WIDER Working Paper 2021/170

Tax effort revisited: new estimates from the Government Revenue Dataset

Kyle McNabb,[1] Michael Danquah,[2] and Abrams M.E. Tagem[2]

November 2021

Abstract: Attention on domestic resource mobilization—particularly in developing countries—has increased significantly in recent years. This stems from, among other things, recognition in the Sustainable Development Goals that further domestic funding is required for development needs, and the Addis Tax Initiative, which aims to foster fairer and more effective domestic resource mobilization. And whilst there is a recognition that many low- and middle-income countries could be collecting more in tax revenues, the answer to the question of just how much more is unclear. So-called tax effort studies have attempted to shed light on this question for many years and recent estimates suggest that many developing countries are not performing anywhere near their potential (i.e. exerting enough effort). This study makes two significant contributions to the tax effort literature. First, we find that the stochastic frontier approaches used in many recent studies are sensitive to empirical specification and the resulting tax effort scores are strongly influenced by outlying input observations. We employ the True Random Effects approach and find that tax effort scores are, on average, higher and more tightly distributed than previous studies suggest. Second, we enjoy improved data coverage due to using the most recent version of the UNU-WIDER Government Revenue Dataset, alongside governance indicators from the V-Dem dataset. This allows us to present tax effort estimates for 50 per cent more observations than the next most complete study.

Key words: tax effort, domestic revenue mobilization, tax, development

JEL classification: H21, H87, O11, O23

Acknowledgements: The authors are grateful to William Greene for offering advice and extra help in the coding of the TRE models.

Note: On 10 October 2022, the tax effort scores were added to the working paper's webpage, as a downloadable file (https://www.wider.unu.edu/publication/tax-effort-revisited-new-estimates-government-revenue-dataset).

[1] Overseas Development Institute, London, UK, corresponding author: k.mcnabb@odi.org.uk; [2] UNU-WIDER, Helsinki, Finland

This study has been prepared within the UNU-WIDER project GRD – Government Revenue Dataset. It is part of UNU-WIDER's Domestic Revenue Mobilization (DRM) programme, which is financed through specific contributions by the Norwegian Agency for Development Cooperation (Norad).

Information and requests: publications@wider.unu.edu

ISSN 1798-7237 ISBN 978-92-9267-110-5

https://doi.org/10.35188/UNU-WIDER/2021/110-5

Typescript prepared by Joseph Laredo.

United Nations University World Institute for Development Economics Research provides economic analysis and policy advice with the aim of promoting sustainable and equitable development. The Institute began operations in 1985 in Helsinki, Finland, as the first research and training centre of the United Nations University. Today it is a unique blend of think tank, research institute, and UN agency—providing a range of services from policy advice to governments as well as freely available original research.

The Institute is funded through income from an endowment fund with additional contributions to its work programme from Finland, Sweden, and the United Kingdom as well as earmarked contributions for specific projects from a variety of donors.

Katajanokanlaituri 6 B, 00160 Helsinki, Finland

1 Introduction

It is well known that many—especially low- and middle-income—countries require significant progress on domestic resource mobilization (DRM) if they wish to, for example, meet the Sustainable Development Goals (SDGs) or support their economies through the economic fallout from the COVID-19 pandemic. The issue has received renewed attention in the past decade, underlined by commitments such as the Addis Tax Initiative, and is enshrined in the SDGs themselves (SDG 17.1 calls for countries to 'strengthen domestic resource mobilisation, including through international support to developing countries, to improve domestic capacity for tax and other revenue collection'). Yet the average tax-to-GDP ratio in low-income countries (LICs) stands at just 13.1 per cent, compared with 23.3 per cent in high income countries (HICs), and the rate of increase is slow in many jurisdictions (UNU-WIDER 2021). How much domestic revenue, then, can poorer countries realistically be expected to collect? The study of *tax effort*, defined as the ratio of actual tax collection to tax *potential*, can provide some clues to the answer. The concept is not new, but it is now more important than ever that policy advice to countries seeking to increase their tax effort is grounded in some notions not only of *where* they can hope to get, but also of *how* they might get there. Whilst this study focuses primarily on the former question, the estimations presented herein provide some useful clues as to which countries are performing 'better' or 'worse'.

We utilize the enhanced coverage of the UNU-WIDER Government Revenue Dataset (GRD; UNU-WIDER 2021) to revisit the question of tax effort, and make a number of methodological innovations, focusing on the variables included, estimation methods, and specification choices. Related literature has recently focused on estimating tax effort via a stochastic frontier approach (SFA) (e.g. Fenochietto et al. 2013; Langford and Ohlenburg 2016; Mawejje and Sebudde 2019), yet the modelling choices have not come under appropriate scrutiny, particularly with regard to the estimation of country-specific inefficiency in tax collection. The tax effort indices resulting from this study will be updated annually and published online alongside the GRD.

Our results show that previous estimates of tax effort have been severely biased by outlying observations amongst the input variables, which can often lead to wildly inaccurate estimates of tax potential. Our diagnostic work shows that tax effort estimates produced by the True Random Effects (TRE) model (Greene 2005) are not influenced by such observations to the same extent. The estimates of tax potential presented herein, therefore, are somewhat more conservative than many of those presented in the literature previously but, we feel, more realistic. Our sample of 161 countries—3,901 observations in total—is by far the largest sample to date for which tax effort has been estimated.

This paper is organized as follows. Section 2 contains an extensive literature review of tax effort studies to date. Section 3 discusses variables, data, and methodological issues before our main analysis is presented in Section 4. In Section 5, we compare our findings with those of previous studies. Section 6 concludes and discusses key political economy and policy takeaways.

2 Literature review

This section provides a literature review of the study and estimation of tax effort, beginning with Oshima in 1957 up to the most recent—to our knowledge—study, Mawejje and Sebudde (2019). We constrain the focus here to include only those studies that examined tax effort in a cross-country setting, and thus exclude a number of studies that focused on a single country. A summary of key findings with regard to explanatory variables in the tax effort regressions reviewed is provided in Table 1.

The genesis of the literature exploring the determinants of tax effort is likely Oshima (1957), whose comparison of a sample of 32 countries between 1948 and 1954 highlighted a number of factors that might account for the relatively poor performance of what he termed 'less developed countries'.[1] On the whole, he suggested that the underlying mechanism at play was simply the level of development of the economy, which is correlated with higher productivity. The study does, however, hint that a few more specific factors might be correlated with level of tax revenue performance, such as the degree of authoritarianism of the government and the importance of subsistence agriculture (with particular reference to the low observed figure for India). Yet no attempt at statistical analysis—beyond simple two-way correlations—was undertaken to prove these theories.

A first regression approach to the question of whether the level of development (as proxied by GNP) might drive government revenue shares comes from Williamson (1961), whose bivariate tests on a sample of 33 developed and developing countries showed a positive relationship between the two variables. Hinrichs (1965) then posited that, broadly, two sets of factors drove the tax level, namely *structural* and *ideological* changes in the economy. The former refers to the fact that processes such as industrialization, urbanization, and specialization lead to higher shares of government revenue in GNP, whilst the latter refers to the process of social mobilization, whereby ideological changes might, for example, lead to the formation of a welfare state whose higher spending needs dictate that the tax level also increases. Hinrichs's study examines 60 countries between 1957 and 1960, finding that GNP per capita is significantly correlated to the revenue ratio in all these countries, but the same relationship did not hold when the sample was divided into developed and less developed sub-samples. Hinrichs (1965) subsequently found that openness—defined as the level of imports as a share of GNP—proved to be a much 'better' index to explain less-developed countries' revenue share; broadly, the estimates showed that less developed countries collected 5 per cent of GNP plus half of the openness-to-GNP ratio, in revenue.[2] The major takeaway from the study, however, is Hinrichs's suggestion that determinants of revenue shares might differ according to level of development, with *openness* a better explanatory variable for less developed countries and *GNP per capita* the better explanator for developed countries.

[1] This is true to the best of our knowledge, although it is, of course, entirely possible that earlier literature may have made similar observations.

[2] Thorn (1967), however, found openness to be an insignificant explanatory variable, but did provide evidence—by including a dummy variable—that former British dependencies had higher revenue shares, suggesting that preferences instilled in such countries for certain levels of public expenditure were also driving revenue efforts.

Table 1: Summary of key findings of recent tax effort studies

Study	Mawejje & Sebudde (2019: T2 Col.1)	Langford & Ohlenburg (2016: T2 Col.1[$])	Cyan et al. (2013: T1 Col. 1)	Cyan et al. (2013: T2 Col. 2)	Fenochietto et al. (2013: T2 Col. 1)	Sen Gupta (2007: T7 Col. V)	Bird et al. (2004: T2 Col. 1)	Ghura (1998: T2 Col. 4)	Stotsky & WoldeMariam (1997: T 6 Col. 2)	Tanzi (1992: T12.5 '1998')
Dependent variable	Non-resource tax excluding social security contributions	Non-resource tax excluding social security contributions	Tax + non-tax revenue	Tax + non-tax revenue	Tax + social security contributions	Central government revenue excluding grants	Tax revenue	Tax revenue	Tax revenue	Tax revenue
Economic variables										
GDP per capita	+ ***		+ ***	+ ***	+ ***	+ ***	- ***	+ ***	+	-
Openness[#]	+ ***	+ **			+ ***	+ **	-	+ ***	+ ***	+ ***
Agriculture share GDP	- ***		- ***	- **	- ***		- ***	- ***	- ***	- **
Manufacturing share GDP		+ *								
Services share GDP			- ***	-						
Construction share GDP			- ***	- ***						
Inflation	- ***	- ***	- ***	- ***				- ***		
Resource revenue	-									
Crude petroleum production			- ***	- **						
Oil-producer dummy								+ ***		
Mining dummy								+ ***		
Grants	+ **		- ***	- ***				- **		
Income inequality	-		- ***	- ***	- ***					
Government debt			+ ***			- **				+ *
Change in government debt								+		
Globalization			+ ***	+ ***						
Gross fixed capital formation			+ ***	+ **						
Aid / GDP						+ *				
Structural reforms								+ ***		
Change in REER								+		
Change in terms of trade								+		

Study	Mawejje & Sebudde (2019: T2 Col.1)	Langford & Ohlenburg (2016: T2 Col.1[$])	Cyan et al. (2013: T1 Col. 1)	Cyan et al. (2013: T2 Col. 2)	Fenochietto et al. (2013: T2 Col. 1)	Sen Gupta (2007: T7 Col. V)	Bird et al. (2004: T2 Col. 1)	Ghura (1998: T2 Col. 4)	Stotsky & WoldeMariam (1997: T 6 Col. 2)	Tanzi (1992: T12.5 '1998')
Dependent variable	Non-resource tax excluding social security contributions	Non-resource tax excluding social security contributions	Tax + non-tax revenue	Tax + non-tax revenue	Tax + social security contributions	Central government revenue excluding grants	Tax revenue	Tax revenue	Tax revenue	Tax revenue
Demographic variables										
Population density			-	- ***						
Share rural population	-									
Health expenditure	+ ***				+ ***					
UN education index		+ ***	+ ***	+ ***						
Age dependency ratio		+ **	+	- *						
Population growth							- ***			
Human capital index								+ **		
Governance / Institutional variables										
Corruption	- ***	- ***	- ***	- ***		+		- ***		
Law and order		- **				-				
Democratic accountability		- *								
Government stability						+				
Governance index							+ ***			
Observations	*2,422*	*1,664*	*1,079*	*1,094*	*?*	*376*	*104*	*415*	*249*	*66*
Estimation method	SFA	SFA	FE	SFA	SFA	SYS-GMM	OLS	IV-GLS	RE	OLS

Note: SFA = Stochastic Frontier Analysis; FE = Fixed Effects; SYS-GMM = System Generalised Method of Moments; IV-GLS = Instrumental Variable Generalised Least Squares; RE = Random Effects; OLS = Ordinary Least Squares. $ The governance variables in this study enter into the inefficiency equation. # Openness is defined as one of (i) imports, (ii) exports, or (iii) imports + exports as a share of GDP.

We show results from only one estimation per study (aside from Cyan et al. 2013). Normally, this is the baseline specification. Other specifications in some studies included further variables but in the interests of space these are not shown. Interested readers should consult the relevant studies for further details. Signs have been changed for a number of variables that were specified as an inverse, or for example a commonly used corruption variable is coded 1–6, with 6 being the highest level of corruption. The interpretation here is, then, that a positive coefficient on corruption would imply a positive relationship with tax collection.

Source: authors' construction.

The first study that might be considered to be in the spirit of the present work—i.e. that attempts to relate a number of factors to tax ratios before estimating revenue effort as compared with actual collection—is that of Lotz and Morss (1970). Considering the role of foreign trade—or *openness*—they find that exports' share of GNP is the best explanatory variable out of three, namely the aforementioned, imports' share of GNP, and the sum of the two. Further, they find that the trade balance (X-M)/GNP is a similarly good predictor of the tax level.[3] Turning to the role of natural resources in an economy's structure, the authors find no significant relationship between the share of mineral and petroleum exports in total exports and the tax ratio.[4] The authors also find a positive relationship between the degree of fiscal decentralization (as measured by the share of tax revenues collected by non-central governments) and tax revenue effort, suggesting that where tax revenue collection is closer to the population, there is higher trust in the collecting authority, leading to a greater willingness to contribute. Having attempted to explain differences in tax effort across countries, the authors make a first attempt at examining the 'unexplained differences', namely the difference between actual tax collected and that predicted by their regression model. They find broad variation, with around half the sample of countries collecting more than predicted by the model, and the other half collecting less. In a similar vein to Hinrichs (1965), however, the authors conclude by suggesting that a different set of variables matter for explaining tax ratios in *developing*, as compared with *developed*, countries, factors that determine the size of the tax base being important in the former, whilst factors influencing the demand for government services are more important in the latter.

