baseline specifications in Table 4, the estimated effects even increased in size and in statistical significance. We also used different measurements of governance provided by the *The Heritage Foundation*. Based on the *Heritage Foundation’ Economic Freedom Index* (and all its subcomponent) we obtain equivalent results. Detailed estimation output is available in the Web Appendix.

In Table 6 we check the sensitivity of our result with respect to the sample chosen. As mentioned above, Portugal appears to be an outlier (with respect to the size of the inherited tax morale) that reduces the strength of our first stage. Specification R2-II shows that the F-statistic on the excluded instrument decreases to 10.18; compared to the baseline specification with an F-statistic of 14.47 (see specification R2-I). Nevertheless, including Portugal does not change the qualitative result in the second stage regression.

²⁰See, for instance, Johnson et al. (1997, 1998); Friedman et al. (2000); Dabla-Norris et al. (2008).

²¹Further details are provided by <http://www.govindicators.org>.

Table 5: Second stages of 2SLS-estimations controlling for governance^a

| | (R1-I) | (R1-II) | (R1-III) | (R1-IV) | (R1-V) | (R1-VI) |
|--|-----------------------------------|-----------------------------------|-----------------------------------|-----------------------------------|-----------------------------------|-----------------------------------|
| Tax morale ^b | -61.173** (30.612) [-0.316] | -51.786** (22.542) [-0.267] | -65.465** (26.043) [-0.338] | -54.309** (24.676) [-0.280] | -58.627** (25.234) [-0.302] | -56.755** (24.415) [-0.293] |
| Voice and accountability ^c | -2.888 (2.456) | | | | | |
| Political stability ^c | | -7.640*** (2.399) | | | | |
| Government effectiveness ^c | | | -5.502*** (1.801) | | | |
| Regulatory quality ^c | | | | -5.000*** (1.915) | | |
| Rule of law ^c | | | | | -4.980*** (1.824) | |
| Control of corruption ^c | | | | | | -3.850*** (1.322) |
| F-statistic on the excluded instrument | 16.27 | 18.20 | 14.99 | 16.28 | 16.37 | 16.32 |

^a The dependent variable is equal to the estimated size of the underground production (UP) as percentage of the official GDP of country c in the year 2000 from Schneider et al. (2010); for details see respective notes to Table 4. Each specifications also controls for 'Historical UP', 'OECD member state', and a constant (see respective notes to Table 4). The number of observations is in each specification equal to 25. Method of estimation is 2SLS. The table shows estimated coefficients with standard error in round parenthesis below. In squared brackets beta coefficients are reported. *, ** and *** indicate statistical significance at the 10-percent level, 5-percent level, and 1-percent level, respectively. ^b This variable is a measure of tax morale in country c in the year $t \in \{1996, 1999, 2000, 2001\}$, given by the estimated coefficients $\hat{\zeta}_c$ from (2); see Table 3. This variable is instrumented by a measure of the inherited part of tax morale of American-born individuals from their ancestors country of origin c , given by the estimated coefficients $\hat{\delta}_c$ from (1); see Table 3. ^c This variable is a measure of the quality of governance from the *Worldwide Governance Indicators Project* provided by the *World Bank*. A higher score indicates a better environment.

Table 6: Sensitivity analysis with respect to sample^a

| | (R2-I) | (R2-II) | (R2-III) | (R2-IV) |
|----------------------------------|----------------------------------|----------------------------------|------------------------------------|------------------------------------|
| | Full | Including | Excluding | Reduced |
| | sample | Portugal | Belgium | sample |
| Tax morale ^b | -56.799* (29.057) [-0.293] | -62.205* (33.207) [-0.327] | -110.540** (51.856) [-0.469] | -168.615* (100.159) [-0.637] |
| Historical UP ^c | 0.833*** (0.151) | 0.819*** (0.160) | 0.750*** (0.188) | 0.690*** (0.266) |
| OECD member state ^d | -6.412** (2.545) | -6.386** (2.563) | -5.256* (2.946) | -6.312 (4.391) |
| Constant | 60.405** (25.755) | 65.115** (29.390) | 105.769** (44.947) | 156.028* (86.734) |
| No. of observations | 25 | 26 | 24 | 21 |
| Mean of UP | 21.30 | 21.35 | 21.26 | 20.76 |
| Standard deviation of UP | 9.69 | 9.50 | 9.90 | 9.80 |
| Mean of TM^{WVS} | 0.83 | 0.82 | 0.83 | 0.83 |
| Standard deviation of TM^{WVS} | 0.05 | 0.05 | 0.04 | 0.04 |
| F-statistic on the excl. instr. | 14.47 | 10.18 | 6.93 | 3.22 |

