Institutional environment and tax performance: empirical evidence from developing economies

PRIANTO BUDI SAPTONO, Dr.*
GUSTOFAN MAHMUD, M.Sc.*

Article**
JEL: H11, H20, H21
https://doi.org/10.3326/pse.46.2.2

* The authors would like to thank to two anonymous referees whose comments have contributed to the final version of the paper.

** Received: October 22, 2021
Accepted: February 4, 2022

Prianto Budi SAPTONO
Faculty of Administrative Science, Universitas Indonesia, Prajudi Atmosudirdjo Building Level 2, FIA UI Depok, Indonesia 16424
e-mail: prianto.saptono@ui.ac.id
ORCiD: 0000-0003-0651-5390

Gustofan MAHMUD
Pratama Kreston Tax Research Institute, Antam Office Park Tower B Level 8, Jakarta, Indonesia 12530
Accounting Study Program, Sekolah Tinggi Ilmu Ekonomi (STIE) Swadaya, Jatiwaringin Raya No. 36, Jakarta, Indonesia 13620
e-mail: gustofan@pratamaindomitra.co.id
ORCiD: 0000-0001-8580-8049

This is an Open Access article distributed under a Creative Commons Attribution-NonCommercial 4.0 International License which permits non commercial use and redistribution, as long as you give appropriate credit, provide a link to the license, and indicate if changes were made.

Abstract

Unlike its predecessor studies, this paper investigates the contemporaneous and lagged effects of institutional variables on tax performance, using unbalanced panel data from 79 developing countries for the 2002-2019 period. The instrumental variable (IV) and system-generalized method of moments (SGMM) estimation models were employed in this study to address potential endogeneity and specification biases in the estimation model. Generally, this study found that countries with low corruption levels and good governance quality could produce more tax revenues. Moreover, the lagged effects of institutional variables, which are much more substantial than their contemporary effects, provide meaningful insight. Efforts directed at fighting corruption and improving the quality of governance must be carried out as early as possible to obtain optimal tax revenues in the future. These efforts can be taken by streamlining tax administration, so that opportunities for bribery and corruption can be reduced.

Keywords: corruption, governance, tax revenue

1 INTRODUCTION

The low ability of governments to generate sufficient revenues to finance vital public provision expenditures is a typical fiscal issue in most developing countries. In the last decade, developing countries have typically collected taxes amounting to only about 13 percent of GDP (World Bank, 2021). It is still below the tipping point of 15 percent of GDP as suggested by the World Bank (Junquera-Varela and Haven, 2018). Apart from economic, institutional factors, such as poor governance and high levels of corruption or perceptions of unfair tax regimes also play an essential role in this problem (Besley and Persson, 2014; Moore and Prichard, 2020). Weak governance delivers complex and inefficient tax systems, which in turn increases tax evasion and the cost of tax correction (Everest-Phillips and Sandall, 2009). The complexity of the taxation system is the cause of rampant corruption in tax administration, especially in the tax collection process (Fjeldstad, 2006; Rahman, 2009). Such situations will undermine the tax structure and revenue collecting capacity of a country, generating a significant loss in availability of funds for the provision of public services (Transparency International, 2014). In the long run, corruption can also ruin taxpayers' morality and erode public trust in government institutions (Nawaz, 2010). Thus, improving the taxation system and building tax administration capacity are two essential components that reinforce each other in generating revenue (Brondolo et al., 2008).

This study investigates the extent to which corruption and governance affect tax performance in developing countries. Several empirical studies have contributed to this topic. However, most of them neglect to address the endogeneity problem of institutional variables. For example, Syadullah (2015) estimated the impact of governance quality and corruption level on tax revenue in 7 Southeast Asian countries during 2003-2012. The authors found that control of corruption has significant adverse effects on the tax ratio, while the rule of law and quality of regulatory variables positively impact the tax ratio. Epaphra and Massawe

(2017), by using a data set for 30 African countries from 1996 to 2016, showed that corruption and institutional quality have more influence on indirect tax revenues than on other types of taxes (direct taxes and trade taxes). Arif and Rawat (2018), based on data from 10 Emerging and Growth-Leading Economies (EAGLE) with a period between 2001 and 2015, confirmed the vital role of the institutional environment in determining the level of tax revenue. Hassan et al. (2021) used data from Pakistan covering the period 1976-2019. They concluded that good governance is an essential resource to increase tax revenue both in the short and in the long run.

Some authors on the same topic may acknowledge the issue of endogeneity in the model, like Ajaz and Ahmad (2010) in developing countries and Imam and Jacobs (2014) in the Middle East. They argued that persistent tax revenues over time are assumed to be endogenous to its lag, so it can potentially create specification bias in the model. They used the dynamic panel SGMM estimator to solve this problem. They found that corruption is the leading cause of low tax revenues in developing countries and the Middle East. The problem with their study is that they fail to answer these questions: for example, do countries with higher tax performance have better governance indicators? Moreover, the SGMM is quite unsatisfactory for dealing with endogeneity when the variables show persistence over time, and the IV approach is much preferred.

Hwang (2002) and Bird, Martinez-Vazquez and Torgler (2008) are probably the only studies on this topic that address institutional variable endogeneity issues. Hwang (2002) found that an increasing in cases of corruption will undermine government revenues. The author used the share of the population professing Protestantism and whether it was a former British colony, and the index of ethnolinguistic fractionalization as instruments for corruption. Meanwhile, Bird, Martinez-Vazquez and Torgler (2008) used the legal origins (English) of a country and fractionalization as instruments of institutional variables and found that enhancing voice or accountability and reducing corruption is an important prerequisite for a more satisfactory level of tax revenue in developing and developed countries. However, their study still contains some limitations. First, they only include total tax revenue as a measure of tax performance without breaking it down into the types of taxation. Dividing tax revenue into its types will sharpen the "knife of analysis." It lets us know which taxes are most affected by corruption and governance (see, e.g., Imam and Jacobs, 2014; Epaphra and Massawe, 2017). Second, the authors ignored the persistent nature of tax revenue, which can create a specification bias in the model (see, e.g., Ajaz and Ahmad, 2010; Imam and Jacobs, 2014). Third, they only control economic factors and ignore other potential control variables, such as population size, public trust, and the shadow economy.

Therefore, the main contribution of this study is expected to be discussion of the intersection of issues that have not been addressed by any previous related studies.

To achieve this goal, we will use an unbalanced panel data set for 79 developing countries covering 2002 to 2019. To measure tax performance, we use aggregate revenues and three general categories of taxation, including direct taxes, indirect taxes, and trade taxes. In addition, we also control for macroeconomic factors (GDP per capita, industrial sector, inflation, and trade openness), demographic factors (population size), and political factors (trust in government), and the shadow economy. To solve the endogeneity problem, we will use two democracy indicators (civil liberties and political rights) as instruments of the troublesome variables. We will analyze the problem using a two-stage least square (2-SLS) estimation. In addition, since tax revenue is persistent over time, it is reasonable to introduce a lagged level of tax revenue as additional independent variable in the model. We will analyze this part using the SGMM estimator to calibrate the auto-correlation and heteroscedasticity in the model. Finally, another contribution of this study is that we also introduce the lagged level of institutional variables in the SGMM estimation. Estimating such a dynamic model allows us to observe both the contemporary and the lagged effects of the institutional environment on tax performance.

The rest of this paper is organized as follows. Section 2 is devoted to discussing research data and variables. Section 3 designs the econometric framework used to study the effects of the institutional environment on tax performance. Section 4 discusses the empirical results of this study. Section 5 concludes and raises some policy implications.

2 DATA

As mentioned in the introduction section, we use panel data from 79 developing countries worldwide covering 2002-2019 to achieve the research objectives. To ensure the robustness of the results, we utilize secondary data obtained from primary sources for each variable studied, such as World Development Indicators (WDI), World Governance Indicators (WGI), Freedom House, World Economic Forum Global Competitiveness Index, and an empirical study conducted by Elgin et al. (2021) aimed at providing detailed information on the construction and sources for the variables included in the database and shows two applications of the database: the stylized facts of informal economic activity around the world and the cyclical features of the informal economy. Their proposed measurement formula has been used by the World Bank to estimate global informal activity over the period 1990-2018.

To measure a country's tax performance, we use aggregated tax revenues and three other categories of tax revenues, namely direct taxes, indirect taxes, and trade taxes. All these tax revenue indicators are expressed as a percentage of GDP and are obtained from WDI.