Closely following the study of Lotz and Morss (1970) was that of Bahl (1971), who makes a number of important theoretical contributions to the tax effort literature; namely, he considers that the modelled variance in tax ratios can be decomposed into factors related to (i) tax capacity and (ii) tax effort. Using a sample of 49 developing countries with data between 1966 and 1968, Bahl (1971) estimates a regression model including the share of agriculture in GDP and the share of mining exports in total exports, finding significant negative and positive relationships, respectively. In a similar vein to Lotz and Morss, he ranks countries according to taxable capacity (i.e. predicted tax ratio given the factors present in the regression model) and tax effort (i.e. the ratio of actual tax collected to estimated tax capacity). Bahl also includes a set of regional dummy variables, finding that the estimated tax ratios in the 'Tropical Africa' region were higher than those in all others.[5]

The notion of estimating tax effort—or taxable capacity—gained a lot of momentum within the International Monetary Fund (IMF) during the 1970s, leading to a series of influential studies on the topic (IMF 1971, 1975, 1979), which were published as part of the Staff Papers series.[6] In each of the three studies, again, openness to trade (export ratio), economic structure (as proxied by the share of mineral exports), and per capita (non-export) income were found to be significant determinants of tax effort. Interestingly, IMF (1979) found that the estimated tax effort indices across all three studies (which took data from, respectively, 1966–68, 1969–71, and 1972–76) showed that many countries saw significant changes in their scores depending on the time period

[3] The judgements over what constitutes a 'good' explanatory variable rely, in this paper, on the size of the estimated R^2.

[4] This likely provides early evidence of the challenges facing many resource-rich countries in mobilizing tax revenue collection—see e.g. Bornhorst et al. (2009) for a recent study on this issue.

[5] 'Tropical Africa' pertains to all African countries in the sample not considered 'Middle East and North Africa'.

[6] By 1979, however, the IMF had adopted the term 'international tax comparisons' in preference to tax effort.

in question, suggesting that the results presented displayed a degree of vulnerability to the sample examined.

Leuthold (1991) was the first to go beyond OLS regression (which all the aforementioned studies had employed), estimating an AR(1) time series model on a panel of eight African countries and finding that it performs much better. The same set of covariates as in previous studies is tested and it is supplemented by the inclusion of a measure of foreign gifts and grants. Tanzi (1992), employing simple bivariate OLS models on a panel of 83 developing countries, regressed the tax ratio on GDP per capita and found a declining importance of the role of GDP per capita between 1978 and 1988, suggesting that following the debt crisis of the 1980s, other determinants such as the structure of the economy became more important. Like previous studies, Tanzi (1992) finds that the share of agriculture in GDP is a superior explanatory variable to GDP per capita (again through bivariate OLS estimations). He also finds, in a multivariate setting, that the debt-to-GDP ratio is a significant explanatory variable.

Stotsky and WoldeMariam (1997) use fixed effects panel estimation to examine tax effort across a panel of 43 sub-Saharan African countries between 1990 and 1995. Again, results are in line with previous studies: that agriculture's share in GDP and openness to trade are positively related to tax effort. In contradiction to much of the previous literature, however, these authors find a negative relationship between the mining share and tax effort.

Ghura (1998) provides, to the best of our knowledge, the first examination of the role of corruption on tax revenue mobilization, marking a shift in the literature, which had previously focused almost exclusively on economic and structural factors. Ghura (1998) first augments the—by this point—'traditional' set of independent variables (namely income, openness, mineral exports, and agricultural value added) with a number designed to capture 'economic policies', namely *inflation*, *change in real effective exchange rate*, *structural reform*, and a measure of *human capital*. Of these, results from an IV-GLS regression on a panel of 39 sub-Saharan African countries between 1985 and 1996 show that inflation has a negative and significant relationship with the tax ratio, whilst both structural reforms and the index of human capital are significantly positively related. The key variable of interest in the paper—namely the level of corruption, as measured by the ICRG—is negatively related to the tax ratio. Finally, a dummy variable to capture members of the CFA Franc is negatively related to the tax ratio, again providing evidence that unobserved regional factors might play a part in the story.

Bird et al. (2004) examine in detail the role of societal institutions in determining tax effort. They consider, specifically, that the 'traditional' set of variables tested capture the 'supply side' factors, or the size/shape of a country's key tax handles, but that institutional variables measuring, for example, control of corruption, rule of law, or bureaucratic quality represent a key set of 'demand-side' variables, which can go some way to explaining revenue performance. The authors find that indicators of institutional quality from both the International Country Risk Guide (ICRG) and the Quality of Governance Institute (QoG) strongly positively determine tax revenue effort. They also find some evidence that a larger shadow economy is associated with a lower tax effort.

An update to earlier studies is provided by Sen Gupta (2007), who both extends the sample under consideration and accounts for a number of outstanding econometric issues. In a fixed and random effects (FE and RE, respectively) panel estimation on 105 developing countries over a 25-year timespan, the results on the 'traditional' economic variables largely confirm those of previous studies; and in some specifications the amount of foreign aid received is significantly positively related to revenue performance. Turning to 'institutional' variables, the results suggest no consistently significant effects of the level of corruption, law and order, political stability, or economic stability. These FE and RE results are broadly confirmed first in a Panel Corrected

Standard Errors (PCSE) model and then in a GMM setup. Evidence is also presented that suggests that countries which rely more on income taxes have, on average, a better revenue performance than those which rely on indirect taxation (both taxes on goods and services and trade taxes). Interestingly, some of these results are found to differ when countries are grouped according to income level. It emerges that the share of agriculture in GDP is significant and negatively related to the tax ratio across all income levels; openness to trade is only robustly positive and significant in LICs and MICs; aid has a statistically significant positive effect on revenue mobilization only in LICs; and measures of corruption and political stability are significant only for LICs and MICs.

Until 2010, all cross-country studies examining tax effort had relied on what Cyan et al. (2013) would later term the 'traditional' method, namely a regression approach. Fenochietto et al. (2013) and Pessino and Fenochietto (2010), however, employed a stochastic frontier model and modelled tax effort with a 'production function' approach. As with the regression models, the estimated 'tax frontier' represents the theoretical maximum revenue a country could collect, given the inputs in the model. However, the tax gap (difference between actual revenue and the frontier) in a stochastic frontier model is decomposed into (i) a level of inefficiency and (ii) a random error term. Factors such as administrative or policy / legislative weaknesses influence the level of (in)efficiency in tax collection. Pessino and Fenochietto (2010) estimate a stochastic frontier model on 96 countries covering 16 years (1991–2006). The model includes a fairly 'standard' set of explanatory variables (GDP per capita, openness, agriculture value added, public education expenditure, income distribution, inflation, corruption), each of which has the expected sign (as previously in the literature) and is statistically significant. The stochastic frontier approach also allows the researcher to include determinants of (i) above, namely technical inefficiency. Pessino and Fenochietto (2010) model inefficiency as a function of the level of corruption, finding that lower levels of corruption predict lower levels of inefficiency. The same authors provide an update to the study in 2013 (Fenochietto et al. 2013) to include an additional 17 resource-rich countries (defined as those where revenue from natural resources was greater than 25 per cent of total tax revenues) over the period 1991–2012, and the results are broadly similar. However, the inclusion of the resource-rich countries leads to a greater estimated inefficiency parameter. Whilst these two studies undoubtedly make methodological advances in the study of tax effort, little attention is given to the role of institutions or governance (with only the corruption variable used in this connection), both studies preferring to focus on the 'economic' determinants of tax effort.

Cyan et al. (2013) provide an in-depth examination of prior tax effort studies, focusing specifically on the estimation methods employed. They consider three approaches to estimating tax effort, namely (i) the 'traditional' method, where tax revenue potential is calculated on the basis of the predicted values of a regression model (as used in all the studies reviewed to this point in the present work); (ii) the stochastic frontier approach; and (iii) a third approach where tax effort is estimated by comparing a country's actual collections to its expenditure level (or budget balance), which, the authors argue, is a good measure of what a country might desire to spend on public goods. Cyan et al. (2013) consider a panel of 94 countries over the period 1970–2009 in order to compare the performance of each of the three approaches to estimating tax effort. When attempting to explain the estimated inefficiency resulting from the stochastic frontier approach, the authors include a range of variables such as complexity of the tax system (an index constructed using the relative shares of different tax handles), a measure of tax morale (from the World Values Survey), political fractionalization (Beck et al. 2001; The Database of Political Institutions), and the level of government debt. The comparison of traditional (fixed effects) and stochastic frontier estimates shows a high correlation, whilst both methods show a lower correlation with the measure of budget balance.

A more recent study to employ stochastic frontier analysis (SFA) is that of Langford and Ohlenburg (2016). These authors were the first to employ the UNU-WIDER GRD as the source of tax revenue data (previous studies had compiled tax data from sources such as the IMF's Government Finance Statistics, OECD's Revenue Statistics, CEPALSTAT, or internal IMF data). In addition to the 'standard' set of economic variables, they include a large number of demographic and institutional factors. Findings—from SFA on a panel of 85 countries, covering the period 1984–2010—are broadly in line with the existing literature, whilst the authors also find significant effects of measures of law and order, corruption, and democratic accountability on the tax ratio.

Finally, the most recent study of note to estimate tax effort is that of Mawejje and Sebudde (2019). The authors restrict the time period to 1996–2015 in order to gain maximum cross-country coverage, with the result that tax effort estimates are provided for 150 countries. The findings are, broadly, in line with other studies.

From the literature reviewed we can draw a number of broad conclusions that will inform the strategy followed herein:

i. A 'core' set of 'economic' variables that explains tax effort emerged fairly early in the literature and have remained important explanators. These include GDP per capita, the degree of openness to trade, the size of the agricultural sector, the rate of inflation, the reliance on grants or aid, and the importance of the natural resources sector.
ii. Authors have increasingly found that both demographic and institutional variables are significant explanatory factors. Whilst there is considerably less consensus on the exact variable to include in the estimation, measures of corruption, population structure (either age structure or urbanization), and level of education have generally been found to be important explanatory variables.
iii. Estimation strategies have evolved over the years, with the consensus seemingly now that SFA is the most useful or appropriate technique for estimating tax effort. Whilst it certainly has advantages, Cyan et al. (2013) showed that there was little difference in the performance of the models (between SFA and a more traditional regression approach). Crucially, however, there was a difference in the eventual tax effort indices computed. Thus, the question of estimation method remains crucial and will be paid due attention in the present work.

Whilst some studies in the late 20th century began to find that results were differing according to region or income group, more recent studies have neglected to continue this line of investigation. It is very likely that what matters for explaining tax effort will differ at different levels of development.

3 Estimation strategy and data

3.1 Estimation strategy

We follow the recent literature and estimate tax effort according to the production function approach. Specifically, this is modelled as:

$$T_{it} = f(X_{it;}\beta) \tag{1}$$

T_{it} is the (observed) tax-to-GDP ratio for country *i* at time *t*. $f(X_{it;}\beta)$ represents the production function, where a vector of inputs, X, is used to generate tax revenue, whilst β represents the vector of parameters to be estimated. This equation describes the scenario where tax policy is perfectly (efficiently) applied in order to maximize potential revenues, *T*, and there are no random shocks to collection. However, SFA models allow us to incorporate the effects of inefficiency (ξ_{it}) in tax collection, as follows:

$$T_{it} = f(X_{it;}\beta) \cdot \xi_{it} \tag{2}$$

where 0 < ξ_{it} < 1.

$\xi_{it} = 1$ represents a scenario where the tax authorities are collecting the maximum potential tax revenue, given the underlying economic, demographic, and institutional factors captured in X, ξ_{it} < 1 describes a situation where there is inefficiency in the process of tax collection, and T is less than potential. Finally, tax collection is also subject to a series of random shocks, $e^{v_{it}}$

$$T_{it} = f(X_{it;}\beta) \cdot \xi_{it} \cdot e^{v_{it}} \tag{3}$$

Equation (3) then illustrates that any deviations from revenue potential are due to (i) technical inefficiencies, ξ_{it}, and (ii) stochastic shocks, $e^{v_{it}}$.

A standard econometric representation of this equation is obtained by following the approach in Aigner et al. (1977). Given that ξ_{it} = $e^{-u_{it}}$, we take natural logarithms of the model in (3) and obtain the following 'base' equation[7]:

$$q_{it} = \alpha + \beta' x_{it} + v_{it} - u_{it} \tag{4}$$

where q_{it} = ln $(\frac{T_{it}}{Y_{it}})$ is the natural logarithm of the tax-to-GDP ratio, $'x_{it} = \ln(X_{it})$ is the vector of independent variables affecting tax revenue performance, and $v_{it} - u_{it}$ is a composite error term including both the random (stochastic) error, v_{it} ~ $N[0,\sigma_v^2]$, and the inefficiency term, $u_{it} = -\ln(\xi_{it})$; u_{it} ~ $N^+[0, \sigma_u^2]$.

However, we also wish to account for the heterogeneity in tax collection via a set of observed factors (z) that affect the inefficiency in tax collection in any country. These factors might influence q_{it} directly and thus enter the model as a direct input of tax capacity (z_c), or as a driver of the inefficiency term (z_e). In order to account for both cases, we augment (4) such that

$$q_{it} = \alpha + \beta' x_{it} + \vartheta_c z_{it,c} + v_{it} - u_{it} \tag{5}$$

where $u_{it} = \vartheta_e z_{it,e}$.

There is no consensus on whether the set of variables contained in z should enter as a component of x, z_c, or z_e. Langford and Ohlenburg (2016) consider the case of a variable which proxies the level of corruption in a country. Corruption may well affect the tax capacity (and thus the placement of the frontier itself) or it might affect the level of tax effort (how close that country is to the tax frontier). Consider also the case of a variable proxying institutional quality. The lack of reliable, accessible institutions may hamper the level of economic activity in any country (and thus the levels of variables contained in x), in which case the proxy might best enter as part of z_c. However, at the same time, a variable capturing a dysfunctional revenue authority in a country with otherwise sound institutions would best be included as part of z_e. Langford and Ohlenburg

[7] The production function is assumed to be Cobb-Douglas in nature and thus, in natural logarithm form, can be estimated via the linear equation shown in (4).

(2016) again highlight that this specification choice has important implications for the interpretation of our results. Specifically, if we include this proxy in z_c, then a country with a low institutional quality score will have a lower estimate of tax capacity (*q*) but a higher tax effort score (as the denominator of $\frac{T_{it}}{q_{it}}$ will, *ceteris paribus*, be smaller).