^a The dependent variable is equal to the estimated size of the underground production (UP) as percentage of the official GDP of country c in the year 2000 from Schneider et al. (2010), where in the full sample $c = \{\text{Austria, Belgium, Canada, China, Czech Republic, Denmark, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Japan, Mexico, Netherlands, Norway, Philippines, Poland, Romania, Russia, Spain, Sweden, Switzerland, United Kingdom}\}$. The reduced sample excludes Belgium, Greece, Japan, and Romania. Method of estimation is 2SLS. The table shows estimated coefficients with standard error in round parenthesis below. In squared brackets beta coefficients are reported. *, ** and *** indicate statistical significance at the 10-percent level, 5-percent level, and 1-percent level, respectively. ^b This variable is a measure of tax morale in country c in the year $t \in \{1996, 1999, 2000, 2001\}$, given by the estimated coefficients $\hat{\zeta}_c$ from (2); see Table 3. This variable is instrumented by a measure of the inherited part of tax morale of American-born individuals from their ancestors country of origin c , given by the estimated coefficients $\hat{\delta}_c$ from (1); see Table 3. ^c This variable is equal to the estimated size of the UP as percentage of the official GDP of country c in the year 1989/90. Sources are given in footnote 9. ^d This binary variable is equal to one if the country is a OECD member state, and zero otherwise.

Table 7: Second stages of 2SLS-estimations controlling for governance – reduced sample^a

| | (R3-I) | (R3-II) | (R3-III) | (R3-IV) | (R3-V) | (R3-VI) |
|---------------------------------------|-----------------------------------|------------------------------------|------------------------------------|------------------------------------|------------------------------------|------------------------------------|
| Tax morale ^b | -178.989* (97.243) [-0.676] | -126.706** (61.061) [-0.479] | -138.040** (60.530) [-0.521] | -123.677** (59.649) [-0.467] | -150.287** (72.308) [-0.568] | -138.930** (65.621) [-0.525] |
| Voice and accountability ^c | -7.671 (5.049) | | | | | |
| Political stability ^c | | -9.255*** (3.322) | | | | |
| Government effectiveness ^c | | | -7.459*** (2.605) | | | |
| Regulatory quality ^c | | | | -6.689*** (2.564) | | |
| Rule of law ^c | | | | | -6.725** (2.786) | |
| Control of corruption ^c | | | | | | -5.002*** (1.922) |
| F-statistic on the excl. instr. | 3.69 | 4.38 | 4.89 | 4.75 | 4.07 | 4.37 |

^a The estimations summarized in this table are equivalent to those presented in Table 5, however, are based on a smaller sample (excluding Belgium, Greece, Japan, and Romania).

We observe an estimated beta coefficient of tax morale equal to minus 0.327, which is statistically significant at the 6.1 percent level. Similarly, specification R2-III shows that excluding Belgium (a further potential outlier, see Figure 2) increases the significance of the estimated impact of tax morale.

Finally, the low number of respondents from certain home countries in the GSS might be a concern. Therefore, we check the robustness of our results and exclude the five home countries (Belgium, Greece, Japan, Portugal and Romania) with the least number of observations, see specification R2-IV. This increases the average number of respondents per home country from 72 to 88. Notably, the lower panel in Table 6 shows that the variables of primary interest have almost identical means in the full and in this reduced sample. Based on the reduced sample we observe again a negative impact of tax morale. Both, the estimated (beta) coefficient and the standard errors increase in size, however, the effect is in this small sample still statistically significant at the 9.2 percent level. We also replicated the analysis controlling for the quality of governance for our reduced sample, see Table 7. Again, our results are robust to this sample modification and compared to the baseline specification in Table 6 we observe an increased statistical significance.