TABLE 1
Variable description and summary statistics

Variable	Description	N	Mean	Std. dev	Min	Max	Source
TR_{it}	Total taxes (% of GDP)	977	15.196	5.624	0.915	39.988	A
DTR_{it}	Direct taxes (% of GDP)	977	5.183	3.099	0.250	20.797	A
$IDTR_{it}$	Indirect taxes (% of GDP)	977	7.676	3.732	0.099	18.622	A
TTR_{it}	Trade taxes (% of GDP)	977	1.942	2.129	0.00003	13.436	A
VA_{it}	Voice and accountability (-2.5 to 2.5 scale)	977	-0.320	0.663	-1.971	1.152	B
PS_{it}	Political stability and no violence/terrorism (-2.5 to 2.5 scale)	977	-0.409	0.734	-2.699	1.263	B
GE_{it}	Government effectiveness (-2.5 to 2.5 scale)	977	-0.338	0.575	-1.848	1.267	B
RQ_{it}	Regulatory quality (-2.5 to 2.5 scale)	977	-0.244	0.564	-2.071	1.127	B
RL_{it}	Rule of law (-2.5 to 2.5 scale)	977	-0.429	0.530	-1.823	1.077	B
GVN_{it}	Governance index	977	0.014	0.996	-2.605	2.842	C
CC_{it}	Control of corruption (-2.5 to 2.5 scale)	977	-0.471	0.543	-1.826	1.647	B
$ICAP_{it}$	Log of GDP per capita (current US$)	977	4,132.364	3,470.211	237.757	22,942.61	A
IND_{it}	Industry, value added (% of GDP)	977	29.293	10.593	9.476	84.349	A
INF_{it}	Inflation, consumer price index (annual %)	977	111.397	3,283.225	-3.749	102,629.8	A
TOP_{it}	Exports plus imports (% of GDP)	977	80.000	32.157	16.141	210.374	A
SHD_{it}	Shadow economy output (% of GDP)	883	34.842	10.384	8.552	66.137	D
POP_{it}	Population size (Persons in thousands)	977	70,300	222,000	108.3	1,400,000	A
TIP_{it}	Trust in politicians (1 = extremely low to 7 = extremely high)	576	2.602	0.806	1	6	E
CVL_{it}	Civil liberty (1 = maximum freedom to 7 = absence of freedom)	968	3.761	1.347	1	7	F
PR_{it}	Political rights (1 = maximum freedom to 7 = absence of freedom)	968	3.863	1.769	1	7	F

Explanation: A = WDI. B = WGI. C = Calculated by the authors using PCA based on five governance indicators / WGI (VA_{it}, PS_{it}, GR_{it}, RQ_{it}, RL_{it}). D = Elgin et al. (2021), E = World Economic Forum Global Competitiveness Index. F = Freedom House.

Source: Own calculation.

Following Arif and Rawat (2018), we construct a governance index using principal component analysis (PCA) of governance indicators introduced by World Governance Indicators (WGI): voice and accountability, political stability and absence of violence/terrorism, government effectiveness, regulatory quality, and the rule of law. WGI also covers corruption control as an indicator of a country's governance. However, we did not include this one indicator in the formation of the governance index. Instead, we use it separately as one measure of corruption. It is done to avoid collinearity between the two key explanatory variables (governance and corruption). Unlike Transparency International's annual Corruption Perceptions Index, this index incorporates different aspects of corruption, from the frequency with which companies make "additional payments to get things done," to the effect of corruption on the business environment, to the corruption measure of "major corruption" in the political arena (Olken and Pande, 2012).

The WGI itself is constructed using the Goldberger (1972) and Efron and Morris (1971, 1972) unobserved components model (UCM) and is expressed in a scale range from -2.5 to 2.5, with higher scores referring to a higher quality of governance and lower level of corruption (Kaufmann, Kraay and Mastruzzi, 2010). Meanwhile, PCA used in this study is a technique used to extract meaningful information from several correlated quantitative variables to represent them as a new set of orthogonal variables called principal components (Hotelling, 1933). PCA also represents the pattern of similarity of observations and variables by displaying them as points on the map (Jackson, 1991; Saporta and Niang, 2009; Jollife and Cadima, 2016). With reference to the literature review presented in section 1, the main hypothesis of this study is formulated as follows:

H1: Corruption has an adverse effect on tax revenues.

H2: Governance has a positive effect on tax revenues.

By including a battery of control variables in the analysis, we intend to consider factors beyond the central question in this study, but which are essential to consider because they can influence the size of tax revenues through mechanisms other than the institutional environment. The level of national development is one of the most potentially confounding determinants. GDP per capita is an indicator widely used in the literature on tax revenues to control different national development levels across countries (Le, Moreno-Dodson and Rojchaichaninthorn, 2008). We argue that per capita income is very close to the "ability to pay" in society. Thus, we suspect that per capita income has a significant positive effect on tax revenue. The inclusion of the economy's composition also seems essential as another determinant in the tax revenue regression. An increasingly industrialized economy structure tends to be easier to tax because companies in this sector have better annual financial reports than other sectors such as agriculture, allowing tax officials to audit it more easily (Gupta, 2007). We suspect that inflation will erode people's purchasing power and thereby hamper the productivity of government revenues (Immervoll, 2005). Therefore, we include inflation in the set of control variables. We also control for the degree of openness of the economy, which affects revenues, particularly international trade taxes (Gaalya, 2015). Data for these variables are obtained from WDI. In addition, since our investigation focused on the case of developing countries faced with harsh shadow economic phenomena, we also include the shadow economy as a control variable. Referring to the results disclosed by Schneider (2005), we suspect that any expansion of the shadow economy will decrease the level of tax revenue. We argue that the shadow economy includes unreported income from legal activities and several illegal activities that will not be counted in GDP (i.e., drug dealing, smuggling, fraud, etc.) (Lippert and Walker, 1997), thereby reducing the taxable revenue bases. To control for demographical factors, we use population size. We consider the population (human resources) as the main producer of the tax revenue bases, so it is estimated to positively affect revenue. Lastly, the issue of trust also seems to influence tax revenue. Accordingly, we also include trust in politicians to control for political factors. We argue that a high level of public trust is the principal capital for the growth of voluntary tax compliance, which leads to an increase in tax revenue. Table 1 delivers a detailed description of all the variables described above.

3 ECONOMETRIC FRAMEWORKS

The first set of models used to estimate the causal relationship between the institutional environment and tax performance is a simple ordinary least squares (OLS) model similar to the basic model used in Epaphra and Massawe (2017):

$$T^{`}_{it} = \beta_1 X^{`}_{it} + \gamma Z^{`}_{it} + d_t + a_i + \varepsilon_{it} \tag{1}$$

Where i and t are the indexes for country and year, respectively. $T^{`}_{it}$ is a $1 \times g_1$ vector of outcome variables (i.e., total tax revenue, direct tax revenue, indirect tax revenue, and trade tax revenue). $X^{`}_{it}$ is a $1 \times k_2$ vector of primary explanatories (i.e., governance and corruption). $Z^{`}_{it}$ is a $1 \times k_3$ vector of observations on the control variables included as covariates to alleviate omitted variable bias (i.e., income per capita, the economic composition, inflation, and trade liberalization). ϕ, β_1, and y are y, $k_1 \times 1$, $k_2 \times 1$ and $k_2 \times 1$ vectors of unknown coefficients. d_t and a_i are the country level and year level effects, respectively. Lastly, ε_{it} is the usual error term and assumed to be non-independently and identically distributed for each i over all t.

However, it should be noted that our OLS estimate as shown in equation (1) is likely to be biased because institutional environment variables are expected to be endogenous to a country's tax performance. For example, countries with higher tax collections may be associated with better governance indicators. By a similar logic, higher tax collections may exacerbate the level of corruption in a country. Equation (1) does not address these concerns, thus potentially biasing our results.

To cope with these issues, we need to incorporate IV in the model with the following characteristics: correlated with $X^{`}_{it}$ (strong instruments) and orthogonal to $T^{`}_{it}$ (valid instruments). We use two indicators of democracy, namely political rights and civil liberties as instruments for corruption and governance. We can confidently say that the instruments are strong because the existing literature concludes that political rights and/or civil liberties can affect the level of corruption (Abu and Staniewski, 2019) and the quality of governance (Benkovskis and Fadejeva, 2014). In addition, we also argue that these two instruments are valid because it seems unlikely that political rights and civil liberties directly affect tax performance without undergoing improvements/deterioration in the institutional environment. For example, when basic individual and political rights are fully granted to citizens it may contribute to political stability and improved institutional quality, which in turn increases the government's efforts to mobilize tax revenues. Data on these instruments are drawn from Freedom House. The index of both instruments ranges from 1 (max freedom) to 7 (total absence of freedom). Table 1 reports summary statistics of these instruments.

The second set of models explores the dynamic relationship between the institutional environment and tax performance. Given that the level of tax collection tends to be highly persistent over time, it is reasonable to assume that current tax revenues depend on past tax levels and institutional quality. Therefore, we estimate the variance of tax performance with a dynamic panel model by introducing the lagged levels of tax revenue and institutional quality on the right side of the equation. The following equation captures that dynamic:

$$T^{`}_{it} = \phi T^{`}_{it-1} + \beta_1 X^{`}_{it} + \beta_2 X^{`}_{it-1} + \gamma Z^{`}_{it} + d_t + a_i + \varepsilon_{it} \tag{2}$$

The main difference shown by equations (1) and (2) is that the latter captures both the contemporary effect and the lagged effect of the institutional environment variables on tax performance. Estimating this dynamic model allows us to perform indirect tests of different causal mechanisms. Again, it makes sense because the economy recognizes a natural phenomenon known as time lag. Policies designed to improve the quality of current institutions may only have an impact in the future, particularly in relation to improving tax performance. Existing studies seem to fail to capture this mechanism.

