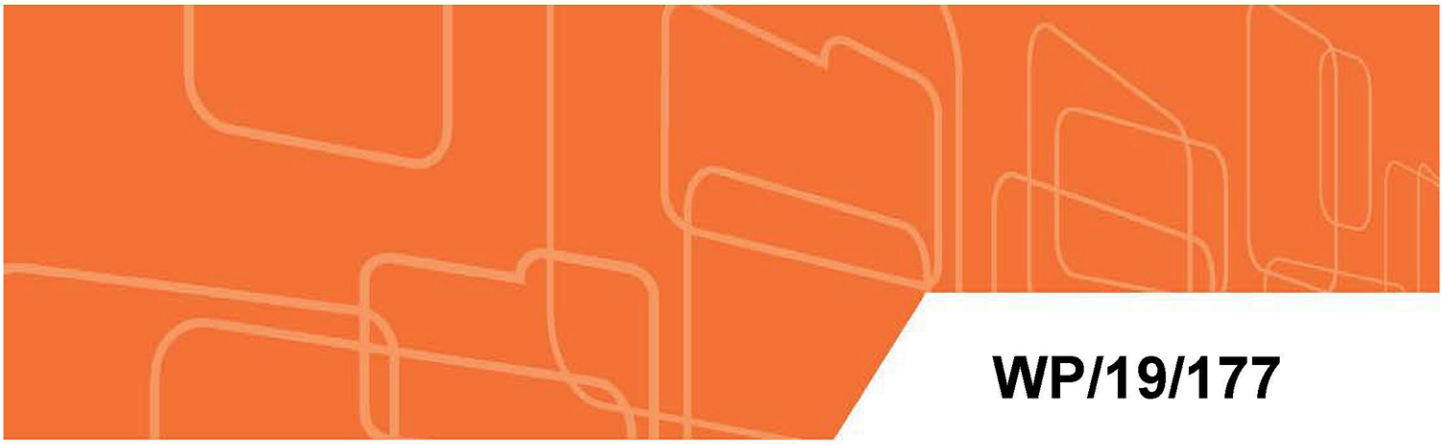


WP/19/177

IMF Working Paper

State Institutions and Tax Capacity: An Empirical Investigation of Causality

by Olusegun Ayodele Akanbi

INTERNATIONAL MONETARY FUND

WP/19/177

IMF Working Paper

State Institutions and Tax Capacity: An Empirical Investigation of Causality

by Olusegun Ayodele Akanbi

INTERNATIONAL MONETARY FUND

 WP/19/177

IMF Working Paper

Fiscal Affairs Department

State Institutions and Tax Capacity: An Empirical Investigation of Causality

Prepared by Olusegun Ayodele Akanbi[1]

Authorized for distribution by Nikolay Gueorguiev

August 2019

Abstract

Would better state institutions increase tax collection, or would higher tax collection help improve state institutions? In the absence of conclusive guidance from theory, this paper searches for an empirical answer to this question, using a panel dataset covering 110 non-resource-rich countries from 1996 to 2017. Employing a panel vector error correction model, the paper finds that tax capacity and state institutions cause and reinforce each other for a wide range of country groups. The bi-directional causality results suggest that developing tax capacity and building state institutions need to go hand in hand for best results, particularly in developing countries. Based on the impulse response analyses, the paper also finds that the causal effects in advanced economies are generally low in both directions, while in developing countries, both tax capacity and institutions shocks have larger positive impacts on institutions and tax capacity, respectively.

JEL Classification Numbers: H11, H26, 023, 055, C23

Keywords: Tax capacity, tax revenue, institutions, causality, GDP

Author's E-Mail Address: OAkanbi@imf.org

[1] I greatly appreciate the special guidance received from Nikolay Gueorguiev, Jiro Honda, and Bernardin Akitoby. I am also grateful for helpful comments and suggestions from Kenji Moriyama, Luca Ricci, Hiro Miyamoto, Mick Thackray, Aksel Sorensen, Enrique Rojas, Cian Ruane, and Aart Kraay (World Bank).

Content **Page**

I. INTRODUCTION AND BACKGROUND

1. **Enhancing tax collection has always been a central theme for economic development, particularly in low-income countries (LICs)**. In recent years the debate on improving LICs' tax capacity has intensified in view of these countries' large development needs (e.g., to meet the Sustainable Development Goals (SDG)). Many developing countries face acute financing needs for critical social and infrastructure spending. According to the estimate by Gaspar et al. (2019), delivering on the SDG agenda would require additional spending in 2030 of US$0.5 trillion (15 percentage points of GDP) in LICs.

2. **To address the paramount need for enhanced tax capacity, a key question remains outstanding**. Should countries develop state institutions first to help them raise more tax revenues, or should sufficient tax revenues be generated first to help develop the institutions? In other words, which way does a causality run between state institutions and tax capacity? To the extent of the author's knowledge, the existing literature (e.g., Brautigam, 2008; Besley and Persson, 2009 and 2014) offers mainly descriptive arguments about the causality between state institutions and tax capacity. The indirect nexus between tax capacity and institutions is undoubted, but only empirical analysis can provide evidence on how exactly they are linked.