3 Conclusions

The causal link between tax morale and actual compliance behavior has not been established yet. To this extent, it was unclear why economic scholars should be interested in (the determinants of) tax morale? In this paper we provide first evidence that tax morale causally affects compliance behavior. Exploiting exogenous variation in tax morale – given by the inherited part of tax morale of American-born from their ancestors country of origin – our 2SLS estimation shows that a higher tax morale reduces the size of the underground production. This evidence increases particularly the significance of the strand of literature studying tax morale. It shows that tax morale can indeed help to explain the puzzle why people pay taxes, despite the existence of low audit probabilities and penalty rates. More generally, this result confirms the supposition that both economic incentives and social norms (or moral consideration) drive individual behavior.

An important caveat about this empirical evidence, however must be mentioned: given that our empirical analysis based on a quite small sample, the magnitudes of the estimated effects should be interpreted with some caution. In general, it would be desirable to verify the effect of tax morale on compliance behavior in better data if any opportunity arises.

Still, our result has also important implications for public policy. It shows that policy makers can in principle alter tax evasion by manipulating tax morale. Naturally, this conclusion raises the important question about good instruments for policy reforms. The literature discusses several institutional arrangements, such as direct democracy (Alm et al., 1999; Feld and Tyran, 2002; Torgler, 2005*a*), that are correlated with a high level of tax morale. Feld and Frey (2002) stress the importance of a respectful treatment of taxpayers by tax authorities. These are important findings, and scholars should pursue along these lines to uncover causal relationships. However, our 2SLS approach reveals that policy reforms intended to increase tax morale may not be very effective in the short-run. Tax morale is persistent, to a large degree it is inherited over generations, and it might take some time to change it.²² This notion is in line with the evidence from natural tax experiments presented in Blumenthal et al. (2001); Fellner et al. (forthcoming) who both find no effect of normative appeals or moral suasion on compliance behavior.

²²See Lindbeck and Nyberg (2006) for an elaborate discussion of parents' incentives for instilling norms in their children.

References

- Algan, Yann and Pierre Cahuc (2009), ‘Civic Virtue and Labor Market Institutions’, *American Economic Journal: Macroeconomics* **1**(1), 111–145.
- Algan, Yann and Pierre Cahuc (2010), ‘Inherited Trust and Growth’, *American Economic Review* **100**(5), 2060–2092.
- 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, Gary H. McClelland and William D. Schulze (1992), ‘Why Do People Pay Taxes?’, *Journal of Public Economics* **48**(1), 21–38.
- Alm, James, Gary H. McClelland and William D. Schulze (1999), ‘Changing the Social Norm of Tax Compliance by Voting’, *Kyklos* **52**(2), 141–71.
- 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.
- Angrist, Joshua D. and Jörn-Steffen Pischke (2009), *Mostly Harmless Econometrics: An Empiricist’s Companion*, Princeton University Press, Princeton, NJ.
- Bajada, Christopher and Friedrich G. Schneider (2005), ‘The Shadow Economies of the Asia-Pacific’, *Pacific Economic Review* **10**(3), 379–401.
- Barone, Guglielmo and Sauro Mocetti (2009), Tax morale and Public Spending Inefficiency, Economic Working Paper 732, Economic Research Department, Bank of Italy.
- Blumenthal, Marsha, Charles W. Christian and Joel Slemrod (2001), ‘Do Normative Appeals Affect Tax Compliance? Evidence from a Controlled Experiment in Minnesota’, *National Tax Journal* **54**(1), 125–138.
- 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.