However, it should be noted that, equation (2) can still produce biased and inconsistent parameters because heteroscedasticity in residuals and autocorrelation within panels (countries) always appear in data involving many units across places. In addition, the bias parameter is also attributed to the potential endogeneity of the lagged tax revenues ($T^{`}_{it-1}$) in a dynamic panel model – when this variable correlates with the random error term of the equation. The potential for this problem to arise is very high, since we cannot fully capture the determinants of tax revenues.

To address those potential issues, we explore model (2) using the dynamic SGMM proposed by Arellano and Bover (1995) and Blundell and Bond (1998). In contrast to the difference-GMM (DGMM) proposed by Arellano and Bond (1991), the SGMM estimation corrects the endogeneity problem by introducing more instruments, thereby dramatically increasing the estimator's efficiency. Generally, there are two types of instruments used in the SGMM, namely first differences of the endogenous variable and the lagged levels of the equation. SGMM then transforms these instruments to make them orthogonal to the fixed effects. Another advantage of using this model is that it allows us to minimize data loss better than the DGMM. This is because instead of subtracting the previous observation from a contemporaneous one, SGMM subtracts the average of all future available observations of a variable. It implies that no matter how many gaps we have in our unbalanced panel data set, such a data is computable for all observations except the last for each individual (country). In addition, Monte Carlo simulation also suggests that when the time span is short and the dependent variable is persistent, there are gains in precision and the small sample bias is reduced when the SGMM is applied (Blundell and Bond, 1998).

Both one-step and two-step SGMM estimators will be used in this study to ensure robust estimation results. In the presence of heteroscedasticity and serial correlation, a two-step GMM estimator should be used by exploiting a weighting matrix using residuals from the first step (Roodman, 2009). However, in finite samples, such standard errors tend to be downward biased. The conventional approach by practitioners in such circumstances is to use what is known as the Windmeijer (2005) adjustment to correct for such a small sample bias.

To test the validity of the instrument, we will run the Hansen test and the difference-in-Hansen test. The null hypothesis for the first test states that all instruments used are exogenous (orthogonal to dependent variables). The null hypothesis for the second test confirms the exogeneity of external instruments (consists of key explanatories and control variables) in the SGMM estimation. However, Roodman (2009) stated that the *p*-value of those tests might be bloated, primarily when the instruments used to overcome the endogeneity problems outnumber the country panels. Due to the relatively large number of periods under study $t = 18$, the SGMM model we build is likely to face the instrument proliferation problem, especially when all lags are exploited as instruments. Therefore, the GMM-style instrument lag needs to be restricted from two to four to prevent overuse of the instruments in the model, as Roodman (2009) suggested. We treat $T^{`}_{it-1}$, $X^{`}_{it}$ and $X^{`}_{it-1}$ as endogenous and generate the GMM-style instruments for the corresponding endogenous variables.

4 RESULTS AND DISCUSSION

4.1 CORRELATION ANALYSIS

According to table 2, all correlation coefficients show the direction of the relationship as hypothesized, except the relationship between per capita income (transformed into logarithm) and trade tax which appears with a negative sign. The likely plausible reason for this result is that most trade taxes are levied on imports rather than exports, while import is an element of leakage for a GDP. In addition, another interesting finding here is the positive relationship between economic openness and trade taxes. It emphasizes that although economic openness indicates a reduction in tariffs, it is compensated by an increase in the trade volume, which encourages the flow of trade tax revenues.

PUBLIC SECTOR ECONOMICS
46 (2) 207-237 (2022)

Table 2
Pearson correlation matrix

	TR_{it}	DTR_{it}	$IDTR_{it}$	TTR_{it}	GVN_{it}	CC_{it}
TR_{it}	1					
DTR_{it}	0.755	1				
$IDTR_{it}$	0.650	0.168	1			
TTR_{it}	0.494	0.295	-0.096	1		
GVN_{it}	0.352	0.354	0.191	0.123	1	
CC_{it}	0.435	0.422	0.212	0.220	0.864	1
LOG_ICAP_{it}	0.162	0.195	0.142	-0.086	0.409	0.297
IND_{it}	0.060	0.187	0.263	0.005	0.032	-0.039
INF_{it}	-0.018	-0.012	-0.095	0.147	-0.100	-0.101
TOP_{it}	0.498	0.333	0.424	-0.183	0.250	0.224
SHD_{it}	-0.062	-0.084	-0.014	-0.067	-0.445	-0.331
LOG_POP_{it}	-0.471	-0.177	-0.448	-0.247	-0.217	-0.303
TIP_{it}	0.142	0.258	0.100	0.188	0.463	0.461
CVL_{it}	-0.331	-0.249	-0.283	-0.068	-0.451	-0.444
PR_{it}	-0.262	-0.230	-0.236	-0.005	-0.387	-0.410

Source: Own calculation.

Table 2 also reports that institutional variables (governance and corruption) appear to be correlated. The governance index has a strong positive relationship with control of corruption (r = 0.86). It implies that there are multicollinearity problems among institutional variables. Hence, these three variables need to be applied separately in the regression analysis to avoid a biasness in the estimation model.

4.2 REGRESSION ANALYSIS OF TOTAL TAX REVENUE

Table 3 (models 1-6) shows the main IV estimation results for governance index and control of corruption with aggregate tax revenues. The two variables representing the institutional factors are tested separately in all regressions to avoid multicollinearity bias. In addition, the shadow economy and public trust in politician variables are also analyzed separately from other control variables because there are fewer observations of them due to the paucity of data. Overall, our estimation results support the central hypothesis of this study. Countries with better governance quality and lower levels of corruption tend to have higher aggregate tax performance. The result is consistent with Bird, Martinez-Vazquez and Torgler (2008), and Ajaz and Ahmad (2010), who explained that efficient governance and an efficient political system are necessary for a profitable tax system in developing countries. Government effort to collect tax revenues will also be disrupted if the administration and governance are poor (Prichard, 2010). The results in this study are also consistent with Alm, Martinez-Vazquez and McClellan (2015), who explained that the presence of tax inspectors who asked for bribes resulted in a higher level of tax evasion. Hunady and Orviska (2015) emphasized that increasing levels of corruption among officials can harm tax administration and tax audits and undermine the state's credibility.

Furthermore, table 3 (models 7-12) also presents the results from the first-stage regression of IV estimations. We use civil liberties and political rights to instrument governance and corruption across all models. The coefficient signs of the excluded instruments (CVL_{it} and PR_{it}) are as predicted and mostly significant at the 1% level. This implies that civil liberties and political rights are particularly strong predictors of the institutional environment. Good governance and effective prevention of corruption will be easier to achieve in countries that uphold the freedom of their citizens to voice and participate in politics. Moreover, our under- and weak-identification tests also reveal congruent results. The *p*-values and *F*-statistics of the respective tests reject the null hypothesis, which indicates that both excluded instruments are relevant and satisfactory (well-performed) for determining institutional variables. The Hansen J test failed to reject the null hypothesis of the joint exogeneity of instruments for all models (*p*-values > 0.1). This indicates that our instruments are valid, meaning orthogonal to the endogenous regressors. Thus, two eligible instrument requirements have been met.

In addition, table 3 also reports the estimations results of a series of control variables on aggregate tax revenue. All the coefficient signs are in accordance with the hypothesis except for the population, which may be because an increase in population is not always accompanied by an increase in compliance, especially in developing countries. Per capita GDP was found to have a positive and significant relationship with total tax revenues. This result is in line with Karagöz (2013) and Ayenew (2016), who noted that economic development improves public services and taxpayers' ability to pay, which increases the efficiency of tax authorities in collecting taxing that intended to finance increased demand for spending. The industrial sector's contribution to GDP was also found to have a positive and significant impact on total tax revenue across all specifications. The industrial sector is typically easier to monitor and tax compared to the agricultural sector, and a larger share of manufacturing in GDP reflects more remarkable economic development and the formal (taxable) sector (Gaalya, 2015; Morrissey et al., 2016). In addition, if production is efficient, manufacturing activities can generate higher taxable income through sales, excise, and corporate income taxes (Ahmed and Mohammed, 2010).