In all, we estimate the base case pooled model in addition to three other models—RE (Pitt and Lee 1981), Battese and Coelli (BC) (Battese and Coelli 1995), and TRE (Greene 2005)—that, together, capture the panel structure of our dataset. The baseline specifications (without introduction of z) for the RE model are generally specified as

$$q_{it} = \alpha + \beta' x_{it} + v_{it} - u_i,\ u_{it}=|U_i|$$

$$v_{it} \sim N[0,\sigma_v^2]$$

$$U_i \sim N^+[0, \sigma_u^2]$$

We note that $u_{it} = |Ui|$; therefore, no time-invariant element is modelled here. The BC model can also be specified as

$$q_{it} = \alpha + \beta' x_{it} + v_{it} - u_{it}$$

$u_{it} = exp[-\eta(t-T)]|U_i|$, or $u_{it} = exp[g(t,T,z_{it})]|U_i|$ if we introduce z variables

$$vit \sim N[0,\sigma v2]$$

$$Ui \sim N + [0, \sigma u2]$$

With the BC model, the time-invariant random component is still a major influence on the model and, therefore, the random parts of the BC model do not vary with time. The time-invariant element in the model, u_{it}, is intended to capture all (and only) the country-specific inefficiency and does not treat unobserved time-invariant effects such as heterogeneity in the data. There is always the likelihood that Ui absorbs large amounts of cross-country heterogeneity that would be inappropriately measured as inefficiency. In the TRE model, a random part that varies with time is introduced in an attempt to separate unobserved time-invariant heterogeneity from inefficiency. This specified as

$$q_{it} = (\alpha + \omega_i) + \beta' x_{it} + v_{it} - u_{it},$$

$$v_{it} \sim N[0,\sigma_v^2]$$

$$u_{it}=|U_{it}| \text{ and } U_{it} \sim N^+[0, \sigma_u^2],$$

$$\omega_i \sim N^+[0, \sigma_\omega^2]$$

where $(\alpha + \omega_i)$ is a time-invariant and country-specific random term meant to capture time-invariant unobserved heterogeneity. Estimation is by maximum simulated likelihood (MSL) by integrating out ω_i using the Monte Carlo method.

3.2 Data

The source of data for the tax and revenue variables is the UNU-WIDER GRD (UNU-WIDER 2021). Thanks to the inclusion of data from IMF Article IV Staff Reports, the GRD contains significantly more complete data for revenue aggregates in LICs and MICs than any other single source.[8] This allows us to study tax effort in a larger sample of countries than has previously been studied in the related literature. The 'economic' and 'demographic' variables are sourced from the World Bank's World Development Indicators (WDI), save for the grants variable, which comes from the OECD Development Assistance Committee (DAC) database, whilst the indicators of governance and institutional quality come from the Varieties of Democracy (V-Dem) dataset (Coppedge et al. 2021). We also, as a robustness check, employ variables from the World Governance Indicators (WGI). The full sample period spans 1980–2019, providing up to 40 years of data, for 192 countries. The total number of potential observations is thus 7,680. Summary statistics are shown in Table 2.

Table 2. Summary statistics

Variable	Obs.	Mean	Std. Dev.	Min.	Max.
Revenue variables (GRD) (all % of GDP)					
Total tax	6,222	16.81	8.29	0	60.95
Economic variables (WDI & DAC)					
GDP per capita (PPP US$)	6,876	11,282.63	16,665.56	164.34	116,232.98
Agriculture (% GDP)	6,168	14.6	13.07	.03	79.04
Exports (% GDP)	6,362	38.58	28.12	.01	228.99
Imports (% GDP)	6,362	44.73	27.22	0	236.39
Trade (% GDP)	6,362	83.31	52.3	.02	442.62
Grants (% GDP)[9]	5,255	6.34	10.47	.00	135.5
Natural resource rents (% GDP)	6,967	6.9	10.75	0	87.46
Demographic variables (WDI)					
Urbanization (%)	7,669	.53	.24	.04	1
Particularistic or public goods	6,635	2.42	0.94	0.028	3.942
V-Dem variables (V-Dem)					
Accountability index	6,635	.48	.96	-1.95	2.06
Rule of law index	6,635	0.53	.31	.014	.998
Public sector corruption index	6,614	.5	.3	.01	.97

Source: authors' calculations.

For the baseline specifications, the dependent variable, q_{it}, is captured by the variable *Total Tax*. This incorporates, broadly, all direct and indirect tax revenues, excluding social security contributions (SSC) and grants.[10] The 'economic' variables contained in x_{it} closely follow the previous literature: we control for GDP per capita and its square, the level of openness to

[8] See McNabb (2017), McNabb et al. (2021), Oppel et al. (2021), and Prichard et al. (2014) for further details on the construction of the data.

[9] We do not present results of models including the *Grants* variable, as the coverage is lower and reduces the number of observations we are able to compute TE scores for. However, when these are included, the scores do not change greatly and our results are qualitatively the same. The results of these estimations are available on request from the authors and online.

[10] There is no best answer to the question of whether the tax variable should be specified inclusive or exclusive of SSC. The main issue arising is that some countries fund social security through SSCs collected by government (e.g. the United Kingdom), some through SSCs collected by a private institution (and thus not reported or captured in government revenue statistics, e.g. Uganda), and some via income taxes (e.g. Denmark). Given that a large share of the countries in our sample are low- or middle-income—where either (i) SSCs are negligibly low (as a % of GDP) or (ii) government revenue data are reported only at central government level and thus social security payments, even if collected by government, are not captured (these are often administered at the 'social security funds' level of government; see McNabb 2017)—we specify the tax variable as exclusive of social security and also include a dummy variable for countries where tax data is reported at the general government level.

international trade (as proxied by the sum of imports (% GDP) + exports (% GDP)), the share of agriculture in GDP, and natural resource rents (as a % of GDP). Whilst a number of previous studies have controlled for the level of government spending on health (e.g. as a % of GDP), the data coverage for this variable is poor and, when tested, it reduced our number of observations substantially. As an alternative, we include a measure of *Particularistic or Public Goods* from the V-Dem dataset (*v2dlencmps*). This index captures the extent to which social and infrastructural spending is deemed to be 'particularistic' (i.e. narrowly focused) or 'public' (i.e. having the characteristics of public goods) in nature. The main demographic factor controlled for is population spread, as proxied by the share of urban population. Turning to measures of governance and institutional quality, we first control for corruption via the *Public Sector Corruption Index* from the V-Dem dataset. This variable considers the extent to which public sector officials are judged to engage in corrupt behaviours.[11] The *Accountability Index* captures the 'constraints on the government's use of political power through requirements for justification for its actions and potential sanctions' (Coppedge et al. 2021: 285), whilst the *Rule of Law Index* captures the extent to which laws are 'transparently, independently, predictably, impartially, and equally enforced' and the extent to which 'the actions of government officials comply with the law' (Coppedge et al. 2021: 269). We also include a dummy variable equal to 1 to capture those countries that report revenue statistics at general government level (equal to zero if at central level).

4 Results

As stated above, four approaches to modelling the frontier are tested, namely Pooled, RE (Pitt and Lee 1981), BC (Battese and Coelli 1995), and TRE (Greene 2005). In Table 3, we display results for a baseline model, which includes only the economic and demographic variables.

The results shown in Table 3 are largely in line with our expectations. Countries reporting tax revenue statistics at general government level have a higher level of total tax, on average (although the RE specification suggests a negative coefficient on this variable). The coefficients on GDP per capita and its square are positive and negative, respectively, suggesting an increasing but nonlinear relationship between level of development and tax revenue mobilization. There is a positive and statistically significant effect on tax revenue of openness to trade, but a higher amount of resource rents in GDP is negatively associated with tax revenue mobilization. This last result might seem counterintuitive—after all, abundant natural resources provide a potentially rich source of tax revenue—but a growing body of literature finds negative effects between resource wealth and domestic revenue mobilization (e.g. Bornhorst et al. 2009; Chachu 2021; Crivelli and Gupta 2014). The finding that countries with a larger agriculture sector collect less tax revenue is well established in the literature and we find confirmation here; the coefficient on agriculture value added (% GDP) is negative and significant. The coefficient on the index capturing the nature of public good provision (where a higher score corresponds to more public rather than particularistic provision) is positive and significant, suggesting that with better public good provision comes higher tax revenue collection (this variable likely proxies for quality of spending on health, education, and infrastructure). Specifications 2, 3, and 4 confirm that, on average, a more urbanized country collects more tax revenue, again in line with our expectations; where the population is more concentrated in urban centres, it is not only easier but also, on average, less costly to collect taxes.

[11] These are defined as 'grant[ing] favors in exchange for bribes, kickbacks, or other material inducements' and in terms of 'how often they steal, embezzle, or misappropriate public funds or other state resources for personal or family use' (Coppedge et al. 2021: 297).

Table 3: SFA baseline results

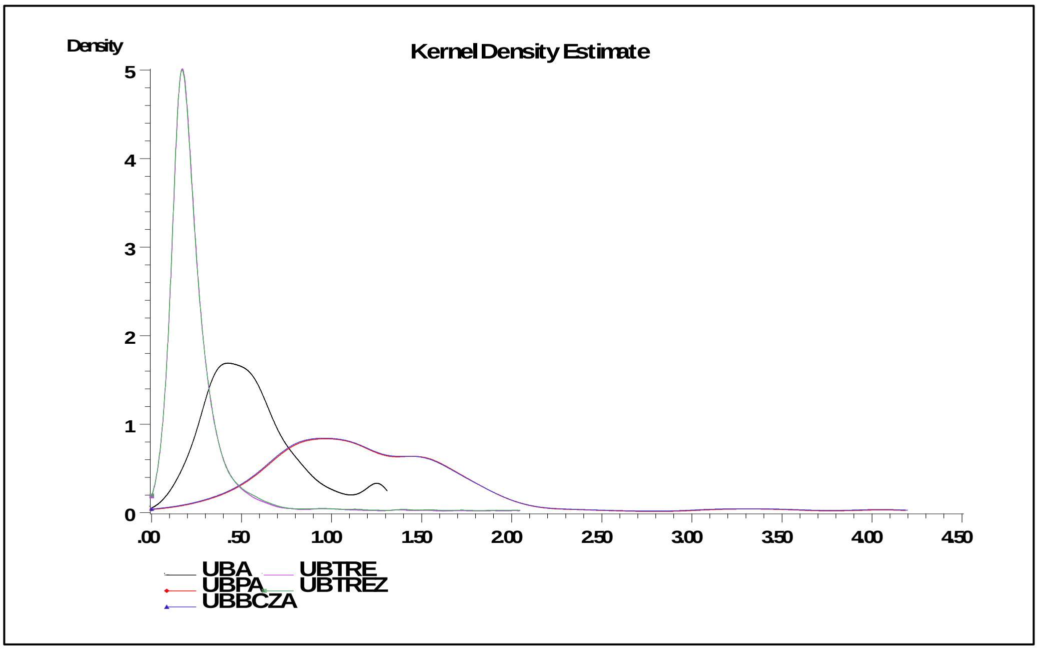
Dependent variable: log total tax	**(I) Pooled**	**(II) RE**	**(III) BC**	**(IV) TRE**
General	0.368***	-0.393***	0.343***	0.207***
	(0.018)	(0.045)	(0.414)	(0.007)
GDP per capita	1.767***	2.270***	2.274***	2.589***
	(0.091)	(0.061)	(0.586)	(0.032)
GDP per capita^2	-0.093***	-0.123***	-0.129***	-0.141***
	(0.005)	(0.004)	(.004)	(0.002)
Trade (% GDP)	0.148***	0.136***	0.140***	0.113***
	(0.013)	(0.005)	(0.005)	(0.004)
Resource rents (% GDP)	-0.015***	-0.019***	-0.024***	-0.024***
	(0.004)	(0.003)	(0.002)	(0.001)
Ag. value added (% GDP)	-0.054***	-0.071***	-0.036***	-0.028***
	(0.014)	(0.004)	(0.005)	(0.004)
Public goods index	0.164***	0.102***	0.095***	0.091***
	(0.009)	(0.004)	(0.004)	(0.003)
Urbanization	-0.194***	0.276***	0.140***	0.317***
	(0.022)	(0.014)	(0.018)	(0.009)
Constant	-6.269***	-6.951***	-7.016***	-9.442***
	(0.407)	(0.256)	(0.248)	(0.407)
Variance parameters				
Sigma	3.059***	1.271***	1.057***	0.287***
	(0.101)	(0.076)	(0.098)	(0.001)
Sigma *w*	-	-	-	1.383***
	-	-	-	(0.007)
Lambda	0.711***	6.505***	5.441***	2.383***
	(0.000)	(1.002)	(0.066)	(0.040)
N	*3901*	*3901*	*3901*	*3901*

Source: authors' estimations.

We now turn to the question of which specification is preferred, i.e. which best captures estimates of countries' tax effort. We start by looking at the kernel densities of the estimated inefficiencies ($-u_{it}$) (Figure 1). The kernel densities show that the TRE model seems more skewed to the left and has a tight variance followed by pooled model. The RE and BC models are more skewed to the right with a higher spread.

Further correlation analysis using estimated inefficiencies for all models (Table 4) shows that the pooled model is highly correlated with the RE and BC models and much less with the TRE model. The RE and BC models are not correlated with the TRE model. The high association between the pooled and the RE and BC models makes these panel models suspect. This is confirmed by observing the variation of the random terms in all models. Sigma, sigma *u*, and sigma *v* are high in the pooled, RE, and BC models compared with the TRE model. This is because in the TRE model some additional variation is now attributed to the time-invariant and country-specific random term (sigma *w*) meant to capture unobserved time-invariant heterogeneity. This means that the TRE is removing or disentangling unobserved heterogeneity from our inefficiency estimates, whilst the RE and BC panel models are not able to treat this. The RE and BC models are carrying both the inefficiency and any time-invariant country-specific heterogeneity. In short, the unobserved time-invariant heterogeneity ends up in the inefficiency estimates of the RE and BC models. In effect, the fact that the random component of the RE and BC models is still time-invariant remains a substantive and detrimental restriction when applying these models to our analysis.