- Dabla-Norris, Era, Mark Gradstein and Gabriela Inchauste (2008), ‘What Causes Firms to Hide Output? The Determinants of Informality’, *Journal of Development Economics* **85**(1-2), 1–27.
- Dell’Anno, Roberto and Friedrich G. Schneider (2006), Estimating the Underground Economy by Using MIMIC Models: A Response to T. Breusch’s Critique, Working Paper 0607, Johannes Kepler University of Linz.
- 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.
- Erard, Brian and Jonathan S. Feinstein (1994), ‘Honesty and Evasion in the Tax Compliance Game’, *RAND Journal of Economics* **25**(1), 1–19.
- Feld, Lars P. and Bruno S. Frey (2002), ‘Trust Breeds Trust: How Taxpayers are Treated’, *Economics of Governance* **3**(2), 87–99.
- Feld, Lars P. and Jean-Robert Tyran (2002), ‘Tax Evasion and Voting: An Experimental Analysis’, *Kyklos* **55**(2), 197–222.
- Fellner, Gerlinde, Rupert Sausgruber and Christian Traxler (forthcoming), ‘Testing Enforcement Strategies in the Field: Threat, Moral Appeal and Social Information’, *Journal of the European Economic Association* .
- Friedman, Eric, Simon Johnson, Daniel Kaufmann and Pablo Zoido-Lobaton (2000), ‘Dodging the Grabbing Hand: The Determinants of Unofficial Activity in 69 Countries’, *Journal of Public Economics* **76**(3), 459–493.
- Gordon, James P. (1989), ‘Individual Morality and Reputation Costs as Deterrents to Tax Evasion’, *European Economic Review* **33**(4), 797–805.
- Halla, Martin (2011), The Link between the Intrinsic Motivation to Comply and Compliance Behavior – A Critical Appraisal of Existing Evidence, in F. G.Schneider, ed., ‘Handbook on the Shadow Economy’, Edward Elgar, Cheltenham.
- Johnson, Simon, Daniel Kaufmann, Andrei Shleifer, Marshall I. Goldman and Martin L. Weitzman (1997), ‘The Unofficial Economy in Transition’, *Brookings Papers on Economic Activity*, **1997**(2), 159–239.
- Johnson, Simon, Daniel Kaufmann and Pablo Zoido-Lobaton (1998), ‘Regulatory Discretion and the Unofficial Economy’, *American Economic Review* **88**(2), 387–392.

- Kleven, Henrik, Martin Knudsen, Claus Kreiner, Soren Pedersen and Emmanuel Saez (2011), 'Unwilling or Unable to Cheat? Evidence from a Tax Audit Experiment in Denmark', *Econometrica* **79**(3), 651–692.
- Lindbeck, Assar and Sten Nyberg (2006), 'Raising Children to Work Hard: Altruism, Work Norms, and Social Insurance', *Quarterly Journal of Economics* **121**(4), 1473–1503.
- Loayza, Norman V. (1996), 'The Economics of the Informal Sector: A Simple Model and Some Empirical Evidence from Latin America', *Carnegie-Rochester Conference Series on Public Policy* pp. 129–162.
- OECD (2002), *Measuring the Non-Observed Economy – A Handbook*, Technical report, Organisation for Economic Co-operation and Development, Paris.
- Posner, Eric A. (2000), 'Law and Social Norms: The Case of Tax Compliance', *Virginia Law Review* **86**(8), 1781–1819.
- Rice, Tom W. and Jan L. Feldman (1997), 'Civic Culture and Democracy from Europe to America', *Journal of Politics* **59**(4), 1143–1172.
- 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.
- Schneider, Friedrich G. (2005), 'Shadow Economies Around the World: What Do We Really Know?', *European Journal of Political Economy* **21**(3), 598–642.
- Stock, James H., Jonathan H. Wright and Motohiro Yogo (2002), 'A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments', *Journal of Business & Economic Statistics* **20**(4), 518–529.
- 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, 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.
- Weck, Hannelore (1983), *Schattenwirtschaft: Eine Möglichkeit zur Einschränkung der öffentlichen Verwaltung? Eine ökonomische Analyse*, number 22 in 'Finanzwissenschaftliche Schriften', Lang, Bern.
- Wenzel, Michael (2005), 'Motivation or Rationalisation? Causal Relations between Ethics, Norms and Tax Compliance', *Journal of Economic Psychology* **26**(4), 491–508.