Table 3

Regression results of total tax revenue and the first stage: IV estimations

Models	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	IV + FE	IV + FE	IV + FE	IV + FE	IV + FE	IV + FE	IV + FE First	IV + FE First	IV + FE First	IV + FE First	IV + FE First	IV + FE First
D.V.	TR_{it}	TR_{it}	TR_{it}	TR_{it}	TR_{it}	TR_{it}	GVN_{it}	GVN_{it}	GVN_{it}	CC_{it}	CC_{it}	CC_{it}
GVN_{it}	3.843***	2.726***	2.159*									
	(1.215)	(0.963)	(1.114)									
CC_{it}				6.284***	4.789***	3.231**						
				(1.813)	(1.620)	(1.542)						
CVL_{it}							-0.083***	-0.116***	-0.128***	-0.042***	-0.055***	-0.071***
							(0.027)	(0.031)	(0.043)	(0.015)	(0.017)	(0.027)
PR_{it}							-0.073***	-0.084***	-0.031	-0.049***	-0.052***	-0.017
							(0.018)	(0.022)	(0.023)	(0.011)	(0.013)	(0.015)
LOG_ICAP_{it}	2.307***			2.060***			0.402***			0.207***		
	(0.796)			(0.700)			(0.067)			(0.043)		
IND_{it}	0.146***			0.130***			-0.011***			-0.004**		
	(0.035)			(0.035)			(0.003)			(0.002)		
INF_{it}	-0.002			-0.003			-0.072			-0.164		
	(0.057)			(0.092)			(0.424)			(0.605)		
TOP_{it}	0.063***			0.063***			0.003			0.002		
	(0.009)			(0.009)			(0.009)			(0.006)		
LOG_POP_{it}	-0.322			-0.904			-0.980***			-0.410***		
	(2.144)			(1.856)			(0.192)			(0.100)		
SHD_{it}		-0.399***			-0.358***			-0.027**			-0.024***	
		(0.092)			(0.095)			(0.011)			(0.007)	

Models	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	IV + FE	**IV + FE**	**IV + FE**	**IV + FE**	**IV + FE**	**IV + FE**	**IV + FE First**	**IV + FE First**	**IV + FE First**	**IV + FE First**	**IV + FE First**	**IV + FE First**
D.V.	TR_{it}	TR_{it}	TR_{it}	TR_{it}	TR_{it}	TR_{it}	GVN_{it}	GVN_{it}	GVN_{it}	CC_{it}	CC_{it}	CC_{it}
TIP_{it}			0.598**			0.598*			0.104***			0.096***
			(0.257)			(0.324)			(0.030)			(0.017)
U.ID (*p*-val.)							0.000	0.000	0.006	0.000	0.000	0.005
W.ID (*F*-stat.)							23,194	26,484	27,160	29,742	32,775	27,377
Hansen (*p*-val.)							0.678	0.382	0.310	0.500	0.287	0.314
Observations	968	874	575	968	874	575	968	874	575	968	874	575
Countries	78	75	66	78	75	66	78	75	66	78	75	66

*Explanation: T-statistics based on HAC standard errors are reported in parentheses. Both country and time effects are included. D.V. = Dependent variables. U.ID = The p-values associated with the Kleibergen-Paap underidentification test. W.ID = The F-statistics associated with the Kleibergen-Paap weak identification test. Hansen = The p-values associated with the Hansen overisdentification test. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.*

Source: Own calculation.

Table 4
Regression results of total tax revenue: SGMM estimations

Models	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	SGMM One-Step	SGMM Two-Step	SGMM One-Step	SGMM Two-Step	SGMM One-Step	SGMM Two-Step	SGMM One-Step	SGMM Two-Step	SGMM One-Step	SGMM Two-Step	SGMM One-Step	SGMM Two-Step
D.V.	TR_{it}	TR_{it}	TR_{it}	TR_{it}	TR_{it}	TR_{it}	TR_{it}	TR_{it}	TR_{it}	TR_{it}	TR_{it}	TR_{it}
TR_{it-1}	0.788***	0.776***	0.869***	0.932***	0.855***	0.856***	0.823***	0.852***	0.866***	0.940***	0.887***	0.982***
	(0.098)	(0.120)	(0.092)	(0.094)	(0.137)	(0.183)	(0.081)	(0.120)	(0.076)	(0.085)	(0.102)	(0.097)
GVN_{it}	1.105**	1.251**	1.158**	1.217**	1.102*	1.424**						
	(0.555)	(0.488)	(0.583)	(0.493)	(0.633)	(0.720)						
CC_{it}							1.457**	1.668**	1.395*	1.174*	1.164	1.061
							(0.716)	(0.789)	(0.771)	(0.707)	(0.822)	(0.804)
GVN_{it-1}	1.525***	1.328***	1.659***	1.473***	1.513***	1.277**						
	(0.392)	(0.361)	(0.411)	(0.365)	(0.547)	(0.551)						
CC_{it-1}							1.677**	1.211**	1.686***	1.438**	1.617***	1.147**
							(0.528)	(0.550)	(0.571)	(0.594)	(0.574)	(0.581)
LOG_ICAP_{it}	1.069***	1.073***					1.057***	1.043***				
	(0.209)	(0.151)					(0.147)	(0.161)				
IND_{it}	0.222*	0.232**					0.220**	0.218*				
	(0.116)	(0.118)					(0.114)	(0.120)				
INF_{it}	-0.005	-0.006					-0.008	-0.007				
	(0.047)	(0.041)					(0.051)	(0.049)				
TOP_{it}	0.111***	0.112***					0.110***	0.112***				
	(0.019)	(0.010)					(0.018)	(0.011)				
LOG_POP_{it}	-0.136	-0.129					-0.057	-0.041				
	(0.106)	(0.107)					(0.089)	(0.107)				

Models	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	SGMM One-Step	**SGMM Two-Step**	**SGMM One-Step**	**SGMM Two-Step**	**SGMM One-Step**	**SGMM Two-Step**	**SGMM One-Step**	**SGMM Two-Step**	**SGMM One-Step**	**SGMM Two-Step**	**SGMM One-Step**	**SGMM Two-Step**
D.V.	TR_{it}	TR_{it}	TR_{it}	TR_{it}	TR_{it}	TR_{it}	TR_{it}	TR_{it}	TR_{it}	TR_{it}	TR_{it}	TR_{it}
SHD_{it}			-0.123***	-0.115***					-0.120**	-0.115***		
			(0.014)	(0.014)					(0.010)	(0.010)		
TIP_{it}					0.514**	0.630**					0.559***	0.644***
					(0.232)	(0.270)					(0.180)	(0.188)
AR2 (*p*-val.)	0.726	0.727	0.588	0.592	0.292	0.303	0.738	0.717	0.595	0.602	0.296	0.307
Hansen (*p*-val.)	0.791	0.791	0.593	0.593	0.541	0.541	0.330	0.330	0.264	0.264	0.614	0.614
Diff. in Hansen (*p*-val.)	0.477	0.477	0.265	0.265	0.342	0.342	0.120	0.120	0.111	0.111	0.335	0.335
Instruments	32	32	27	27	22	22	32	32	27	27	22	22
Observations	875	875	785	785	544	544	875	875	785	785	544	544
Countries	77	77	75	75	67	67	77	77	75	75	67	67

*Explanation: HAC and Windmeijer's (2005) robust standard errors are reported in parentheses on one-step and two-step estimations, respectively. Both country and time effects are included. D.V. = Dependent variables. U.ID = The p-values associated with the Kleibergen-Paap underidentification test. W.ID = The F-statistics associated with the Kleibergen-Paap weak identification test. Hansen = The p-values associated with the Hansen overidentification test. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.*

Source: Own calculation.

In all specifications, inflation has a weak negative effect on total tax revenue. It is evident that inflation undermines people's purchasing power, thereby distorting the measurement of tax income bases (Immervoll, 2005; Ayenew, 2016). These results are congruent with Ghura (1998) and Agbeyegbe, Stotsky and WoldeMariam (2004), who posited that with skyrocketing prices, the intention of entrepreneurs to avoid taxes would increase, such as splitting the parent company into several smaller companies engaged in the informal economy, while consumers can switch to spending on goods that are less likely to be taxed. Our findings also reveal a positive and significant relationship between trade openness and aggregate tax revenue. A possible plausible reason is that trade liberalization can generate broader revenue bases since the transaction volume in such an economic environment tends to be greater (Agbeyegbe, Stotsky and WoldeMariam, 2004; Castro and Camarillo, 2014). In addition, imports and exports are easier to collect because they are carried out in specific locations (Gupta, 2007). The shadow economy as expected worsens tax performance. This result is in line with Cobham (2005), who estimated that developing countries lose USD 285 billion per year due to tax avoidance in the domestic shadow economy. Lastly, political trust has proven to be an important determinant in increasing a country's tax revenues.

Table 4 displays results from the SGMM estimations, which now investigate the contemporaneous and lagged effects of institutional variables on total tax revenue in a dynamic framework as defined in equation 2. According to the results shown in table 4, we find that the Hansen J test across all specifications failed to reject the null hypothesis of the exogeneity of all instruments since their p-values are above 0.1. It provides some evidence that our instruments are uncorrelated with the outcome of interest. The difference-in-Hansen tests also reveal that our GMM-style instruments are valid, meaning that they are orthogonal to tax performance variables (p-value > 0.1). Moreover, the Arellano-Bond test for AR (2) across all regressions failed to reject the null hypothesis of no second-order serial correlation (p-value > 0.1). This implies that the original disturbances are serially uncorrelated, and the moment conditions are correctly specified, making our estimations safe from bias.

Some essential findings emerge in table 4. The positive and significant effect (p-value < 0.01) of the first-lag of tax revenue in its contemporaneous form in all estimations (both one-step and two-step) provides strong evidence for the persistent nature of tax revenue over time. Overall, the estimated (positive) lagged effects of governance and corruption on total tax revenues are more prominent in magnitude and statistically more significant than their respective contemporaneous effects. These results are robust across all models. Quantitatively, a 1-point increase in the lagged and contemporary governance indexes resulted in a rise in approximately 1.1-1.4 percentage points and 1.3-1.7 percentage points in aggregate tax revenue, respectively (based on models 1-6). Under the same scenario, a 1-point increase in the lagged and contemporary corruption control indexes can be associated with an increase of approximately 1.2-1.7 percentage points and 1.1-1.7 percentage points in total tax generation, respectively (based on models 7-12). As noted earlier, the lagged effects of governance and control of corruption are preferred because they will provide better and more logical guidance for public authorities. Any public policy that is being executed today can generate benefit only in the future. Thus, improving the institutional environment by fighting corruption and increasing governance effectiveness must be considered and implemented in advance so that the tax collection can be realized immediately.