Figure 1: Kernel density estimates of estimated inefficiencies



Source: authors' estimations.

Table 4: Correlation (*-uit*)

	Pooled	**RE**	**BC**	**BCz**	**TRE**	**TREz(I)**	**TREz(II)**	**TREz(III)**
Pooled	**1**	0.67	0.81	0.67	0.50	0.50	0.51	0.51
RE	0.67	**1**	0.86	1.00	0.23	0.23	0.24	0.24
BC	0.81	0.86	**1**	0.86	0.25	0.25	0.26	0.26
BCz	0.67	1.00	0.86	**1**	0.24	0.24	0.24	0.25
TRE	0.50	0.23	0.25	0.24	**1**	0.99	0.99	0.99
TREz(I)	0.50	0.23	0.25	0.24	0.99	**1**	1.00	0.99
TREz(II)	0.51	0.24	0.26	0.24	0.99	1.00	**1**	0.99
TREz(III)	0.51	0.24	0.26	0.25	0.99	0.99	0.99	**1**

Source: authors' estimations.

In the existing literature that has used SFA to estimate TE, Langford and Ohlenburg (2016) and Mawejje and Sebudde (2019) have preferred estimates from RE specifications, whilst Fenochietto et al. (2013) focused on results from BC and Mundlak Random Effects models. The results from each of the four approaches shown in Table 5, however, provide starkly differing estimates of tax effort and tax potential. These are shown in the left-hand panel under 'Standard model'. We display results only for the most recent observation for each country for which all four approaches yielded a result. Estimates are shown for 161 countries in total, a significant improvement in terms of coverage, compared with existing studies.

Table 5: Tax effort estimates across different SFA approaches

Country	Year	Standard model			‘z’ model			
		Pooled	RE	BC	TRE	TREz(I)	TREz(II)	TREz(III)
Albania	2019	0.76	0.54	0.40	0.88	0.88	0.87	0.89
Algeria	2017	0.69	0.24	0.40	0.84	0.84	0.85	0.84
Angola	2019	0.86	0.58	0.89	0.50	0.51	0.50	0.50
Argentina	2019	0.83	0.50	0.47	0.92	0.92	0.92	0.92
Armenia	2012	0.58	0.45	0.34	0.83	0.83	0.82	0.83
Australia	2018	0.85	0.73	0.74	0.91	0.90	0.92	0.91
Austria	2019	0.72	0.69	0.69	0.87	0.86	0.86	0.88
Azerbaijan	2019	0.73	0.30	0.48	0.80	0.79	0.80	0.80
Bahrain	2018	0.28	0.03	0.08	0.40	0.40	0.39	0.39
Bangladesh	2018	0.42	0.22	0.29	0.72	0.71	0.71	0.72
Barbados	2005	0.80	0.42	0.66	0.89	0.88	0.88	0.88
Belarus	2019	0.71	0.60	0.51	0.75	0.74	0.74	0.73
Belgium	2019	0.76	0.56	0.59	0.88	0.87	0.87	0.87
Benin	2016	0.48	0.23	0.32	0.75	0.73	0.74	0.75
Bhutan	2018	0.60	0.21	0.30	0.94	0.94	0.94	0.94
Bolivia	2019	0.75	0.51	0.38	0.86	0.86	0.87	0.86
Bosnia and Herzegovina	2019	0.76	0.68	0.51	0.82	0.82	0.82	0.83
Botswana	2019	0.69	0.33	0.58	0.69	0.69	0.70	0.69
Brazil	2019	0.85	0.57	0.51	0.90	0.90	0.90	0.90
Bulgaria	2019	0.62	0.44	0.39	0.88	0.86	0.87	0.87
Burkina Faso	2019	0.77	0.44	0.51	0.84	0.85	0.89	0.85
Burundi	2014	0.87	0.99	0.97	0.78	0.77	0.80	0.76
Cabo Verde	2018	0.78	0.29	0.44	0.88	0.87	0.87	0.88
Cambodia	2019	0.82	0.32	0.38	0.96	0.96	0.96	0.96
Cameroon	2017	0.87	0.32	0.44	0.92	0.92	0.92	0.91
Canada	2016	0.79	0.66	0.67	0.87	0.87	0.87	0.86
Central African Rep.	2012	0.79	0.39	0.44	0.85	0.85	0.85	0.84
Chad	2018	0.59	0.40	0.43	0.62	0.62	0.60	0.61
Chile	2019	0.77	0.24	0.46	0.92	0.91	0.92	0.91
China	2018	0.59	0.50	0.37	0.77	0.75	0.77	0.75
Colombia	1999	0.62	0.39	0.30	0.87	0.88	0.87	0.89
Comoros	2019	0.43	0.21	0.28	0.84	0.83	0.82	0.83
Congo	2018	0.57	0.16	0.27	0.80	0.78	0.78	0.77
Congo, Dem. Rep.	2018	0.78	0.21	0.25	0.88	0.88	0.87	0.86
Costa Rica	2019	0.56	0.19	0.36	0.83	0.84	0.85	0.86
Côte d’Ivoire	2018	0.57	0.26	0.37	0.78	0.78	0.79	0.80
Croatia	2019	0.73	0.64	0.56	0.88	0.87	0.87	0.87
Cyprus	2019	0.83	0.31	0.60	0.90	0.89	0.91	0.90
Czech Rep.	2019	0.58	0.40	0.40	0.90	0.89	0.89	0.90
Denmark	2019	0.90	0.98	0.98	0.87	0.88	0.88	0.85
Djibouti	2018	0.56	0.17	0.27	0.82	0.81	0.81	0.78
Dominican Rep.	2019	0.71	0.21	0.37	0.89	0.88	0.89	0.89
Ecuador	2019	0.67	0.18	0.31	0.93	0.94	0.93	0.93
Egypt	2019	0.67	0.33	0.52	0.76	0.75	0.77	0.74
El Salvador	2019	0.74	0.22	0.37	0.91	0.91	0.91	0.91
Equatorial Guinea	2019	0.53	0.15	0.31	0.80	0.82	0.81	0.80
Estonia	2019	0.58	0.42	0.40	0.93	0.90	0.92	0.91
Eswatini	2019	0.88	0.52	0.69	0.91	0.91	0.91	0.91
Ethiopia	2019	0.41	0.71	0.38	0.64	0.64	0.66	0.64

Country	Year	Standard model				'z' model		
		Pooled	RE	BC	TRE	TREz(I)	TREz(II)	TREz(III)
Fiji	2018	0.80	0.34	0.55	0.88	0.88	0.89	0.90
Finland	2019	0.82	0.71	0.71	0.89	0.87	0.87	0.88
France	2019	0.82	0.60	0.61	0.93	0.92	0.92	0.92
Gabon	1996	0.82	0.26	0.47	0.92	0.92	0.92	0.93
Gambia, Rep. of The	2019	0.64	0.22	0.31	0.89	0.91	0.89	0.91
Georgia	2019	0.58	0.48	0.37	0.83	0.82	0.84	0.82
Germany	2019	0.68	0.48	0.52	0.92	0.91	0.91	0.92
Ghana	2019	0.60	0.19	0.28	0.92	0.92	0.92	0.93
Greece	2019	0.78	0.55	0.52	0.91	0.91	0.90	0.91
Guatemala	2019	0.65	0.23	0.35	0.86	0.84	0.86	0.86
Guinea	2019	0.75	0.33	0.42	0.84	0.85	0.85	0.83
Guinea Bissau	2019	0.66	0.16	0.22	0.96	0.96	0.96	0.96
Guyana	2005	0.63	0.33	0.44	0.86	0.86	0.85	0.86
Haiti	2019	0.49	0.17	0.24	0.78	0.78	0.81	0.77
Honduras	2019	0.75	0.50	0.36	0.90	0.90	0.90	0.91
Hong Kong	2018	0.36	0.17	0.21	0.93	0.91	0.92	0.92
Hungary	2019	0.78	0.58	0.52	0.88	0.86	0.88	0.88
Iceland	2019	0.81	0.69	0.67	0.83	0.82	0.81	0.82
India	2016	0.58	0.66	0.41	0.78	0.76	0.75	0.74
Indonesia	2019	0.47	0.24	0.39	0.62	0.63	0.64	0.63
Iran, Islamic Rep.	2014	0.38	0.09	0.18	0.93	0.91	0.91	0.87
Iraq	2009	0.28	0.02	0.02	0.97	0.96	0.97	0.97
Ireland	2019	0.52	0.56	0.58	0.78	0.74	0.77	0.76
Israel	2018	0.76	0.55	0.56	0.87	0.87	0.87	0.88
Italy	2019	0.78	0.72	0.70	0.83	0.84	0.84	0.84
Jamaica	2019	0.82	0.35	0.58	0.88	0.88	0.89	0.86
Japan	2018	0.61	0.38	0.40	0.91	0.91	0.91	0.91
Jordan	2019	0.68	0.23	0.42	0.72	0.71	0.72	0.70
Kazakhstan	2018	0.72	0.48	0.39	0.91	0.90	0.89	0.90
Kenya	2019	0.81	0.48	0.58	0.80	0.81	0.81	0.80
Korea, Rep.	2019	0.57	0.37	0.37	0.93	0.93	0.93	0.94
Kuwait	2018	0.28	0.02	0.05	0.87	0.92	0.91	0.89
Kyrgyzstan	2019	0.62	0.57	0.37	0.87	0.88	0.88	0.89
Lao PDR	2016	0.39	0.29	0.19	0.91	0.92	0.91	0.91
Latvia	2019	0.61	0.44	0.40	0.91	0.90	0.91	0.90
Lebanon	2019	0.74	0.21	0.39	0.85	0.87	0.86	0.86
Lesotho	2019	0.90	0.93	0.98	0.59	0.60	0.62	0.61
Liberia	2019	0.81	0.32	0.40	0.92	0.92	0.92	0.91
Libya	2008	0.28	0.07	0.07	0.92	0.91	0.91	0.93
Lithuania	2019	0.57	0.40	0.37	0.91	0.90	0.92	0.92
Luxembourg	2019	0.76	0.63	0.75	0.88	0.88	0.89	0.89
Madagascar	2019	0.71	0.32	0.39	0.87	0.87	0.88	0.88
Malawi	2019	0.85	0.60	0.61	0.94	0.93	0.94	0.93
Malaysia	2019	0.59	0.25	0.47	0.67	0.68	0.69	0.69
Maldives	2019	0.60	0.29	0.46	0.82	0.84	0.82	0.83
Mali	2019	0.78	0.37	0.46	0.86	0.84	0.85	0.84
Mauritania	2019	0.60	0.20	0.31	0.90	0.90	0.90	0.92
Mauritius	2019	0.71	0.23	0.39	0.95	0.95	0.95	0.96
Mexico	2019	0.68	0.16	0.32	0.94	0.94	0.94	0.95
Moldova	2019	0.68	0.52	0.38	0.90	0.89	0.89	0.89
Mongolia	2019	0.67	0.46	0.36	0.89	0.88	0.88	0.88

Country	Year	Standard model				'z' model		
		Pooled	RE	BC	TRE	TREz(I)	TREz(II)	TREz(III)
Montenegro	2019	0.84	0.64	0.53	0.90	0.90	0.91	0.91
Morocco	2019	0.82	0.41	0.61	0.75	0.74	0.74	0.72
Mozambique	2019	0.95	0.57	0.64	0.96	0.97	0.97	0.97
Myanmar	2018	0.28	0.22	0.27	0.33	0.32	0.33	0.31
Namibia	2019	0.89	0.52	0.76	0.84	0.83	0.82	0.83
Nepal	2019	0.80	0.36	0.41	0.95	0.96	0.96	0.96
Netherlands	2019	0.69	0.50	0.53	0.91	0.90	0.91	0.91
New Zealand	2017	0.86	0.81	0.78	0.87	0.87	0.86	0.88
Nicaragua	2019	0.78	0.43	0.31	0.95	0.95	0.95	0.95
Niger	2019	0.63	0.42	0.45	0.86	0.85	0.86	0.83
Nigeria	2007	0.40	0.35	0.22	0.75	0.76	0.75	0.74
North Macedonia	2019	0.67	0.30	0.50	0.78	0.75	0.77	0.76
Norway	2019	0.82	0.81	0.85	0.84	0.83	0.87	0.81
Oman	2013	0.28	0.04	0.10	0.67	0.64	0.63	0.61
Pakistan	2015	0.77	0.32	0.42	0.89	0.89	0.89	0.90
Panama	2019	0.32	0.23	0.22	0.81	0.79	0.79	0.82
Papua New Guinea	2004	0.72	0.48	0.53	0.92	0.92	0.93	0.93
Paraguay	2019	0.57	0.18	0.30	0.88	0.86	0.88	0.87
Peru	2019	0.68	0.25	0.43	0.79	0.79	0.80	0.79
Philippines	2019	0.61	0.24	0.37	0.87	0.85	0.87	0.88
Poland	2019	0.62	0.48	0.44	0.90	0.88	0.89	0.88
Portugal	2019	0.63	0.52	0.49	0.86	0.87	0.88	0.88
Qatar	2008	0.31	0.04	0.12	0.97	0.97	0.97	0.97
Romania	2019	0.55	0.51	0.44	0.78	0.72	0.75	0.78
Russian Federation	2019	0.84	0.67	0.61	0.90	0.88	0.89	0.89
Rwanda	2019	0.77	0.60	0.63	0.77	0.79	0.80	0.77
Saudi Arabia	2018	0.31	0.04	0.09	0.98	0.98	0.98	0.98
Senegal	2019	0.82	0.34	0.46	0.91	0.90	0.92	0.90
Serbia	2018	0.71	0.58	0.47	0.88	0.86	0.86	0.86
Seychelles	2019	0.87	0.40	0.71	0.90	0.90	0.90	0.89
Sierra Leone	2019	0.74	0.30	0.38	0.92	0.92	0.92	0.93
Singapore	2019	0.47	0.13	0.36	0.92	0.91	0.90	0.92
Slovak Republic	2019	0.48	0.39	0.36	0.89	0.88	0.88	0.88
Slovenia	2019	0.56	0.51	0.47	0.86	0.85	0.85	0.85
Solomon Islands	2006	0.81	0.58	0.62	0.85	0.87	0.86	0.89
South Africa	2019	0.77	0.61	0.52	0.87	0.88	0.87	0.89
Spain	2019	0.67	0.48	0.48	0.89	0.88	0.88	0.88
Sri Lanka	2019	0.42	0.30	0.42	0.73	0.71	0.73	0.73
Sudan	2019	0.41	0.22	0.30	0.67	0.70	0.67	0.71
Suriname	2010	0.69	0.28	0.47	0.81	0.81	0.80	0.82
Sweden	2019	0.86	0.73	0.75	0.91	0.90	0.89	0.91
Switzerland	2019	0.61	0.44	0.48	0.91	0.91	0.90	0.91
Tajikistan	2017	0.86	0.78	0.44	0.94	0.94	0.94	0.95
Tanzania	2017	0.70	0.33	0.40	0.87	0.87	0.88	0.88
Thailand	2019	0.57	0.46	0.36	0.84	0.82	0.84	0.83
Timor-Leste	2015	0.28	0.13	0.16	0.61	0.61	0.57	0.62
Togo	2019	0.76	0.34	0.42	0.88	0.88	0.89	0.89
Tunisia	2018	0.60	0.47	0.38	0.87	0.87	0.87	0.87
Turkey	2019	0.69	0.48	0.42	0.84	0.82	0.82	0.83
Turkmenistan	2008	0.84	0.19	0.27	0.96	0.96	0.96	0.96
Uganda	2019	0.62	0.36	0.41	0.84	0.86	0.86	0.85

