Opening the Black Box - Determinants of Tax Morale in Africa

Wiebke Weiger & Jasmin Cantzler

13 November 2015

Contents

1 Introduction

A country's ability to efficiently mobilize its national resources is a topic that has recently been gaining traction in the discussion surrounding development. Since the Monterrey Consensus was adopted in 2002, in which countries recognized the importance of raising domestic revenues, development aid has been increasingly shifting from direct financial assistance to improving tax system and national resource management. This shift in strategy was officially stated in 2008, in the Doha Declaration on Financing for Development, which pledged to enhance national tax revenues; one of the recognized strategies pledged to achieve this was combating tax evasion.

High levels of tax evasion lead to a misallocation of resources and hinder the ability of the government to invest in the provision of public goods. Therefore, understanding the rationale behind tax evaders becomes essential to national development and state building. Traditionally, the standard way to explain tax compliance has been the economics-of-crime approach, which assumes a rational taxpayer maximizing his or her utility by balancing the risk of detection and punishment with the benefit of tax evasion. Today, this approach is increasingly regarded as too narrow to fully explain tax compliance, and many argue for the need to include social factors, which are said to explain why people conform to paying taxes even in the absence of strong deterrence mechanisms. These social factors constitute and influence the individual's "intrinsic motivation to pay taxes", hereafter referred to as tax morale, by increasing moral costs of tax evasion and thus increasing tax compliance.

Against this background, the purpose of our research paper is to identify the determinants of tax morale in African countries, thereby opening the door for new policies fostering tax morale and thus tax compliance and increased public revenues in African countries. Using data from the Afrobarometer, we conduct a regression analysis using the commonly identified determinants of tax morale to test whether these also hold in developing countries; or more specifically, in African countries. Generally, the few studies that do look inside the black box of tax morale find that factors such as trust in government, the level of corruption, interactions with other taxpayers and tax administrators, tax burden, perception about benefits of public spending, social norms, detection, and punishment, gender all determine tax morale and hence influence tax compliance. Our conceptual framework is therefore based on the determinants of tax morale in developed countries, whereby:

$$\begin{aligned} TaxMorale_i = & \alpha_i + \beta_1 TrustinGovernment/PublicOfficials_i + \beta_2 LevelofCorruption_i \\ & + \beta_3 interactionswithothertaxpayers_i + \beta_4 taxburden_i \\ & + \beta_5 detectionandpunishment_i + \beta_5 SocialNorms_i + \epsilon_i \end{aligned} \tag{1}$$

This paper briefly outlines the design, operationalization and data selection, before conducting a first regression analysis and briefly describing its results. More extensive information will be available in the final report.

2 Method

2.1 Data Selection and Operationalization

As outlined in Assignment 1, we make use of the Afrobarometer Surveys to investigate the determinants of tax morale in African countries. We merged the survey results from Round 3, 4 and 5, resulting in a data set that spans the years 2005-2006, 2008-2010 and 2011-2013. For a step-by-step protocol please consult this R-file.

2.2 Descriptive Statistics of Our Data

Table 1 in the columns 2, 3 and 4 shows for each country the percentage of individuals saying that tax evasion is never justifiable. Columns 5, 6, and 7 give the mean value for all countries

based on a scale from 0 to 3, were 3 is the highest tax morale (value 0 integrates the values 4 to 10).

Country	meanage	meantaxmorale
Algeria	38.45598	2.351005
Benin	37.09086	2.351005
Botswana	38.86013	2.351005
Burkina Faso	37.16331	2.351005
Burundi	37.18000	2.351005
Cameroon	33.54746	2.351005
Cap Verde	38.08714	2.351005
Cote dIvoire	35.57383	2.351005
Egypt	39.05892	2.351005
Ghana	38.37247	2.351005
Guinea	39.57465	2.351005
Kenya	35.64767	2.351005
Lesotho	42.11250	2.351005
Liberia	36.10462	2.351005
Madagascar	38.76808	2.351005
Malawi	35.32700	2.351005
Mali	40.21890	2.351005
Mauritius	43.63917	2.351005
Morocco	36.11204	2.351005
Mozambique	33.10077	2.351005
Namibia	35.10292	2.351005
Niger	38.84448	2.351005
Nigeria	31.94566	2.351005
Senegal	38.39282	2.351005
Sierra Leone	36.15267	2.351005
South Africa	38.79433	2.351005
Sudan	33.93086	2.351005
Swaziland	38.39181	2.351005
Tanzania	37.84155	2.351005
Togo	34.70519	2.351005
Tunisia	42.29764	2.351005
Uganda	33.81181	2.351005
Zambia	34.37360	2.351005
Zimbabwe	37.67819	2.351005

The average values in Table 1 give a first overview about tax morale in African countries; in general .... seems to have higher tax morale than...

2.3 Inferential Statistics

A regression analysis is used to investigate the determinants of tax morale in African countries, using R Studio (R Core Team 2015; Hlavac 2015; Xie 2015) . Our dependent variable[1] asks the participants to agree or disagree with the statement "The tax department always has the right to make people pay taxes." The variable is measured from "Strongly Disagree", "Disagree", "Neither Agree or Disagree", "Agree", "Strongly Agree". Since we believe that the "distances" between these five points are not equal, an OLS regression in

[1]For Round 3 (Codebook 2005) it is variable Q52D, for Round 4 (Codebook 2008) it is Q44C, and for Round 5 (2015) it is Q48C.

this case is problematic because the assumptions of OLS are violated when used with a non-interval outcome variable. Instead, we opted for an ordinal logistic regression.