4.3 REGRESSION ANALYSIS OF THREE TYPES OF TAX REVENUE

To further explore which types of taxes are most influenced by the institutional environment, we also break down the regression according to three familiar sources of tax revenue: direct tax, indirect tax, and trade tax. The analysis was executed with IV estimation and the results are reported in table 5. Civil liberties and political rights show consistency as adequate instruments for governance and corruption across all regressions. It is indicated by the *p*-value of the over-identification test results which are not significant at the 5% level almost in all regressions. The impact of governance and corruption are evidenced in direct taxes, indirect taxes, and trade taxes across models until public trust in politicians is used as the only control variable. Statistically, the significance levels of the governance and corruption effects vary depending on the type of tax. As we see, the tax revenue response to governance and corruption appears to be more substantial in indirect taxes (p-value < 0.01) than in trade taxes (p-value < 0.05) and direct taxes (p-value < 0.1). These results are somewhat similar to what Epaphra and Massawe (2017) found in Africa. They consider institutional variables are important determinants of indirect taxes and trade taxes but not direct taxes. It implies that direct taxes tend to be more resilient when corruption is rampant, and the quality of governance is unsatisfactory. Thus, if governments need to increase tax revenue by minimizing distortions and maximizing social welfare, they must implement reforms that reduce corruption or increase revenues from tax categories less prone to corruption (Imam and Jacobs, 2014). In the case of this study, revenue-raising efforts might focus on direct taxes, which have the weakest response to corruption compared to other types of taxes.

Table 5 also delivers the estimation results of other exogenous variables. The percentage of industrial value-added in GDP, economic openness, shadow economy, and public trust have coefficient signs under the hypothesis in all regressions but with varying effect sizes and significance levels. Income per capita has a positive and significant effect on direct taxes (p-value < 0.01) and indirect taxes (p-value < 0.05) but has an opposite direction and insignificant effect on trade taxes (p-value > 0.1). The negative relationship of GDP per capita and trade tax revenue confirms the findings of Epaphra and Massawe (2017), who argued that developing countries that are still in the early stages of development have poor tax administration capabilities. Consequently, they rely heavily on trade taxes as their source of revenue which is easier to collect and enforce than domestic taxes. Inflation has a weak negative effect on direct taxes and indirect taxes. However, the opposite empirical results were found on inflation in the trade tax revenue regression. The positive impact of inflation may indicate that an increase in the price of domestic goods triggers a shift in consumer preferences towards imported goods, which in turn increases the trade tax base. Population has a positive and significant effect (p-value < 0.05) only on indirect taxes and has the opposite direction to direct taxes. These results corroborate the two earlier analyzes: (1) the direct taxes are very vulnerable to evasion, and (2) population growth does not necessarily go hand in hand with increased taxpayer compliance.

Table 5

Regression results of three types of tax revenue: IV estimations

Models	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
	IV + FE	IV + FE	IV + FE	IV + FE	IV + FE	IV + FE	IV + FE	IV + FE	IV + FE	IV + FE	IV + FE	IV + FE	IV + FE	IV + FE	IV + FE	IV + FE	IV + FE	IV + FE
D.V.	DTR_{it}	DTR_{it}	DTR_{it}	DTR_{it}	DTR_{it}	DTR_{it}	$IDTR_{it}$	$IDTR_{it}$	$IDTR_{it}$	$IDTR_{it}$	$IDTR_{it}$	$IDTR_{it}$	TTR_{it}	TTR_{it}	TTR_{it}	TTR_{it}	TTR_{it}	TTR_{it}
GVN_{it}	0.888*	0.789*	0.534				2.635***	1.570***	1.251				1.010**	0.824**	0.669			
	(0.538)	(0.461)	(0.697)				(0.772)	(0.587)	(0.976)				(0 482)	(0.334)	(0.749)			
CC_{it}				1.442*	1.378*	0.987				4.343***	2.763***	2.483				1.635**	1.404**	1.190
				(0.868)	(0.796)	(1.296)				(1.184)	(1.013)	(1.788)				(0.758)	(0.598)	(1.335)
LOG_ICAP_{it}	1.208***			1.233***			1.142**			1.094**			-0.158			-0.227		
	(0.418)			(0.396)			(0.024)			(0.433)			(0.363)			(0.338)		
IND_{it}	0.171***			0.170***			0.084***			0.075***			0.029**			0.023**		
	(0.023)			(0.023)			(0.017)			(0.016)			(0.012)			(0.011)		
INF_{it}	-0.003			-0.004			-0.016			-0.021			0.006			0.003		
	(0.002)			(0.003)			(0.403)			(0.627)			(0.016)			(0.024)		
TOP_{it}	0.030***			0.030***			0.020***			0.020***			0.012***			0.012***		
	(0.005)			(0.005)			(0.005)			(0.005)			(0.003)			(0.003)		
LOG_POP_{it}	-1.359			-1.465			0 011**			0.017**			1.003			0.674		
	(1.143)			(1.053)			(0 005)			(0.008)			(0.798)			(0.734)		
SHD_{it}		-0.233***			-0.227***			-0.096*			-0.072			-0.064**			-0.052*	
		(0.051)			(0.053)			(0.052)			(0.056)			(0.026)			(0.027)	
TIP_{it}			0.087**			0.048**			0.243*			0.263			0.324**			0.377**
			(0.036)			(0.024)			(0.143)			(0.192)			(0.149)			(0.181)
U.ID (*p*-val.)	0.000	0.000	0.006	0.000	0.000	0.005	0.000	0.000	0.005	0.000	0.000	0.005	0.000	0.000	0.006	0.000	0.000	0.005
W.ID (*F*-stat.)	23,194	26,484	7,160	29,742	32,775	7,377	23,194	32,775	7,377	29,742	32,775	7,377	23,194	26,484	7,160	29,742	32,775	7,377

Models	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
	IV + FE	IV + FE	IV + FE	IV + FE	IV + FE	IV + FE	IV + FE	IV + FE	IV + FE	IV + FE	IV + FE	IV + FE	IV + FE	IV + FE	IV + FE	IV + FE	IV + FE	IV + FE
D.V.	DTR_{it}	DTR_{it}	DTR_{it}	DTR_{it}	DTR_{it}	DTR_{it}	$IDTR_{it}$	$IDTR_{it}$	$IDTR_{it}$	$IDTR_{it}$	$IDTR_{it}$	$IDTR_{it}$	TTR_{it}	TTR_{it}	TTR_{it}	TTR_{it}	TTR_{it}	TTR_{it}
Hansen (*p*-val.)	0.615	0.294	0.578	0.591	0.266	0.583	0.968	0.340	0.095	0.767	0.340	0.100	0.640	0.939	4.696	0.555	0.934	0.027
Observations	968	874	575	968	874	575	968	874	575	968	874	575	968	874	575	968	874	575
Countries	78	75	66	78	75	66	78	75	66	78	75	66	78	75	66	78	75	66

*Explanation: T-statistics based on HAC standard errors are reported in parentheses. Both country and time effects are included. D.V. = Dependent variables. U.ID = The p-values associated with the Kleibergen-Paap underidentification test. W.ID = The F-statistics associated with the Kleibergen-Paap weak identification test. Hansen = The p-values associated with the Hansen overidentification test. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.*

Source: Own calculation.

Table 6

Regression results of direct tax revenue: SGMM estimations

Models	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	SGMM One-Step	SGMM Two-Step	SGMM One-Step	SGMM Two-Step	SGMM One-Step	SGMM Two-Step	SGMM One-Step	SGMM Two-Step	SGMM One-Step	SGMM Two-Step	SGMM One-Step	SGMM Two-Step
D.V.	DTR_{it}	DTR_{it}	DTR_{it}	DTR_{it}	DTR_{it}	DTR_{it}	DTR_{it}	DTR_{it}	DTR_{it}	DTR_{it}	DTR_{it}	DTR_{it}
DTR_{it-1}	0.681***	0.864***	0.456*	0.803***	0.515*	0.895***	0.776***	0.949***	0.591***	0.913***	0.613***	0.796***
	(0.139)	(0.132)	(0.241)	(0.242)	(0.280)	(0.263)	(0.132)	(0.159)	(0.223)	(0.306)	(0.229)	(0.241)
GVN_{it}	0.575**	0.647**	0.734	0.752**	0.837*	0.821**						
	(0.294)	(0.258)	(0.451)	(0.371)	(0.478)	(0.392)						
CC_{it}							0.613*	0.537	0.677**	0.582	0.825*	0.715
							(0.357)	(0.422)	(0.311)	(0.425)	(0.465)	(0.687)
GVN_{it-1}	1.303***	0.979***	1.234**	0.890***	1.376**	0.992***						
	(0.296)	(0.239)	(0.486)	(0.267)	(0.600)	(0.329)						
CC_{it-1}							0.980*	0.725*	1.199*	0.895**	1.292*	0.884
							(0.520)	(0.431)	(0.644)	(0.433)	(0.734)	(0.629)
LOG_ICAP_{it}	1.723***	1.666***					1.632***	1.540***				
	(0.172)	(0.176)					(0.124)	(0.131)				
IND_{it}	0.233***	0.203***					0.200***	0.170***				
	(0.024)	(0.018)					(0.018)	(0.014)				
INF_{it}	-0.008	-0.002					-0.003	-0.003				
	(0.034)	(0.041)					(0.032)	(0.036)				
TOP_{it}	0.030***	0.020***					0.030***	0.040***				
	(0.004)	(0.005)					(0.004)	(0.005)				
LOG_POP_{it}	-0.007	-0.021					-0.044	-0.033				
	(0.077)	(0.082)					(0.060)	(0.072)				
SHD_{it}			-0.160***	-0.160***					-0.090***	-0.060***		
			(0.016)	(0.018)					(0.011)	(0.015)		