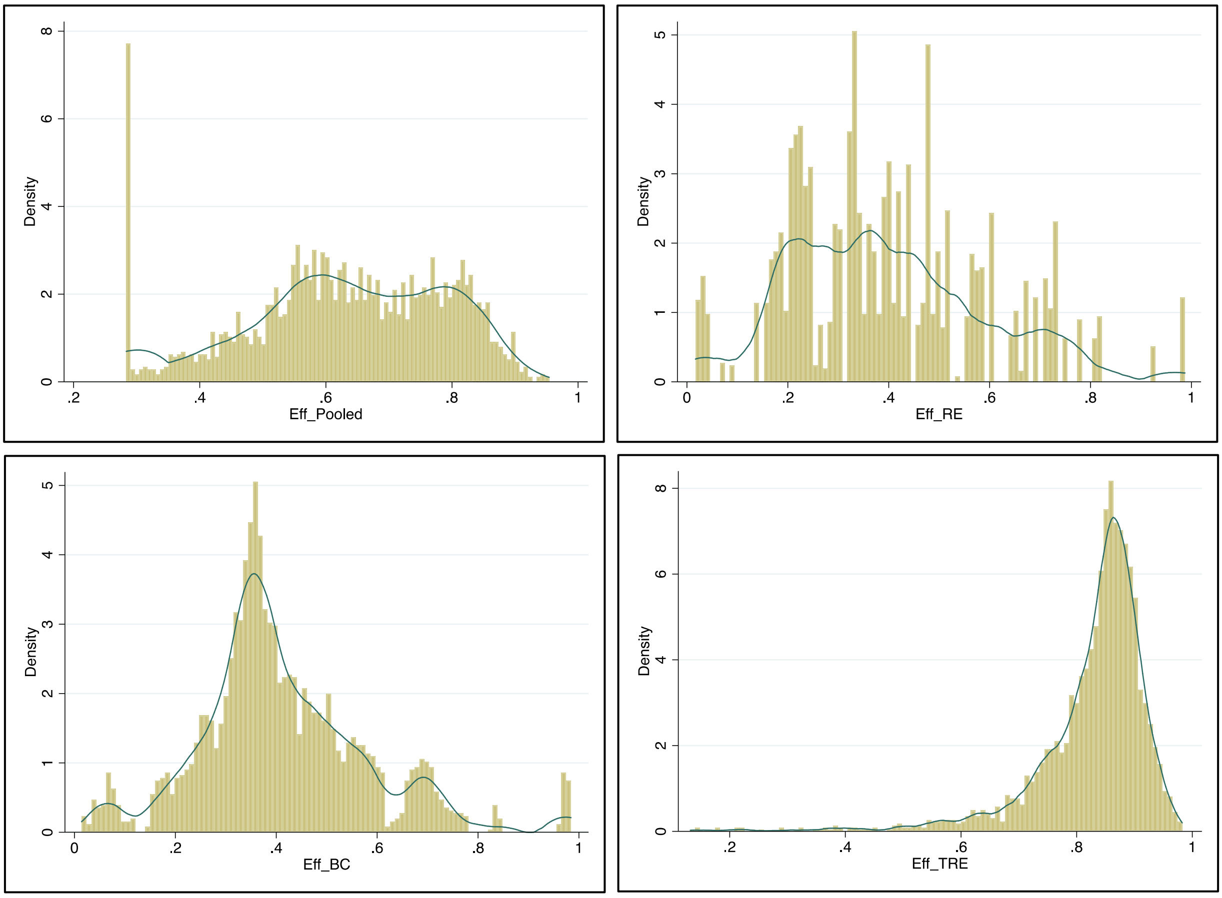
Country	Year		Standard model			'z' model		
		Pooled	RE	BC	TRE	TREz(I)	TREz(II)	TREz(III)
Ukraine	2019	0.77	0.60	0.49	0.86	0.85	0.86	0.85
United Arab Emirates	2019	0.55	0.39	0.44	0.80	0.78	0.78	0.77
United Kingdom	2019	0.77	0.59	0.61	0.88	0.88	0.89	0.88
United States	2018	0.73	0.58	0.62	0.86	0.84	0.85	0.86
Uruguay	2016	0.67	0.34	0.32	0.95	0.95	0.95	0.96
Uzbekistan	2019	0.69	0.75	0.50	0.65	0.65	0.66	0.64
Vanuatu	2014	0.77	0.45	0.54	0.87	0.87	0.88	0.87
Viet Nam	2019	0.42	0.49	0.33	0.66	0.66	0.69	0.65
Zambia	2019	0.71	0.38	0.51	0.74	0.74	0.75	0.74
Zimbabwe	2018	0.77	0.42	0.56	0.64	0.64	0.63	0.63
World average	*2017*	*0.67*	*0.41*	*0.44*	*0.84*	*0.84*	*0.84*	*0.84*

Source: authors' calculations.

We see that, on average, the TRE and Pooled estimates (in columns 4 and 1, respectively) show the highest TE scores (at 0.84 and 0.67), followed by BC (0.44) and RE (0.41). The couple of outliers could explain the low estimated efficiency of the RE and BC models (see Kumbhakar et al. 2015).

In Figure 2, we plot the distribution of the TE scores (n=3,901) across the four estimation methods tested.

Figure 2: Distribution of tax effort estimates, by estimation method



Source: authors' calculations.

We see that the distribution of scores under the TRE model is a lot tighter and more rightward skewed than the other approaches. The variance is also lower.

In assessing which model is most appropriate for tax effort calculations, it is also, we feel, important to use intuition. Whilst the diagnostics suggest that TRE performs better, we also believe it to be the case that scores of the magnitude of those estimated by TRE are more likely to represent the gap between actual and potential revenue collection across countries. We must remember that the tax effort scores are estimated using a set of underlying economic, demographic, and institutional variables from each country. It is important to examine these input variables to help us better understand the most appropriate modelling approach. A couple of country-specific cases are illustrative.

Take the case of Slovakia. It currently collects 19.82 per cent of GDP in taxes (excluding SSC). The BC estimates suggest that, with a tax effort score of 0.36, its tax potential is 55 per cent. This estimate of tax effort seems extremely low, especially compared with the TRE score of 0.89, which puts tax potential at around 22.3 per cent. Slovakia does not, for example, have abundant natural resources (rents stand at 0.25 per cent of GDP in 2018), nor a particularly large GDP per capita (at \$31,888, it ranks around 40th worldwide), nor an overwhelmingly large urban population (54 per cent). Slovakia *does*, however, have a very high ratio of trade to GDP, at 191 per cent (for 2018, this ranked 9th in the world, according to the WDI data at hand). However, high trade-to-GDP values are observed overwhelmingly in small landlocked countries or island nations. Furthermore, the vast majority of Slovakia's trade is intra-EU, on which no VAT or customs duties are levied. Thus, it may be the case that the BC models are overly sensitive to extreme values of particular input variables—in this case the high ratio of trade to GDP—and produce a skewed estimate of tax effort. In reality, for the case of Slovakia, that trade does not allow additional taxes to be levied. On the other hand, the TRE approach appears to be less affected by outlying variables.

Take, as another example, Burundi. The most recent year for which we have TE estimates is 2014, in which actual tax collections were 13.67 per cent. None of the underlying input variables shows particularly extreme values, save for the urbanization rate, where Burundi ranks bottom of every country in the sample at just 11.8 per cent. The BC estimate of tax effort is at 0.97, suggesting that Burundi's tax potential is just 14.1 per cent at present. Again, it seems that the estimate of TE is heavily skewed upward by one extreme value of one input variable. The TRE estimate of 0.78 seems somewhat more realistic for a country of Burundi's economic stature, suggesting that tax potential lies closer to 17.5 per cent.

Thus, the lower variance of the TE estimates from the TRE models are less skewed by outlying values of input variables. They may, then, prove to be more conservative estimates of tax effort, but should also, in such cases, be more 'realistic'.

4.1 Incorporating inefficiency components

We proceed to estimate the stochastic tax frontier according to the TRE procedure and, in Table 6, show results when we include institutional factors z_c into the equation. These are included as inputs (see Faust and Baranzini 2014).[12] We test three separate specifications. In column I, we control for the level of public sector corruption, accountability, and the rule of law. The signs are as expected, and each statistically significant. In columns II and III we omit, respectively, the Rule

[12] The TRE approach only allows for components of z to enter as inputs in the production frontier.

of Law and Accountability indices; the results do not substantively change, although the sign flips on the Rule of Law Index to become negative in column III.

Table 6: Incorporating inefficiency components as inputs

Dependent variable: Log total tax	**TRE(I)**	**TRE(II)**	**TRE(III)**
General	0.350***	0.044***	0.194***
	(0.006)	(0.007)	(0.007)
GDP per capita	2.378***	2.405***	2.544***
	(0.033)	(0.032)	(0.033)
GDP per capita^2	-0.127***	-0.130***	-0.137***
	(0.002)	(0.002)	(0.004)
Trade (% GDP)	0.124***	0.113***	0.137***
	(0.004)	(0.004)	(0.004)
Resource rents (% GDP)	-0.028***	-0.015	-0.034***
	(0.001)	(0.001)	(0.001)
Ag. value added (% GDP)	-0.024***	-0.030***	-0.016***
	(0.004)	(0.004)	(0.003)
Public goods index	0.086***	0.080***	0.088***
	(0.003)	(0.003)	(0.003)
Urbanisation	0.287***	0.276***	0.267****
	(0.009)	(0.008)	(0.008)
Constant	-8.875***	-8.962***	-9.632***
	(0.141)	(0.140)	(0.142)
Public sector corruption	-0.133***	-0.022***	-0.039**
	(0.023)	(0.012)	(0.018)
Accountability index	0.028*	-0.015***	
	(0.015)	(0.011)	
Rule of law index	0.175***		-0.083***
	(0.029)		(0.021)
N	*3901*	*3901*	*3901*
Sigma	0.287***	0.290***	0.295***
	(0.001)	(0.001)	(0.001)
Sigma *w*	1.120**	1.536***	1.277***
	(0.001)	(0.010)	(0.008)
Lambda	2.430***	2.494***	2.656***
	(0.043)	(0.044)	(0.045)

Source: authors' calculations.

The resulting estimates for tax effort are shown above in Table 4, under the 'z' model columns. We see that the inclusion of the inefficiency variables very slightly affects the estimated tax effort scores.

What, then, do these results tell us about tax *potential*? This is simply estimated as $\frac{T_{it}}{q_{it}}$. We show the results for our favoured specification, TRE, both with and without the controls for institutional quality found in z_c. In Table 7, we calculate tax potential for the TRE specifications IV (Table 3) and I (Table 6), as estimated above, and show the percentage change between tax potential (TREz) and total tax (current).