In order to use an ordinal logistic regression the following four assumptions need to hold:

1. The dependent variable is measured at the ordinal level.
2. One or more independent variables that are continuous, ordinal, or categorical
3. There is no multicollinearity
4. Ordinal odds

As indicated above, the first assumption holds true; our tax morale variable is measured at the ordinal level. The second assumption also holds true; all of our independent variables, are continuous, ordinal or categorical. To test for multicollinearity, we applied the variance inflation factor; the result are shown in the table below. Since all numbers exept year are very close to 1, we can assert that there is no multicollinearity.

Table 2: Variance Inflation Factor

Variables	VIF
Country	1.03
TrustPresident	1.01
CorruptionTax	1.01
SelfEmployedTax	1.02
Religion	1.01
Gender	1.01
EconomicSituation	1.01

We also created a simply correlation matrix to test for multicollinearity. As a rule of thumb, any correlation with a value of .5 and above will present multicollinearity, when 1 is perfect correlation. As seen in the correlation matrix, not a singly correlation yields a value above .5. The highest correlation, between country and paying taxes as a self-employed person is only -0.15, far from close to a high, multicollinear correlation.

Table 3: Correlation Matrix

	Country	TrustPresident	CorruptionTax	SelfEmployedTax	Religion	Gender	EconomicSituat
Country	1.00	0.03	0.01	-0.15	-0.07	0.00	0.00
TrustPresident	0.03	1.00	-0.07	-0.01	0.01	0.00	0.00
CorruptionTax	0.01	-0.07	1.00	0.03	0.00	-0.01	-0.05
SelfEmployedTax	-0.15	-0.01	0.03	1.00	-0.01	0.05	-0.01
Religion	-0.07	0.01	0.00	-0.01	1.00	0.04	-0.03
Gender	0.00	0.00	-0.01	0.05	0.04	1.00	-0.01
EconomicSituation	0.00	0.00	-0.05	-0.01	-0.03	-0.01	1.00

We had some difficulties to test for ordinal odds and as a consequence we have not yet decided which test or approach to chose, but will resolve this for the final project. Essentially the assumption of ordinal logistic regression is that the "relationship between each pair of outcome groups is the same" (UCLA –). There are some statistical test available but their weakness is said to lay in the tendency to reject the null hypothesis, "and hence, indicate that there the parallel slopes assumption does not hold, in cases where the assumption does hold" (UCLA –).

Nevertheless, we went ahead and created two models using ordinal logistic regression, whose output is shown in the table below.

Table 4: Ordinal Logistic Regression Results of Tax Morale

	Dependent variable:	
	TaxMorale	
	(1)	(2)
CorruptionPresidentMost of them	-0.14^{***}	-0.14^{***}
	(0.03)	(0.01)
CorruptionPresidentNone	-0.11^{***}	-0.12^{***}
	(0.03)	(0.01)
CorruptionPresidentSome of them	-0.20^{***}	-0.20^{***}
	(0.03)	(0.01)
CorruptionTaxMost of them	-0.29^{***}	-0.29^{***}
	(0.02)	(0.01)
CorruptionTaxNone	-0.16^{***}	-0.16^{***}
	(0.03)	(0.01)
CorruptionTaxSome of them	-0.33^{***}	-0.33^{***}
	(0.02)	(0.01)
EconomicSituationFairly good	0.07^{***}	0.07^{***}
	(0.02)	(0.02)
EconomicSituationNeither good nor bad	0.03	0.03
	(0.02)	(0.02)
EconomicSituationVery bad	0.30^{***}	0.30^{***}
	(0.02)	(0.02)
EconomicSituationVery good	0.36^{***}	0.36^{***}
	(0.03)	(0.01)
Year		0.01^{***}
		(0.0000)
Gendermale		0.07^{***}
		(0.01)
Age		0.001
		(0.001)
Observations	73,878	73,381
Note:	*p<0.1; **p<0.05; ***p<0.01	

Since β's in logistic regression are a bit weird, we need to transform these results into predicted probabilities in your final project.

3 Discussion & Open Questions

Wiebke - please have a look - only collecting my doubts here so I don't forget. 1. Since many of our dependent variables are also nominal we are still investigating whether there are more suitable regression models applicable than ordinal logistic regression. After all, at this point we are treating the dependent variables as numerical. Academic opinions seem strongly divided whether it is okay or highly problematic - what would be your take on this? 2. Maybe there is a way to group our corruption variables together per respondent? Or not advisable

References

Hlavac, Marek. 2015. "Stargazer: Well-Formatted Regression and Summary Statistics Tables."

R Core Team. 2015. "A Language and Environment for Statistical Computing. R Foundation for Statistical Computing." Vienna, Austria.

UCLA. –. "R Data Analysis Examples: Ordinal Logistic Regression."

Xie, Yihui. 2015. *Knitr: A General-Purpose Package for Dynamic Report Generation in R.R Package Version 1.11.*