Models	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	SGMM One-Step	**SGMM Two-Step**	**SGMM One-Step**	**SGMM Two-Step**	**SGMM One-Step**	**SGMM Two-Step**	**SGMM One-Step**	**SGMM Two-Step**	**SGMM One-Step**	**SGMM Two-Step**	**SGMM One-Step**	**SGMM Two-Step**
D.V.	DTR_{it}	DTR_{it}	DTR_{it}	DTR_{it}	DTR_{it}	DTR_{it}	DTR_{it}	DTR_{it}	DTR_{it}	DTR_{it}	DTR_{it}	DTR_{it}
TIP_{it}					0.401*	0.472*					0.600**	0.708**
					(0.229)	(0.269)					(0.284)	(0.355)
AR2 (*p*-val.)	0.140	0.185	0.165	0.216	0.211	0.377	0.153	0.194	0.174	0.221	0.215	0.416
Hansen (*p*-val.)	0.265	0.265	0.198	0.198	0.376	0.376	0.083	0.083	0.140	0.140	0.123	0.123
Diff. in Hansen (*p*-val.)	0.086	0.086	0.061	0.061	0.158	0.158	0.081	0.081	0.213	0.213	0.088	0.088
Instruments	32	32	27	27	22	22	32	32	27	27	22	22
Observations	875	875	785	785	544	544	875	875	785	785	544	544
Countries	77	77	75	75	67	67	77	77	75	75	67	67

*Explanation: HAC and Windmeijer's (2005) robust standard errors are reported in parentheses on one-step and two-step estimations, respectively. Both country and time effects are included. D.V. = Dependent variables. U.ID = The p-values associated with the Kleibergen-Paap underidentification test. W.ID = The F-statistics associated with the Kleibergen-Paap weak identification test. Hansen = The p-values associated with the Hansen overidentification test. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.*

Source: Own calculation.

Table 7

Regression results of indirect tax revenue: SGMM estimations

Models	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	SGMM One-Step	SGMM Two-Step	SGMM One-Step	SGMM Two-Step	SGMM One-Step	SGMM Two-Step	SGMM One-Step	SGMM Two-Step	SGMM One-Step	SGMM Two-Step	SGMM One-Step	SGMM Two-Step
D.V.	$IDTR_{it}$	$IDTR_{it}$	$IDTR_{it}$	$IDTR_{it}$	$IDTR_{it}$	$IDTR_{it}$	$IDTR_{it}$	$IDTR_{it}$	$IDTR_{it}$	$IDTR_{it}$	$IDTR_{it}$	$IDTR_{it}$
$IDTR_{it-1}$	0.767***	0.751***	0.839***	0.760***	0.695***	0.527*	0.753***	0.769***	0.842***	0.781***	0.764***	0.781***
	(0.144)	(0.163)	(0.127)	(0.133)	(0.184)	(0.291)	(0.147)	(0.143)	(0.106)	(0.163)	(0.153)	(0.257)
GVN_{it}	1.356***	1.366***	1.178***	1.388***	1.616***	1.681***						
	(0.402)	(0.460)	(0.444)	(0.420)	(0.521)	(0.546)						
CC_{it}							1.565***	1.948***	1.453**	1.916**	1.332*	1.370
							(0.606)	(0.692)	(0.612)	(0.898)	(0.690)	(0.968)
GVN_{it-1}	2.153***	2.324***	2.206***	2.242***	2.079***	2.346***						
	(0.264)	(0.256)	(0.268)	(0.262)	(0.440)	(0.435)						
CC_{it-1}							2.024***	2.129***	2.353***	2.299***	2.327***	2.232***
							(0.393)	(0.435)	(0.412)	(0.453)	(0.533)	(0.422)
LOG_ICAP_{it}	1.125***	1.112***					1.118***	1.020***				
	(0.131)	(0.121)					(0.132)	(0.116)				
IND_{it}	0.144***	0.140***					0.144***	0.135***				
	(0.024)	(0.024)					(0.025)	(0.022)				
INF_{it}	-0.005	-0.007					-0.007	-0.001				
	(0.003)	(0.005)					(0.005)	(0.007)				
TOP_{it}	0.100***	0.111***					0.100***	0.120***				
	(0.007)	(0.009)					(0.008)	(0.010)				
LOG_POP_{it}	0.160**	0.176**					0.143*	0.144				
	(0.079)	(0.087)					(0.082)	(0.106)				

Models	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	SGMM One-Step	**SGMM Two-Step**	**SGMM One-Step**	**SGMM Two-Step**	**SGMM One-Step**	**SGMM Two-Step**	**SGMM One-Step**	**SGMM Two-Step**	**SGMM One-Step**	**SGMM Two-Step**	**SGMM One-Step**	**SGMM Two-Step**
D.V.	$IDTR_{it}$	$IDTR_{it}$	$IDTR_{it}$	$IDTR_{it}$	$IDTR_{it}$	$IDTR_{it}$	$IDTR_{it}$	$IDTR_{it}$	$IDTR_{it}$	$IDTR_{it}$	$IDTR_{it}$	$IDTR_{it}$
SHD_{it}			-0.120***	-0.131***					-0.118***	-0.121***		
			(0.017)	(0.017)					(0.012)	(0.013)		
TIP_{it}					0.645*	0.804*					0.324	0.232
					(0.340)	(0.482)					(0.283)	(0.422)
AR2 (*p*-val.)	0.609	0.604	0.569	0.613	0.607	0.751	0.629	0.596	0.582	0.592	0.583	0.561
Hansen (*p*-val.)	0.233	0.233	0.146	0.146	0.290	0.290	0.270	0.270	0.270	0.270	0.193	0.193
Diff. in Hansen (*p*-val.)	0.567	0.567	0.239	0.239	0.107	0.107	0.160	0.160	0.140	0.140	0.130	0.130
Instruments	32	32	27	27	22	22	32	32	27	27	22	22
Observations	875	875	785	785	544	544	875	875	785	785	544	544
Countries	77	77	75	75	67	67	77	77	75	75	67	67

*Explanation: HAC and Windmeijer's (2005) robust standard errors are reported in parentheses on one-step and two-step estimations, respectively. Both country and time effects are included. D.V. = Dependent variables. U.ID = The p-values associated with the Kleibergen-Paap underidentification test. W.ID = The F-statistics associated with the Kleibergen-Paap weak identification test. Hansen = The p-values associated with the Hansen overidentification test. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.*

Source: Own calculation.

Table 8

Regression results of trade tax revenue: SGMM estimations

Models	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	SGMM One-Step	**SGMM Two-Step**	**SGMM One-Step**	**SGMM Two-Step**	**SGMM One-Step**	**SGMM Two-Step**	**SGMM One-Step**	**SGMM Two-Step**	**SGMM One-Step**	**SGMM Two-Step**	**SGMM One-Step**	**SGMM Two-Step**
D.V.	TTR_{it}	TTR_{it}	TTR_{it}	TTR_{it}	TTR_{it}	TTR_{it}	TTR_{it}	TTR_{it}	TTR_{it}	TTR_{it}	TTR_{it}	TTR_{it}
TTR_{it-1}	0.740***	0.775***	0.749***	0.872***	0.710***	0.771***	0.750***	0.776***	0.735***	0.795***	0.725***	0.817***
	(0.076)	(0.109)	(0.099)	(0.132)	(0.101)	(0.132)	(0.069)	(0.088)	(0.084)	(0.080)	(0.079)	(0.091)
GVN_{it}	0.834***	0.817***	0.955***	0.953***	1.097***	1.042***						
	(0.169)	(0.151)	(0.161)	(0.147)	(0.189)	(0.178)						
CC_{it}							0.707***	0.761***	0.808***	0.827***	0.918***	0.915***
							(0.268)	(0.186)	(0.269)	(0.203)	(0.344)	(0.238)
GVN_{it-1}	1.001***	1.030***	1.007***	1.113***	1.116***	1.103***						
	(0.175)	(0.173)	(0.183)	(0.178)	(0.234)	(0.156)						
CC_{it-1}							1.021***	1.082***	1.092***	0.936***	1.214***	0.938***
							(0.349)	(0.198)	(0.310)	(0.200)	(0.450)	(0.270)
LOG_ICAP_{it}	-0.158	-0.104					-0.142	-0.117				
	(0.136)	(0.107)					(0.119)	(0.095)				
IND_{it}	0.022***	0.022***					0.030***	0.020***				
	(0.008)	(0.004)					(0.008)	(0.005)				
INF_{it}	0.005	0.009					0.001	0.001				
	(0.025)	(0.025)					(0.033)	(0.022)				
TOP_{it}	0.020***	0.010***					0.010***	0.030***				
	(0.002)	(0.002)					(0.002)	(0.001)				
LOG_POP_{it}	0.081	0.074					0.078	0.048				
	(0.068)	(0.050)					(0.050)	(0.034)				