Table 7: Tax potential calculations

Country	Year	TRE	TRE (z)	Total tax (current)	Tax potential (TRE)	Tax potential (TREz)	% change
Albania	2019	0.88	0.88	19.40%	21.93%	21.94%	13.10%
Algeria	2017	0.84	0.84	14.11%	16.85%	16.70%	18.36%
Angola	2019	0.50	0.51	18.53%	36.80%	36.14%	95.05%
Argentina	2019	0.92	0.92	22.93%	24.89%	24.82%	8.23%
Armenia	2012	0.83	0.83	17.96%	21.64%	21.57%	20.12%
Australia	2018	0.91	0.90	29.40%	32.42%	32.51%	10.60%
Austria	2019	0.87	0.86	27.70%	31.83%	32.12%	15.96%
Azerbaijan	2019	0.80	0.79	14.26%	17.84%	18.15%	27.26%
Bahrain	2018	0.40	0.40	1.15%	2.89%	2.84%	147.46%
Bangladesh	2018	0.72	0.71	7.90%	10.98%	11.06%	40.02%
Barbados	2005	0.89	0.88	26.73%	30.17%	30.26%	13.22%
Belarus	2019	0.75	0.74	23.80%	31.90%	32.25%	35.50%
Belgium	2019	0.88	0.87	29.44%	33.58%	33.81%	14.84%
Benin	2016	0.75	0.73	9.15%	12.24%	12.47%	36.20%
Bhutan	2018	0.94	0.94	16.40%	17.48%	17.50%	6.68%
Bolivia	2019	0.86	0.86	18.50%	21.53%	21.46%	16.00%
Bosnia and Herzegovina	2019	0.82	0.82	22.71%	27.70%	27.67%	21.85%
Botswana	2019	0.69	0.69	19.77%	28.66%	28.61%	44.72%
Brazil	2019	0.90	0.90	24.09%	26.63%	26.86%	11.50%
Bulgaria	2019	0.88	0.86	21.20%	24.18%	24.67%	16.36%
Burkina Faso	2019	0.84	0.85	15.93%	18.91%	18.83%	18.22%
Burundi	2014	0.78	0.77	13.67%	17.55%	17.78%	30.06%
Cabo Verde	2018	0.88	0.87	21.27%	24.29%	24.59%	15.59%
Cambodia	2019	0.96	0.96	20.00%	20.84%	20.80%	3.98%
Cameroon	2017	0.92	0.92	13.39%	14.63%	14.61%	9.08%
Canada	2016	0.87	0.87	28.37%	32.55%	32.61%	14.96%
Central African Rep.	2012	0.85	0.85	8.17%	9.64%	9.63%	17.90%
Chad	2018	0.62	0.62	7.13%	11.50%	11.45%	60.49%
Chile	2019	0.92	0.91	17.80%	19.35%	19.62%	10.20%
China	2018	0.77	0.75	18.52%	23.93%	24.75%	33.65%
Colombia	1999	0.87	0.88	14.19%	16.24%	16.10%	13.43%
Comoros	2019	0.84	0.83	6.95%	8.32%	8.33%	19.83%
Congo	2018	0.80	0.78	7.81%	9.77%	9.96%	27.50%
Congo, Dem. Rep.	2018	0.88	0.88	7.08%	8.05%	8.08%	14.17%
Costa Rica	2019	0.83	0.84	13.43%	16.18%	15.92%	18.51%
Côte d’Ivoire	2018	0.78	0.78	12.05%	15.50%	15.45%	28.24%
Croatia	2019	0.88	0.87	26.85%	30.68%	30.70%	14.32%
Cyprus	2019	0.90	0.89	24.21%	26.81%	27.11%	11.99%
Czech Rep.	2019	0.90	0.89	19.46%	21.66%	21.84%	12.22%
Denmark	2019	0.87	0.88	45.89%	52.85%	52.42%	14.22%
Djibouti	2018	0.82	0.81	12.89%	15.65%	15.95%	23.72%
Dominican Rep.	2019	0.89	0.88	13.30%	15.01%	15.18%	14.16%
Ecuador	2019	0.93	0.94	14.04%	15.04%	15.01%	6.94%
Egypt	2019	0.76	0.75	13.83%	18.12%	18.47%	33.55%
El Salvador	2019	0.91	0.91	17.99%	19.72%	19.75%	9.79%
Equatorial Guinea	2019	0.80	0.82	9.59%	12.03%	11.67%	21.67%
Estonia	2019	0.93	0.90	21.49%	23.20%	23.92%	11.31%
Eswatini	2019	0.91	0.91	25.33%	27.72%	27.95%	10.36%
Ethiopia	2019	0.64	0.64	9.98%	15.53%	15.62%	56.57%

Country	Year	TRE	TRE (z)	Total tax (current)	Tax potential (TRE)	Tax potential (TREz)	% change
Fiji	2018	0.88	0.88	23.99%	27.24%	27.30%	13.80%
Finland	2019	0.89	0.87	30.46%	34.26%	34.83%	14.35%
France	2019	0.93	0.92	30.47%	32.92%	33.05%	8.46%
Gabon	1996	0.92	0.92	20.49%	22.18%	22.18%	8.26%
Gambia, Rep. of The	2019	0.89	0.91	10.91%	12.20%	12.03%	10.24%
Georgia	2019	0.83	0.82	23.18%	28.08%	28.14%	21.42%
Germany	2019	0.92	0.91	24.12%	26.34%	26.49%	9.81%
Ghana	2019	0.92	0.92	12.28%	13.31%	13.34%	8.61%
Greece	2019	0.91	0.91	27.39%	30.12%	30.17%	10.16%
Guatemala	2019	0.86	0.84	10.50%	12.24%	12.45%	18.53%
Guinea	2019	0.84	0.85	12.52%	14.89%	14.76%	17.92%
Guinea Bissau	2019	0.96	0.96	9.38%	9.80%	9.80%	4.46%
Guyana	2005	0.86	0.86	15.46%	18.04%	17.98%	16.29%
Haiti	2019	0.78	0.78	6.09%	7.81%	7.81%	28.37%
Honduras	2019	0.90	0.90	18.32%	20.38%	20.46%	11.68%
Hong Kong	2018	0.93	0.91	13.81%	14.90%	15.11%	9.45%
Hungary	2019	0.88	0.86	23.77%	27.14%	27.54%	15.85%
Iceland	2019	0.83	0.82	31.90%	38.60%	39.12%	22.64%
India	2016	0.78	0.76	17.35%	22.20%	22.89%	31.91%
Indonesia	2019	0.62	0.63	9.76%	15.74%	15.53%	59.16%
Iran, Islamic Rep.	2014	0.93	0.91	6.16%	6.64%	6.80%	10.35%
Iraq	2009	0.97	0.96	1.38%	1.42%	1.44%	4.20%
Ireland	2019	0.78	0.74	18.24%	23.54%	24.64%	35.08%
Israel	2018	0.87	0.87	25.76%	29.45%	29.54%	14.67%
Italy	2019	0.83	0.84	29.14%	35.32%	34.62%	18.81%
Jamaica	2019	0.88	0.88	26.96%	30.69%	30.73%	13.99%
Japan	2018	0.91	0.91	18.87%	20.72%	20.85%	10.50%
Jordan	2019	0.72	0.71	14.81%	20.55%	20.75%	40.13%
Kazakhstan	2018	0.91	0.90	17.94%	19.74%	20.04%	11.73%
Kenya	2019	0.80	0.81	15.86%	19.75%	19.67%	24.04%
Korea, Rep.	2019	0.93	0.93	20.01%	21.57%	21.55%	7.70%
Kuwait	2018	0.87	0.92	1.41%	1.62%	1.53%	8.27%
Kyrgyzstan	2019	0.87	0.88	20.31%	23.23%	23.09%	13.70%
Lao PDR	2016	0.91	0.92	12.43%	13.67%	13.58%	9.29%
Latvia	2019	0.91	0.90	21.69%	23.87%	24.22%	11.66%
Lebanon	2019	0.85	0.87	15.59%	18.36%	18.01%	15.55%
Lesotho	2019	0.59	0.60	33.22%	56.13%	55.22%	66.22%
Liberia	2019	0.92	0.92	12.60%	13.72%	13.73%	8.96%
Libya	2008	0.92	0.91	3.91%	4.25%	4.28%	9.63%
Lithuania	2019	0.91	0.90	20.47%	22.43%	22.62%	10.52%
Luxembourg	2019	0.88	0.88	28.45%	32.38%	32.17%	13.07%
Madagascar	2019	0.87	0.87	10.23%	11.70%	11.72%	14.55%
Malawi	2019	0.94	0.93	17.28%	18.42%	18.49%	7.02%
Malaysia	2019	0.67	0.68	11.95%	17.78%	17.58%	47.12%
Maldives	2019	0.82	0.84	19.05%	23.18%	22.79%	19.65%
Mali	2019	0.86	0.84	14.75%	17.08%	17.48%	18.48%
Mauritania	2019	0.90	0.90	12.17%	13.59%	13.46%	10.60%
Mauritius	2019	0.95	0.95	19.73%	20.75%	20.76%	5.22%
Mexico	2019	0.94	0.94	13.15%	14.01%	13.98%	6.33%
Moldova	2019	0.90	0.89	19.06%	21.12%	21.33%	11.89%

Country	Year	TRE	TRE (z)	Total tax (current)	Tax potential (TRE)	Tax potential (TREz)	% change
Mongolia	2019	0.89	0.88	23.49%	26.49%	26.56%	13.06%
Montenegro	2019	0.90	0.90	27.17%	30.28%	30.16%	11.01%
Morocco	2019	0.75	0.74	21.45%	28.71%	29.00%	35.19%
Mozambique	2019	0.96	0.97	25.05%	25.96%	25.94%	3.54%
Myanmar	2018	0.33	0.32	2.98%	9.06%	9.25%	210.46%
Namibia	2019	0.84	0.83	30.25%	36.09%	36.42%	20.38%
Nepal	2019	0.95	0.96	19.36%	20.28%	20.26%	4.65%
Netherlands	2019	0.91	0.90	25.85%	28.43%	28.78%	11.35%
New Zealand	2017	0.87	0.87	32.31%	37.15%	37.35%	15.61%
Nicaragua	2019	0.95	0.95	19.23%	20.20%	20.30%	5.57%
Niger	2019	0.86	0.85	10.36%	12.01%	12.13%	17.08%
Nigeria	2007	0.75	0.76	7.17%	9.57%	9.48%	32.21%
North Macedonia	2019	0.78	0.75	16.93%	21.76%	22.50%	32.93%
Norway	2019	0.84	0.83	29.15%	34.59%	34.93%	19.82%
Oman	2013	0.67	0.64	2.53%	3.77%	3.95%	56.05%
Pakistan	2015	0.89	0.89	11.40%	12.81%	12.87%	12.85%
Panama	2019	0.81	0.79	8.20%	10.09%	10.32%	25.80%
Papua New Guinea	2004	0.92	0.92	15.92%	17.33%	17.27%	8.48%
Paraguay	2019	0.88	0.86	9.82%	11.19%	11.38%	15.87%
Peru	2019	0.79	0.79	14.17%	17.98%	17.84%	25.93%
Philippines	2019	0.87	0.85	14.49%	16.60%	16.96%	17.01%
Poland	2019	0.90	0.88	21.95%	24.32%	24.95%	13.69%
Portugal	2019	0.86	0.87	24.95%	28.85%	28.60%	14.63%
Qatar	2008	0.97	0.97	4.33%	4.48%	4.49%	3.63%
Romania	2019	0.78	0.72	15.43%	19.72%	21.33%	38.26%
Russian Federation	2019	0.90	0.88	26.00%	28.95%	29.40%	13.06%
Rwanda	2019	0.77	0.79	16.74%	21.67%	21.18%	26.53%
Saudi Arabia	2018	0.98	0.98	4.75%	4.83%	4.83%	1.82%
Senegal	2019	0.91	0.90	17.65%	19.43%	19.54%	10.69%
Serbia	2018	0.88	0.86	23.69%	27.02%	27.63%	16.62%
Seychelles	2019	0.90	0.90	33.46%	37.00%	37.20%	11.19%
Sierra Leone	2019	0.92	0.92	12.28%	13.41%	13.31%	8.38%
Singapore	2019	0.92	0.91	13.24%	14.46%	14.61%	10.33%
Slovak Republic	2019	0.89	0.88	19.82%	22.39%	22.54%	13.73%
Slovenia	2019	0.86	0.85	21.67%	25.33%	25.51%	17.70%
Solomon Islands	2006	0.85	0.87	16.57%	19.60%	18.94%	14.32%
South Africa	2019	0.87	0.88	28.47%	32.76%	32.47%	14.05%
Spain	2019	0.89	0.88	22.40%	25.22%	25.50%	13.86%
Sri Lanka	2019	0.73	0.71	11.56%	15.87%	16.31%	41.06%
Sudan	2019	0.67	0.70	5.44%	8.15%	7.75%	42.35%
Suriname	2010	0.81	0.81	15.70%	19.49%	19.29%	22.87%
Sweden	2019	0.91	0.90	33.65%	37.14%	37.24%	10.67%
Switzerland	2019	0.91	0.91	20.97%	22.97%	23.04%	9.89%
Tajikistan	2017	0.94	0.94	21.55%	23.02%	22.84%	5.98%
Tanzania	2017	0.87	0.87	11.84%	13.67%	13.62%	15.02%
Thailand	2019	0.84	0.82	16.13%	19.24%	19.69%	22.06%
Timor-Leste	2015	0.61	0.61	3.87%	6.37%	6.30%	62.73%
Togo	2019	0.88	0.88	13.15%	14.89%	14.89%	13.26%
Tunisia	2018	0.87	0.87	22.59%	26.02%	26.08%	15.46%
Turkey	2019	0.84	0.82	15.84%	18.92%	19.32%	21.97%

Country	Year	TRE	TRE (z)	Total tax (current)	Tax potential (TRE)	Tax potential (TREz)	% change
Turkmenistan	2008	0.96	0.96	18.62%	19.42%	19.42%	4.32%
Uganda	2019	0.84	0.86	11.47%	13.61%	13.41%	16.89%
Ukraine	2019	0.86	0.85	25.46%	29.73%	29.92%	17.51%
United Arab Emirates	2019	0.80	0.78	14.78%	18.55%	18.93%	28.09%
United Kingdom	2019	0.88	0.88	26.53%	30.10%	30.19%	13.80%
United States	2018	0.86	0.84	18.35%	21.24%	21.73%	18.40%
Uruguay	2016	0.95	0.95	21.82%	22.90%	22.86%	4.79%
Uzbekistan	2019	0.65	0.65	19.19%	29.72%	29.71%	54.82%
Vanuatu	2014	0.87	0.87	18.34%	21.06%	21.05%	14.76%
Viet Nam	2019	0.66	0.66	14.55%	22.07%	22.06%	51.61%
Zambia	2019	0.74	0.74	16.68%	22.57%	22.61%	35.57%
Zimbabwe	2018	0.64	0.64	11.63%	18.07%	18.10%	55.59%
World average	*2017*	*0.84*	*0.84*	*17.65%*	*20.83%*	*20.91%*	*21.53%*

Source: authors' calculations.