Models	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	SGMM One-Step	SGMM Two-Step	SGMM One-Step	SGMM Two-Step	SGMM One-Step	SGMM Two-Step	SGMM One-Step	SGMM Two-Step	SGMM One-Step	SGMM Two-Step	SGMM One-Step	SGMM Two-Step
D.V.	TTR_{it}	TTR_{it}	TTR_{it}	TTR_{it}	TTR_{it}	TTR_{it}	TTR_{it}	TTR_{it}	TTR_{it}	TTR_{it}	TTR_{it}	TTR_{it}
SHD_{it}			-0.012**	-0.014***					-0.013***	-0.012***		
			(0.005)	(0.005)					(0.005)	(0.004)		
TIP_{it}					0.172*	0.140					0.176*	0.106
					(0.103)	(0.099)					(0.094)	(0.108)
AR2 (*p*-val.)	0.170	0.175	0.130	0.138	0.165	0.169	0.171	0.177	0.131	0.138	0.168	0.172
Hansen (*p*-val.)	0.324	0.324	0.450	0.450	0.417	0.417	0.391	0.391	0.275	0.275	0.433	0.433
Diff. in Hansen (*p*-val.)	0.464	0.464	0.354	0.354	0.432	0.432	0.604	0.604	0.558	0.558	0.570	0.570
Instruments	32	32	27	27	22	22	32	32	27	27	22	22
Observations	875	875	785	785	544	544	875	875	785	785	544	544
Countries	77	77	75	75	67	67	77	77	75	75	67	67

*Explanation: HAC and Windmeijer's (2005) robust standard errors are reported in parentheses on one-step and two-step estimations, respectively. Both country and time effects are included. D.V. = Dependent variables. U.ID = The p-values associated with the Kleibergen-Paap underidentification test. W.ID = The F-statistics associated with the Kleibergen-Paap weak identification test. Hansen = The p-values associated with the Hansen overidentification test. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.*

Source: Own calculation.

We also report the SGMM estimation results of direct tax (table 6), indirect tax (table 7), and trade tax (table 8) regressions. In each table, models 1-6 and models 7-12 summarize the contemporary and lagged effects of the governance and corruption variables on tax performance, respectively. As in the total tax revenue regression, all valid instrument requirements have been met and concerns on the issue of autocorrelation in the disturbance can be ignored in all regressions (see the *p*-values of AR2, Hansen J statistics, and Difference-in-Hansen tests in tables 6-8). The results are still consistent when the dependent variable is split into several taxation categories. The lagged effect of the corruption and governance variables is superior to the contemporary effect. However, there are variations in the magnitude of the effect and the level of significance. In addition, consistent with the results of IV estimations, the contemporary influence of institutional variables (corruption and governance) is more dominant in indirect taxes and trade taxes. Also, the coefficient signs of all control variables included in all regressions have a similar pattern to that of the IV estimations but with minor differences in terms of effect magnitudes and significance levels. Industrial activity, an open economy, and the shadow economy always have a significant effect (at least at the 5% level) and, according to the hypothesis, on all types of taxation. Although public trust in politicians positively affects the three types of tax revenue, the impact is not always significant. Inflation hurts direct and indirect taxes but has a positive and insignificant effect on trade taxes. The income per capita only has a positive and significant effect (p-values < 0.01) on direct and indirect taxes. Lastly, population size only has a positive and slightly significant effect on indirect taxes (p-values $<$ 0.1). Thus, it is safe to declare that our estimation results are robust.

5 CONCLUSION

This paper investigates whether the institutional environment, i.e., governance and corruption, affected tax performance in developing countries in the 2002-2019 period. To obtain more specific results, three types of taxation i.e., direct tax, indirect tax, and trade tax are also used to measure tax performance. In addition, several control variables covering macroeconomic indicators, population size, shadow economy, and public trust are included in the analysis. Since we suspect a bidirectional relationship of institutional variables and tax revenues, we incorporate civil liberties and political rights to instrument corruption and governance. The initial analysis of this study was carried out using the IV estimation model. However, different from previous studies, we provide further analysis by examining the contemporary and lagged effects of institutional variables on tax performance. Therefore, we use both one-step and two-step SGMM models to achieve the final objective of this study.

This study's main result shows that better governance quality and lower levels of corruption benefit governments' overall tax collection. However, the effect of these two institutional variables is more significant in the case of indirect taxes and trade taxes and relatively weak in the case of direct taxes, according to the IV and SGMM models. Intuitively, good governance is frequently associated with an efficient tax administration, speeding up the tax revenue collection process, and encouraging citizens to be more compliant in paying taxes. On the other hand, poor governance is closely related to complex tax regimes, creating more loopholes for tax evasion.

Similarly, a high level of corruption reflects low government accountability, which distorts public trust; hence, voluntary compliance in paying taxes among citizens is undermined. In addition, the SGMM regression results show that the lagged effect is much more substantial than the contemporaneous effect, both in terms of effect magnitude and significance level across all tax revenue measures (total tax, direct tax, indirect tax, and trade tax). This result confirms our assumption of a time-lag symptom of policies directed at institutional environment improvement to support the acquisition of government revenues.

Our empirical results also reveal the critical role of other exogenous variables in determining tax revenue in developing countries. Industrial output as a share of GDP, trade openness, and public trust in politicians have a strong positive effect on all types of tax revenue. On the other hand, the shadow economy consistently has a negative and significant impact on the three categories of tax performance in all regression models. Income per capita only has a positive and significant impact on direct and indirect taxes. In a similar pattern, inflation weakly worsens the performance of direct and indirect taxes but tends to enhance that of trade taxes. Finally, population has a weak positive effect on indirect taxes and trade taxes but has the opposite direction to direct taxes. These results tend to be consistent across all regression models but remain with variation in effect sizes and significance levels.

Our findings may point to some policy implications for governments in developing countries. Reducing corruption and improving governance are the primary efforts that must be implemented to encourage the rate of tax revenue. Given that our empirical results provide evidence that the effect of corruption is substantial on indirect taxes, the government should focus more on fighting bribery practices related to indirect taxes such as value-added tax, sales tax, excise tax, and customs duties. Alternatively, the governments could turn to other sources of tax revenue such as direct taxes which were found to have the weakest response to corruption, implying that this type of tax is less prone to corruption. Corruption in tax administration can be curbed by establishing semi-autonomous tax agencies, higher salaries for tax officers, improved tax services, and reduced interaction of taxpayers and tax officials, for example, by investing in technology and taxes. Moreover, the lagged effects of institutional variables on tax revenue that are much more substantial than their contemporaneous effects in most regressions deliver a valuable insight. The efforts to improve the institutional environment that have been mentioned before must be executed as early as possible to achieve optimal tax performance in the future. Efforts to increase tax revenues also need to be addressed by increasing economic development through increasing per capita income, transforming the economic structure towards industrialization, increasing economic openness, improving public trust, curbing inflation and the proliferation of shadow economic activity, and keep the population size at the ideal level.

Disclosure statement
The authors declare that there is no conflict of interest.

REFERENCES

1. Abu, N. and Staniewski, M. W., 2019. Determinants of corruption in Nigeria: evidence from various estimation techniques. *Economic Research-Ekonomska istrazivanja*, 32(1), pp. 3052-3076. https://doi.org/10.1080/1331677X.2019.1655467
2. Agbeyegbe, T. D., Stotsky, J. G. and WoldeMariam, A., 2004. Trade Liberalization, Exchange Rate Changes, and Tax Revenue in Sub-Saharan Africa. *IMF Working Papers*, WP/04/178. https://doi.org/10.5089/9781451858983.001
3. Ahmed, Q. M. and Mohammed, S. D., 2010. Determinant of tax buoyancy: Empirical evidence from developing countries. *European Journal of Social Sciences*, 13(3), pp. 408-414.
4. Ajaz, T. and Ahmad, E., 2010. The Effect of Corruption and Governance on Tax Revenues. *The Pakistan Development Review*, 49(4), pp. 405-417.
5. Alm, J., Martinez-Vazquez, J. and McClellan, C., 2015. Corruption and firm tax evasion. *Journal of Economic Behavior and Organization*, 124, pp. 146-163. https://doi.org/10.1016/j.jebo.2015.10.006
6. Arellano, M. and Bond, S., 1991. Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *Review of Economic Studies*, 58(2), pp. 277-297. https://doi.org/10.2307/2297968
7. Arellano, M. and Bover, O., 1995. Another look at the instrumental variable estimation of error-components models. *Journal of Econometrics*, 68(1), pp. 29-51. https://doi.org/10.1016/0304-4076(94)01642-D
8. Arif, I. and Rawat, A. S., 2018. Corruption, governance, and tax revenue: Evidence from EAGLE countries. *Journal of Transnational Management*, 23(2-3), pp. 119-133. https://doi.org/10.1080/15475778.2018.1469912
9. Ayenew, W., 2016. Determinants of Tax Revenue in Ethiopia (Johansen Co-Integration Approach). *International Journal of Business, Economics and Management*, 3(6), pp. 69-84. https://doi.org/10.18488/journal.62/2016.3.6/62.6.69.84
10. Benkovskis, K. and Fadejeva, L., 2014. The effect of VAT rate on inflation in Latvia: Evidence from CPI microdata. *Applied Economics*, 46(21), pp. 2520-2533. https://doi.org/10.1080/00036846.2014.904492
11. Besley, T. and Persson, T., 2014. Why do developing countries tax so little? *Journal of Economic Perspectives*, 28(4), pp. 99-120. https://doi.org/10.1257/jep.28.4.99
12. Bird, R. M., Martinez-Vazquez, J. and Torgler, B., 2008. Tax Effort in Developing Countries and High Income Countries: The Impact of Corruption, Voice and Accountability. *Economic Analysis and Policy*, 38(1), pp. 55-71. https://doi.org/10.1016/S0313-5926(08)50006-3
13. Blundell, R. and Bond, S., 1998. Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics*, 87(1), pp. 115-143. https://doi.org/10.1016/S0304-4076(98)00009-8

14. Brondolo, J. [et al.], 2008. Tax Administration Reform and Fiscal Adjustment: The case of Indonesia. *IMF Working Paper,* No. 08/129.
15. Castro, G. Á. and Camarillo, D. B. R., 2014. Determinants of tax revenue in OECD countries over the period 2001-2011. *Contaduria y Administracion*, 59(3), pp. 35-59. https://doi.org/10.1016/s0186-1042(14)71265-3
16. Cobham, A., 2005. Tax evasion, tax avoidance and development finance. *QEH Working Paper Series*, (129), pp. 1-20.
17. Efron, B. and Morris, C., 1971. Limiting the risk of bayes and empirical bayes estimators - part I: The bayes case. *Journal of the American Statistical Association*, 66(336), pp. 807-815. https://doi.org/10.1080/01621459.1971.10482348
18. Efron, B. and Morris, C., 1972. Limiting the risk of Bayes and empirical bayes estimators - Part II: The empirical bayes case. *Journal of the American Statistical Association*, 67(337), pp. 130-139. https://doi.org/10.1080/01621459.1972.10481215
19. Elgin, C. [et al.], 2021. Understanding informality. *C.E.P.R Discussion Paper,* No. 16497. https://doi.org/10.2139/ssrn.3916568
20. Epaphra, M. and Massawe, J., 2017. Corruption, governance and tax revenues in Africa. *Business and Economic Horizons*, 13(4), pp. 439-467. https://doi.org/10.15208/beh.2017.31
21. Everest-Phillips, M. and Sandall, R., 2009. Linking Business Tax Reform with Governance: How to Measure Success. In: *Investment Climate in Practice, No. 2*.
22. Fjeldstad, O. H., 2006. Corruption in tax administration: Lessons from institutional reforms in Uganda. In S. Rose-Ackerman, ed. *International Handbook on the Economics of Corruption*. London: Edward Elgar Publishing, pp. 484-511.
23. Gaalya, M. S., 2015. Trade Liberalization and Tax Revenue Performance in Uganda. *Modern Economy*, 6(2), pp. 228-244. https://doi.org/10.4236/me.2015.62021
24. Ghura, D., 1998. Tax Revenue in Sub-Saharan Africa: Effects of Economic Policies and Corruption. *IMF Working Papers*, WP/98/135. https://doi.org/10.5089/9781451855685.001
25. Goldberger, A. S., 1972. Maximum-Likelihood Estimation of Regressions Containing Unobservable Independent Variables. *International Economic Review*, 13(1), pp. 1-15. https://doi.org/10.2307/2525901
26. Gupta, A. S., 2007. Determinants of Tax Revenue Efforts in Developing Countries. *IMF Working Paper*, WP/07/184.
27. Hassan, M. S. [et al.], 2021. Governance: A Source to Increase Tax Revenue in Pakistan. *Complexity*, 2021. https://doi.org/10.1155/2021/6663536
28. Hotelling, H., 1933. Analysis of a complex of statistical variables into Principal Components. *The Journal of Educational Psychology*, 24, pp. 417-441.
29. Hunady, J. and Orviska, M., 2015. Effect of Corruption on Tax Revenue in the OECD and Latin America Countries. *Theoretical and Practical Aspects of Public Finance*, 1(1), pp. 80-85.

30. Hwang, J., 2002. A Note on the Relationship Between Corruption and Government Revenue. *Journal of Economic Development*, 27(2), pp. 161-178.
31. Imam, P. A. and Jacobs, D., 2014. Effect of Corruption on Tax Revenues in the Middle East. *Review of Middle East Economics and Finance*, 10(1), pp. 1-24. https://doi.org/10.1515/rmeef-2014-0001
32. Immervoll, H., 2005. Falling up the Stairs: The Effects of "Bracket Creep" on Household Incomes. *Review of Income and Wealth*, 51(1), pp. 37-62. https://doi.org/10.1111/j.1475-4991.2005.00144.x
33. Jackson, J. E., 1991. A User's Guide to Principal Components. *The Journal of the Operational Research Society,* 43(6). https://doi.org/10.2307/2583020
34. Jollife, I. T. and Cadima, J., 2016. Principal component analysis: A review and recent developments. *Philosophical Transactions of the Royal Society: A Mathematical, Physical and Engineering Sciences*, 374(2065). https://doi.org/10.1098/rsta.2015.0202
35. Junquera-Varela, R.-F. and Haven, B., 2018. *Getting to 15 percent: addressing the largest tax gaps*. World Bank Blogs.
36. Karagöz, K., 2013. Determinants of Tax Revenue: Does Sectorial Composition Matter? *Journal of Finance, Accounting and Management*, 4(2), pp. 50-63.
37. Kaufmann, D., Kraay, A. and Mastruzzi, M., 2010. The worldwide governance indicators: Methodology and analytical issues. *World Bank Policy Research Working Paper,* No. 5430. https://doi.org/10.1017/S1876404511200046
38. Le, T. M., Moreno-Dodson, B. and Rojchaichaninthorn, J., 2008. Expanding Taxable Capacity and Reaching Revenue Potential: Cross-Country Analysis. *World Bank Policy Research Paper*, No. 4459.
39. Lippert, O. and Walker, M., 1997. *The Underground Economy: Global Evidences of its Size and Impact*. Vancouer: T. F. Institute.
40. Moore, M. and Prichard, W., 2020. *How Can Governments of Low-Income Countries Collect More Tax Revenue?* https://doi.org/10.1007/978-3-030-37595-9_4
41. Morrissey, O. [et al.], 2016. Tax Revenue Performance and Vulnerability in Developing Countries. *Journal of Development Studies*, 52(12), pp. 1689-1703. https://doi.org/10.1080/00220388.2016.1153071
42. Nawaz, F., 2010. Exploring the Relationships between Corruption and Tax Revenue. In: *U4 Helpdesk*.
43. Olken, B. A. and Pande, R., 2012. Corruption in developing countries. *Annual Review of Economics*, 4, pp. 479-509. https://doi.org/10.1146/annurev-economics-080511-110917
44. Prichard, W., 2010. Taxation and State Building: Towards a Governance Focused Tax Reform Agenda. *IDS Working Papers,* No. 341. https://doi.org/10.1111/j.2040-0209.2010.00341_2.x
45. Rahman, A., 2009. Tackling Corruption Through Tax Administration Reform. In: *Investments Climate In Practise,* No. 3.

46. Roodman, D., 2009. How to do xtabond2: An introduction to difference and system GMM in Stata. *Stata Journal*, 9(1), pp. 86-136. https://doi.org/10.1177/1536867x0900900106
47. Saporta, G. and Niang, N., 2009. Principal Component Analysis: Application to Statistical Process Control. In G. Govaert, ed. *Data Analysis*. John Wiley & Sons, pp. 1-23. https://doi.org/https://doi.org/10.1002/9780470611777.ch1
48. Schneider, F., 2005. Shadow economies around the world: What do we really know? *European Journal of Political Economy*, 21(3), pp. 598-642. https://doi.org/10.1016/j.ejpoleco.2004.10.002
49. Syadullah, M., 2015. Governance and Tax Revenue in Asean Countries. *Journal of Social and Development Sciences*, 6(2), pp. 76-88. https://doi.org/10.22610/jsds.v6i2.845
50. Transparency International, 2014. *Tax Administration and Corruption Topic Guide*.
51. Windmeijer, F., 2005. A finite sample correction for the variance of linear efficient two-step GMM estimators. *Journal of Econometrics*, 126(1), pp. 25-51. https://doi.org/10.1016/j.jeconom.2004.02.005
52. World Bank, 2021. *World Development Indicators.* Washington D.C: World Bank.