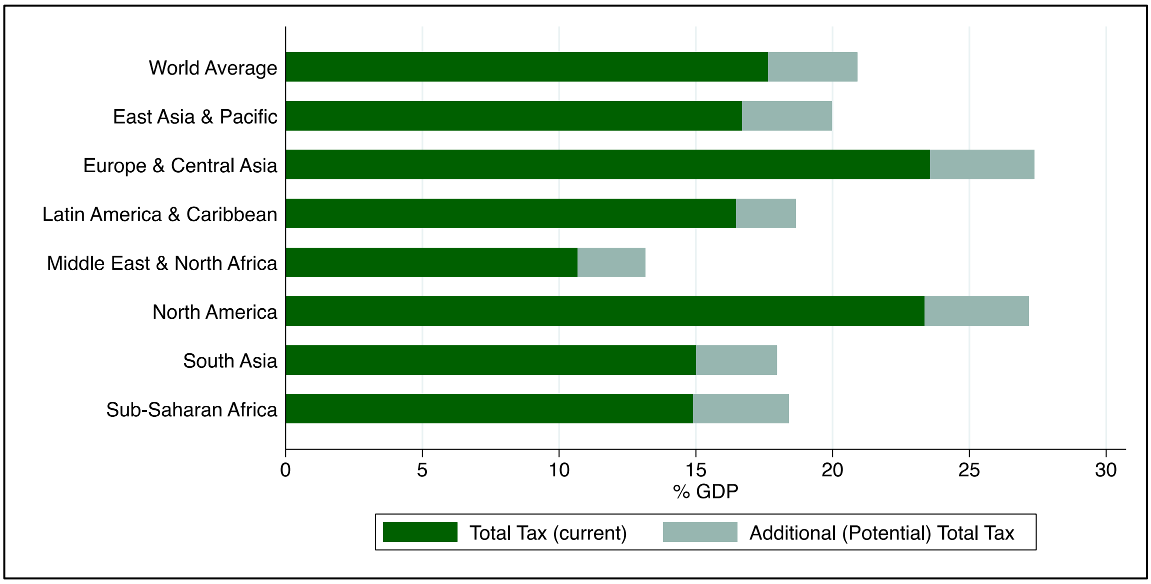
The average (worldwide) tax effort score is 0.84 (across both specifications) and tax potential is 20.91 (TREz specification). This represents an average increase of 21.53 per cent in tax revenues (or 3.26 percentage points). In Table 8 and Figures 3 and 4, we summarize these scores by averaging across regions and income groups (as defined by the World Bank). We see that tax effort is lowest, on average, in the East Asia and Pacific region, at around 0.81, and highest in the Latin America and Caribbean region, at 0.87. Interestingly, tax effort is not found to be lowest in LICs, but in LMICs, at 0.82 on average.

Table 8: Tax potential averages across income groups

Region / Income group	TRE	TREz	Total tax (current)	Tax potential (TRE)	Tax potential (TREz)	% change
LIC	0.84	0.84	12.31%	14.52%	14.46%	20.28%
LMIC	0.82	0.82	16.00%	19.79%	19.82%	27.35%
UMIC	0.84	0.84	18.06%	21.51%	21.61%	20.68%
HIC	0.87	0.87	21.80%	24.73%	24.94%	17.65%
East Asia & Pacific	0.81	0.81	16.70%	19.91%	19.98%	30.75%
Europe & Central Asia	0.87	0.86	23.55%	27.15%	27.37%	16.74%
Latin America & Caribbean	0.88	0.88	16.47%	18.63%	18.65%	14.45%
Middle East & North Africa	0.82	0.82	10.67%	13.07%	13.15%	27.42%
North America	0.87	0.86	23.36%	26.90%	27.17%	16.68%
South Asia	0.83	0.83	15.01%	17.89%	17.97%	21.81%
Sub-Saharan Africa	0.82	0.82	14.90%	18.43%	18.40%	23.87%
World average	0.84	0.84	17.65%	20.83%	20.91%	21.53%

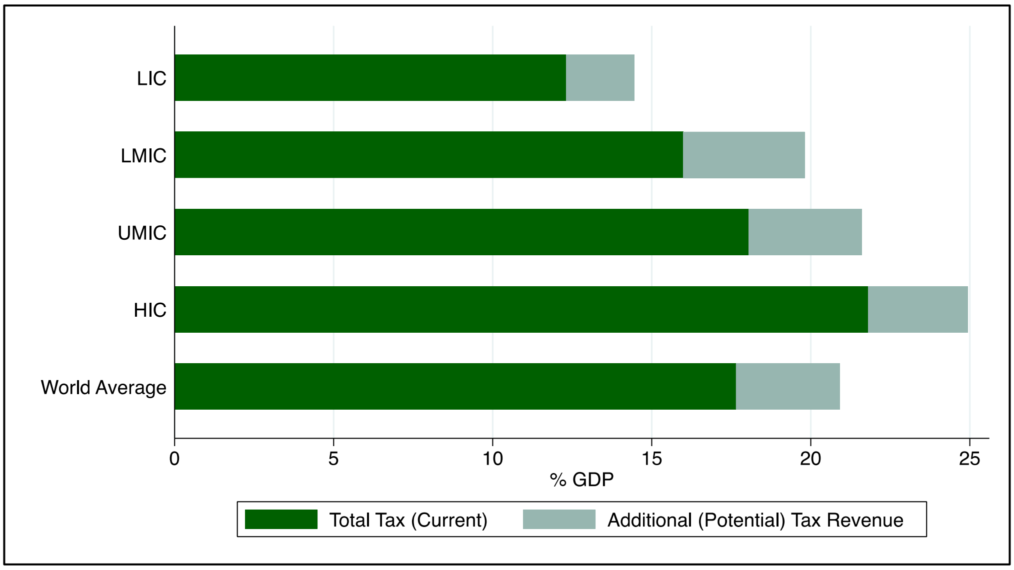
Source: authors' calculations.

Figure 3: Total tax and tax potential, by region



Source: authors' calculations.

Figure 4: Total tax and tax potential, by income group



Source: authors' calculations.

5 Comparison of estimates with existing literature

In this section, we consider how our estimates of tax effort compare with those in the existing literature. Specifically, we Compare with two recent studies that have employed SFA (Langford and Ohlenburg 2016 [LO] and Mawejje and Sebudde 2019 [MS]) and with the estimates of tax effort from USAID's Collecting Taxes Database (CTD), which are estimated according to the approach set out in Fenochietto et al. (2013).[13] Table 9 contrasts our estimates with MS and LO.

[13] See USAID (2021) for further information on the approach taken to calculate tax effort.

Table 9: Comparison of TE estimates with MS (2019) and LO (2016)

Country	MS (2019)		This study	LO (2016)		This study
	Year	Tax effort	Tax effort (TREz)	Year	Tax effort	Tax effort (TREz)
Albania	2015	0.50	0.88			
Algeria	2015	0.35	0.59			
Antigua and Barbuda	2014	0.37				
Argentina	2014	0.51	0.94	2009	0.73	0.89
Armenia	2015	0.51		2008	0.54	
Australia	2015	0.74	0.88	2008	0.8	0.85
Austria	2015	0.71	0.88	2010	0.78	0.86
Azerbaijan	2015	0.38	0.87			
Bahamas	2015	0.34		2010	0.64	
Bahrain	2011	0.03	0.39			
Bangladesh	2015	0.31	0.79	2010	0.42	0.83
Barbados	2009	0.55				
Belarus				2010	0.66	0.81
Belgium	2015	0.66	0.87	2010	0.74	0.85
Belize	2015	0.47				
Benin	2013	0.55	0.88			
Bhutan	2015	0.29	0.90			
Bolivia	2015	0.44				
Bosnia and Herzegovina	2015	0.49	0.87			
Botswana	2014	0.41	0.86			
Brazil	2015	0.62	0.85	2009	0.75	0.85
Brunei Darussalam	2015	0.08				
Bulgaria	2015	0.45	0.83			
Burkina Faso	2013	0.56	0.91	2010	0.65	0.86
Cambodia	2015	0.40	0.92			
Cameroon	2014	0.40	0.89			
Canada	2015	0.76	0.85	2009	0.85	0.86
Central African Republic	2012	0.51	0.85			
Chad	2015	0.28				
Chile	2015	0.35	0.90			
China	2014	0.46	0.79	2010	0.58	0.81
Colombia	2015	0.43		2010	0.54	
Congo, Dem. Rep.	2010	0.38	0.96			
Congo, Rep.	2014	0.31	0.77			
Costa Rica	2015	0.28	0.85	2010	0.47	0.85
Croatia	2014	0.54	0.82	2009	0.67	0.86
Cyprus	2015	0.62	0.89	2008	0.96	0.93
Czech Republic	2015	0.43	0.85	2010	0.53	0.82
Denmark				2010	0.83	0.86
Djibouti	2007	0.59				
Dominica	2015	0.45				
Dominican Rep.	2015	0.32	0.84	2010	0.47	0.79
Ecuador	2015	0.28	0.95			
Egypt, Arab Rep.	2015	0.33	0.75			
El Salvador	2015	0.28	0.90	2009	0.42	0.88
Equatorial Guinea	2015	0.08	0.93			
Eritrea	2002	0.56				
Estonia	2015	0.43	0.88	2010	0.57	0.83

Country	MS (2019)		This study	LO (2016)		This study
	Year	Tax effort	Tax effort (TREz)	Year	Tax effort	Tax effort (TREz)
Eswatini	2015	0.49	0.90			
Ethiopia				2009	0.48	
Fiji	2013	0.48	0.84			
Finland	2015	0.73	0.88	2009	0.82	0.82
France	2015	0.68	0.88	2009	0.76	0.84
Gabon	2015	0.37				
Gambia, The	2014	0.55	0.93			
Georgia	2015	0.50	0.88			
Germany	2015	0.52	0.88	2009	0.66	0.86
Ghana	2012	0.44	0.88	2010	0.52	0.86
Greece	2015	0.53	0.88	2010	0.64	0.84
Grenada	2015	0.41				
Guatemala	2013	0.30	0.89	2009	0.35	0.85
Guinea	2005	0.44	0.83			
Guyana	2014	0.47				
Honduras	2015	0.37	0.88	2009	0.39	0.80
Hong Kong				2010	0.53	0.87
Hungary	2015	0.54	0.85	2010	0.71	0.84
Iceland	2015	0.78	0.81	2009	0.85	
India	2010	0.59	0.79	2010	0.58	0.79
Indonesia	2015	0.31	0.70			
Iran, Islamic Rep.	2015	0.17				
Ireland	2015	0.62	0.72	2010	0.55	0.80
Israel	2015	0.67	0.87			
Italy	2015	0.83	0.87	2009	0.86	0.88
Jamaica				2010	0.85	0.89
Japan	2015	0.51	0.89	2010	0.54	0.79
Jordan	2009	0.37	0.84	2010	0.47	0.79
Kazakhstan	2011	0.34	0.95			
Kenya	2015	0.57	0.84	2010	0.61	0.84
Korea, Rep.	2015	0.44	0.88	2009	0.56	0.86
Kuwait	2015	0.03	0.85			
Kyrgyz Republic	2015	0.58	0.88			
Lao PDR	2010	0.39	0.94			
Latvia	2015	0.43	0.85	2010	0.53	0.82
Lebanon	2015	0.32	0.80	2010	0.64	0.91
Lesotho	2014	0.97	0.84			
Liberia	2015	0.91	0.88			
Lithuania	2015	0.34	0.79	2009	0.47	0.78
Luxembourg	2015	0.82	0.83	2009	0.87	0.89
Macedonia, FYR	2015	0.41	0.78			
Madagascar	2008	0.46	0.92	2009	0.55	0.81
Malawi	2015	0.50	0.92	2010	0.77	0.93
Malaysia	2008	0.27	0.78	2008	0.39	0.78
Maldives	2014	0.27	0.86			
Malta	2015	0.49		2008	0.88	
Mauritania	2012	0.43	0.76			
Mauritius	2015	0.37	0.51			
Mexico	2015	0.25	0.93			
Moldova	2015	0.54	0.85	2010	0.62	0.79

Country	MS (2019)		This study	LO (2016)		This study
	Year	Tax effort	Tax effort (TREz)	Year	Tax effort	Tax effort (TREz)
Mongolia	2007	0.43	0.91			
Morocco	2015	0.57	0.79	2010	0.84	0.85
Mozambique	2015	0.58	0.91	2010	0.78	0.83
Myanmar	2015	0.30	0.82			
Namibia	2014	0.60	0.90	2010	0.83	0.77
Nepal	2015	0.48	0.92			
Netherlands	2015	0.55	0.84	2008	0.7	0.87
New Zealand	2010	0.74	0.82	2009	0.86	0.83
Nicaragua	2015	0.35	0.90	2010	0.65	0.84
Nigeria	2009	0.20				
Norway	2015	0.96	0.80			
Pakistan	2014	0.41	0.81	2010	0.49	0.85
Panama	2015	0.20	0.84	2010	0.42	0.88
Papua New Guinea	2004	0.39	0.92			
Paraguay	2015	0.30	0.86	2010	0.47	0.77
Peru	2015	0.39	0.83			
Philippines	2015	0.37	0.83	2010	0.42	0.79
Poland	2015	0.49	0.81	2008	0.67	0.89
Portugal	2015	0.47	0.88	2008	0.69	0.86
Romania	2015	0.46	0.85			
Russian Federation	2015	0.58	0.80			
Rwanda	2015	0.52	0.80			
Saudi Arabia	2015	0.05	0.71			
Senegal	2014	0.56	0.87	2010	0.75	0.90
Serbia	2015	0.55	0.81	2010	0.73	0.87
Seychelles	2014	0.57	0.83			
Sierra Leone	2015	0.54	0.89	2010	0.64	0.85
Singapore				2009	0.46	0.82
Slovak Republic	2015	0.38	0.85	2010	0.45	0.75
Slovenia	2015	0.52	0.84	2009	0.6	0.85
Solomon Islands	2006	0.36	0.87			
South Africa	2015	0.59	0.88	2008	0.77	0.87
Spain	2015	0.55	0.87			
Sri Lanka	2015	0.34	0.76	2010	0.51	0.74
St. Kitts and Nevis	2014	0.50				
St. Lucia	2012	0.47				
St. Vincent and the Grenadines	2015	0.48				
Sudan	2011	0.28	0.82			
Sweden	2015	0.80	0.88	2009	0.9	0.85
Switzerland	2015	0.60	0.89	2009	0.64	0.88
Tajikistan	2011	0.59	0.85			
Tanzania	2015	0.44	0.82	2010	0.67	0.83
Thailand	2015	0.36	0.82	2010	0.5	0.78
Togo	2015	0.68	0.92	2010	0.54	0.86
Tonga	2015	0.42				
Trinidad and Tobago	2005	0.43				
Tunisia	2014	0.42	0.87	2010	0.64	0.86
Turkey	2015	0.47	0.88	2009	0.62	0.87
Uganda	2015	0.50	0.88	2010	0.51	0.76
Ukraine	2015	0.53	0.87	2010	0.75	0.87

Country	MS (2019)		This study	LO (2016)		This study
	Year	Tax effort	Tax effort (TREz)	Year	Tax effort	Tax effort (TREz)
United Kingdom	2015	0.64	0.88	2010	0.79	0.87
United States	2015	0.59	0.86	2010	0.59	0.78
Uruguay	2015	0.40	0.89	2010	0.72	0.89
Vanuatu	2014	0.40	0.87			
Venezuela, RB	2013	0.34				
Viet Nam	2015	0.43	0.71			
Zambia	2015	0.42	0.70			
Zimbabwe	2005	0.41	0.57	2010	0.9	0.86
World average	*2014*	*0.47*	*0.84*	*2010*	*0.64*	*0.84*

Source: authors' calculations.

We see that, on average—and as expected given the earlier discussion—our estimates of tax effort (of 0.84 on average) are much higher than those previously reported in the literature. The average TE estimate in MS was just 0.47 and in LO 0.64. We also see, again, a much larger variance in the scores reported in these other studies. The estimates in MS range from 0.03 to 0.97, whilst in LO they range from 0.35 to 0.96; the estimates from our own estimations range between 0.74 and 0.94.[14]

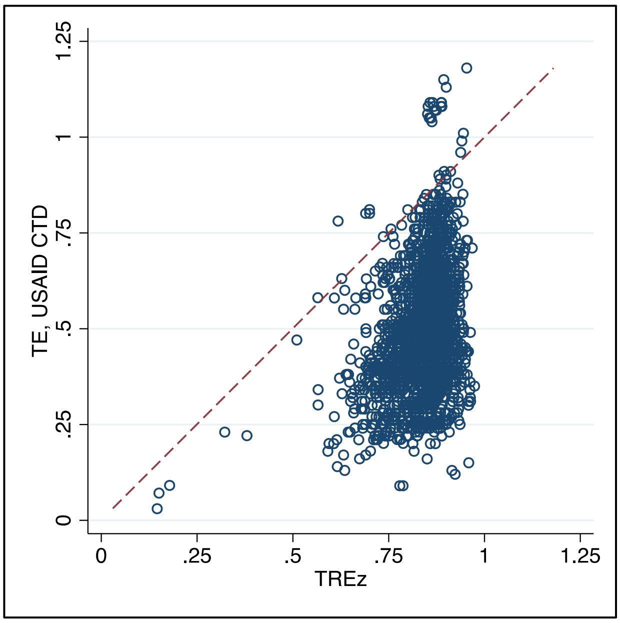
To the best of our knowledge, the USAID CTD 2020/21 is the only other source that publishes a time series of tax effort scores. The database contains some 2,601 observations for some 153 countries or territories. The estimates presented here share some 2,289 common estimates with those in the USAID CTD. Figure 5 shows how these estimates compare.[15]

In just 36 observations (or 1.6 per cent of available comparison points) were the TE scores computed in the present study higher than those presented in the USAID CTD. This likely stems from the estimation method employed by the latter (BC) and the reasons discussed above. Figure 5 also shows a much wider variance of the estimates from USAID, again in line with what we presented above.

[14] One potential limitation of comparing with existing studies is that the outcome variable—namely the tax-to-GDP ratio—is often subject to revision whenever countries rebase their GDP. Thus, to compare (for example) Albania's TE in 2015 (from a study carried out in 2016) with Albania's TE in 2015 (from a study carried out in 2021) may involve comparing two different scores entirely. However, as many of the input variables are also expressed as a percentage of GDP, this is unlikely to severely bias our results and certainly does not explain the differences in the variance of the observed tax effort scores.

[15] We do not show these in a table form due to the sheer volume of observations.

Figure 5: Comparison of TE estimates from USAID and the TREz model (this study)



Source: authors' calculations.

6 Discussion and conclusion: the political economy of tax effort estimates

This study has sought to revisit the question of tax effort across countries. Employing more appropriate estimation techniques, along with better, more complete data, has led to dramatically different estimates of tax effort from those presented in some recent literature. By employing the True Random Effects approach to estimate tax effort via Stochastic Frontier Analysis, we find our estimates to be much less influenced by outlying input data, and the resulting tax effort scores are thus much more tightly fit and on average, a lot higher.

The study does, however, have limitations with respect to our output variable, which is total tax collections excluding social security contributions. First, a number of countries—particularly those dependent on revenues from extractive industries—collect next to nothing in tax revenues, but depend on royalties in the form of non-tax payments. In order to account for such revenues, one would need to estimate *revenue effort*, employing total government revenue as the outcome variable. This, however, comes with further limitations: some non-tax revenues are less under the control of policymakers than tax revenues (e.g. non-tax revenues from extraction of oil might be more influenced by the world price of oil than the rates charged to extractive companies). Second, the choice to fund social security via earmarked contributions, or through the tax system, is country-specific. Thus, by defining our output variable as *excluding* social security contributions, we may see higher TE scores for countries that fund social security through taxes versus those that charge high rates of social security contributions.

Finally, a word on the interpretation—and use—of tax effort scores more generally. We believe that the estimates presented herein have gone some way to offering clarity on the exercise of estimating tax effort. It is crucial to interpret these figures whilst also eyeballing the input data and using one's intuition. Previous TE scores have been heavily influenced by outlying observations. For those involved in advocacy, technical assistance, or policy advisory roles, we hope that this study is informative and provides more clarity on the question of *how much tax revenue can country X collect, given its underlying characteristics*? However, just as tax revenues may fluctuate in the future, so too do the underlying characteristics that produce these estimates. The models and results are not forward-looking. They do not take into account, for example, the fact that country X may have just found oil reserves that will transform its economy and tax potential (for better or worse). TE scores are a useful, high-level benchmark for the level of tax that a country *might* be able to collect *today*. They should form part of a toolkit for assessing tax performance in a given country, but should not be solely relied upon. They say little of *why* countries are performing better or worse. Other diagnostics, such as the IMF's Tax Administration Diagnostic Assessment Tool (TADAT), country-specific studies, or Medium Term Revenue Strategies should give a more rounded view of the underlying determinants of tax revenue performance—and future areas for improvement—in a given country.

References

Aigner, D., C.K. Lovell, and P. Schmidt (1977). 'Formulation and estimation of stochastic frontier production function models'. *Journal of Econometrics*, 6(1): 21–37. https://doi.org/10.1016/0304-4076(77)90052-5

Bahl, R.W. (1971). 'A regression approach to tax effort and tax ratio analysis'. *IMF Staff Papers*, 18(3): 570–612. https://doi.org/10.2307/3866315

Battese, G.E., and T.J. Coelli (1995). 'A model for technical inefficiency effects in a stochastic frontier production function for panel data'. *Empirical Economics*, 20(2): 325–32. https://doi.org/10.1007/BF01205442

Beck, T., G. Clarke, A. Groff, P. Keefer, and P. Walsh (2001). 'New tools in comparative political economy: the database of political institutions'. *The World Bank Economic Review*, 15(1): 165–76. https://doi.org/10.1093/wber/15.1.165

Bird, R.M., J. Martinez-Vazquez, and B. Torgler (2006). 'Societal institutions and tax effort in developing countries'. CREMA Working Paper 2004-21. Zurich: Crema Vermögensverwaltung & Research. https://doi.org/10.2139/ssrn.662081

Bornhorst, F., S. Gupta, and J. Thornton (2009). 'Natural resource endowments and the domestic revenue effort'. *European Journal of Political Economy*, 25(4): 439–46. https://doi.org/10.1016/j.ejpoleco.2009.01.003

Chachu, D. (2021). 'Good institutions and tax revenue outcomes in resource-rich countries: When "good" is not enough'. WIDER Working Paper 2021/75. Helsinki: UNU-WIDER. https://doi.org/10.35188/UNU-WIDER/2021/013-9

Coppedge, M., et al. (2021). 'V-Dem Dataset v11.1'. Varieties of Democracy Project. https://doi.org/10.23696/vdemds21

Crivelli, E., and S. Gupta (2014). 'Resource blessing, revenue curse? Domestic revenue effort in resource-rich countries'. *European Journal of Political Economy*, 35: 88–101. https://doi.org/10.1016/j.ejpoleco.2014.04.001

Cyan, M., J. Martinez-Vazquez, and V. Vulovic (2013). 'Measuring tax effort: does the estimation approach matter and should effort be linked to expenditure goals?'. ICEPP Working Paper 39. International Center for Public Policy. Available at: https://scholarworks.gsu.edu/icepp/39 (accessed 18 November 2021).

Faust, A.K., and A. Baranzini (2014). 'The economic performance of Swiss drinking water utilities'. *Journal of Productivity Analysis*, 41(3): 383–97. https://doi.org/10.1007/s11123-013-0344-0

Fenochietto, R., C. Pessino, M. Keen, and V.J. Perry (2013). 'Understanding countries' tax effort'. IMF Working Paper 2013(244). Washington, DC: International Monetary Fund. https://doi.org/10.5089/9781484301272.001

Ghura, M.D. (1998). 'Tax revenue in sub-Saharan Africa: effects of economic policies and corruption'. IMF Working Paper 1998(135). Washington, DC: International Monetary Fund. https://doi.org/10.5089/9781451855685.001

Greene, W. (2005). 'Reconsidering heterogeneity in panel data estimators of the stochastic frontier model'. *Journal of Econometrics*, 126: 269–303. https://doi.org/10.1016/j.jeconom.2004.05.003

Hinrichs, H. (1965). 'Determinants of government revenue share among less developed countries'. *Economic Journal*, 75(299): 546–56. https://doi.org/10.2307/2228952

IMF (1971). 'Trends in taxation in developing countries'. *IMF Staff Papers*, 18(2). Washington, DC: International Monetary Fund.

IMF (1975). 'Tax ratios and tax effort in developing countries, 1969–71'. *IMF Staff Papers*, 22(1). Washington, DC: International Monetary Fund.

IMF (1979). 'International comparisons of taxation for selected developing countries, 1972–76'. *IMF Staff Papers*, 26(1). Washington, DC: International Monetary Fund. https://doi.org/10.2307/3866567

Kumbhakar, S.C., H. Wang, and A.P. Horncastle (2015). 'A practitioner's guide to stochastic frontier analysis using Stata'. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139342070

Langford, B., and T. Ohlenburg (2016). 'Tax revenue potential and effort: an empirical investigation'. IGC Working Paper S-43202-UGA-1. London: International Growth Centre. Available at: https://www.theigc.org/wp-content/uploads/2016/01/Langford-Ohlenburg-2016-Working-paper.pdf (accessed 18 November 2021).

Leuthold, J.H. (1991). 'Tax shares in developing economies a panel study'. *Journal of development Economics*, 35(1): 173–85. https://doi.org/10.1016/0304-3878(91)90072-4

Lotz, J.R., and E.R. Morss (1970). 'A theory of tax level determinants for developing countries'. *Economic Development and Cultural Change*, 18(3): 328–41. https://doi.org/10.1086/450436

Mawejje, J., and R. Sebudde (2019). 'Tax revenue potential and effort: worldwide estimates using a new dataset'. *Economic Analysis and Policy*, 63: 119–29. https://doi.org/10.1016/j.eap.2019.05.005

McNabb, K. (2017). 'Toward closer cohesion of international tax statistics: the ICTD/UNU-WIDER GRD 2017'. WIDER Working Paper 184/2017. Helsinki: UNU-WIDER. https://doi.org/10.35188/UNU-WIDER/2017/410-0

McNabb, K., A. Oppel, D. Chachu (2021). 'Government Revenue Dataset (2021): source selection'. WIDER Technical Note 2021/10. Helsinki: UNU-WIDER. https://doi.org/10.35188/UNU-WIDER/WTN/2021-10

Oppel, A., K. McNabb, and D. Chachu (2021). 'Government revenue dataset (2021): variable description'. WIDER Technical Note 2021/11. Helsinki: UNU-WIDER. https://doi.org/10.35188/UNU-WIDER/WTN/2021-11

Oshima, H.T. (1957). 'Share of government in gross national product for various countries'. *The American Economic Review*, 47(3): 381–90.

Pessino, C., and R. Fenochietto (2010). 'Determining countries' tax effort'. *Hacienda Pública Española/Revista de Economía Pública*, 195(4): 65–87.

Pitt, M.M., and L.F. Lee (1981). 'The measurement and sources of technical inefficiency in the Indonesian weaving industry'. *Journal of Development Economics*, 9(1): 43–64.

Prichard, W., A. Cobham, and A. Goodall (2014). 'The ICTD government revenue dataset'. ICTD Working Paper 19. Brighton, UK: International Centre for Tax and Development. https://doi.org/10.2139/ssrn.2496442

Sen Gupta, A. (2007). 'Determinants of tax revenue efforts in developing countries'. IMF Working Paper WP/07/184. Washington, DC: International Monetary Fund. https://doi.org/10.5089/9781451867480.001

Stotsky, J.G., and A. WoldeMariam (1997). 'Tax effort in sub-Saharan Africa'. IMF Working Paper WP/97/107. Washington, DC: International Monetary Fund. https://doi.org/10.5089/9781451852943.001

Tanzi, V. (1992). 'Structural factors and tax revenue in developing countries: a decade of evidence'. In I. Goldin and A.L. Winters (eds), *Open economies: structural adjustment and agriculture*. New York: Cambridge University Press. https://doi.org/10.1017/CBO9780511628559.023

Thorn, R.S. (1967). 'The evolution of public finances during economic development'. *The Manchester School*, 35(1): 19–53. https://doi.org/10.1111/j.1467-9957.1967.tb01204.x

UNU-WIDER (2021). 'UNU-WIDER government revenue dataset'. Version 2021. Helsinki: UNU-WIDER. https://doi.org/10.35188/UNU-WIDER/GRD-2021

USAID (2021). 'Collecting taxes database 2020–2021. Technical note'. Available at: https://s3.amazonaws.com/files.devdata.devtechlab.com/Collecting+Taxes+Database/Collecting+Taxes+Database+Technical+Note+2020-2021.pdf (accessed 18 November 2021).

Williamson, J.G. (1961). 'Public expenditure and revenue: an international comparison'. *The Manchester School*, 29(1): 43–56. https://doi.org/10.1111/j.1467-9957.1961.tb01187.x