3. **This paper investigates this question by empirically testing the long run causality between tax capacity and state institutions**. Specifically, it examines whether state institutions affect tax capacity and/or if changes in tax capacity lead to changes in state institutions. Throughout the paper, tax capacity is defined as the ability to raise domestic revenue, measured as the ratio of tax revenue to GDP (Gaspar et al., 2016a), while state institutions (or institutions)—i.e., rule of law, effectiveness of government, corruption control—are defined broadly as the traditions by which authority in a country is exercised.[2]

4. **The analysis in this paper builds on the technique (based on a panel vector error correction model) pioneered by Hurlin and Venet (2001)**. In the absence of sufficiently long time-series to conduct a standard Granger causality test, this technique allows us to estimate the long-run causality between tax capacity and state institutions, using a panel data of 110 non-resource-rich countries between 1996 and 2017. The empirical methodology adopts a multivariate approach to causality testing rather than a bivariate approach and focused on detecting long run rather than short run causality.

5. **The empirical results reveal a strong existence of a long run bi-directional causality between tax capacity and institutions for all country groups**. It confirms that a simultaneous change in institutions and tax capacity is needed to achieve best results. The finding also

[2] Institutions are also the rules of the game in a society or the humanly devised constraints that shaped human interaction (North, 1990; Kaufmann *et al,* 1999). They are regarded as an important part of state-building. From a broader perspective, state-building is defined by Fjeldstad and Moore (2008) as the increasing capacity of governments to interact constructively with societal interests, to obtain support and resources from those interests, and to pursue consistent lines of action.

supports the argument for the existence of a tipping point in tax-to-GDP levels, above which a significant acceleration of growth and development is observed (Gaspar et al., 2016a). In terms of magnitudes, this paper finds relatively low quantitative interaction between tax capacity and institutions in high-income countries (HICs), while the effects are generally high in developing countries.

6. **These results have important policy implications.** The bidirectional causality suggests that countries with low tax capacity and weak institutions may well be trapped in a vicious cycle, with underdeveloped institutions thwarting tax capacity, which in turn further weakens institutions. Enhanced efforts on multiple fronts may facilitate a virtuous cycle with enhanced tax capacity generating funds to improve institutions and with better institutions enhancing tax capacity. To achieve best outcomes, countries should intensify efforts to develop institutions and raise tax capacity simultaneously, as they reinforce each other. To enhance tax capacity for further institutional building, a strong and coherent medium-term revenue strategy (MTRS) can provide a robust framework for the necessary structural reforms of the tax system.

7. **The rest of the paper is organised as follows**: Section II presents the channels of causality between state institutions and tax capacity and develops the hypotheses to be tested; Section III presents the data analysis and stylised facts; Section IV describes the empirical strategy (methodology and estimation techniques); Section V discusses empirical results; Section VI concludes with policy implications.

II. POTENTIAL CHANNELS OF CAUSALITY

8. **This section discusses hypotheses on the link between institutions and tax capacity, based on the literature**. Specifically, it focuses on how changes in the states' institutional development could lead to changes in tax capacity and how changes in tax capacity could affect state institutions. Both potential directions of causality between quality of institutions and tax capacity have support in the theoretical literature.

(i) ***Hypothesis 1:*** *Changes in the tax capacity causes a subsequent change in the quality of state institutions*.

- *This hypothesis is not fully confirmed in the empirical literature*. Prichard and Leonard (2010) empirically test this hypothesis (based on cross-country dataset for sub-Saharan African countries) and find that improvements in tax revenue tended to precede state institutional improvements for the period 1973 to 1990, though such effects disappear post-1990. They therefore conclude that these results provide tentative support for the hypothesis that improvements in tax collection can be a catalyst for broader gains in state institutions, but that such linkages are not guaranteed and depend on the character of reform. Some studies (Brautigam, 2008; Bornhorst, Gupta, and Thornton, 2008; Prichard and Leonard, 2010; Gaspar, Jaramillo, and Wingender, 2016b) argued that taxation is a central means to building state

institutions. Based on case studies, these studies argue that governments tend to be more accountable and effective when they demand more taxes from their populace.

- *There are several potential reasons to expect that improvement in institutions may follow tax capacity*. First, improvement in tax capacity may increase the scope for spending to strengthen institutions. Any further improvement in tax capacity may require more efficient institutional structures, thereby creating stronger incentives for public investment in improved institutions (Brautigam, 2008). Second, enhanced tax capacity can provide governments with resources required to implement potentially expensive overall reforms which will eventually strengthen institutions in the economy. In addition, governments will be able to shift towards appropriate revenue sources that are more progressive, thereby creating more effective tax administrations (Fjeldstad and Moore, 2008). Third, improvement in tax capacity can help create a constituency of businesses and citizens with the interest and ability to advocate improvements in institutions which encourages constructive state-society engagement around taxes. Since rising tax revenue indicates shrinking profits of businesses and incomes of households, it creates awareness among the populace to hold their governments accountable for the use of the tax proceeds. This process generates critical mass of specific societal institutions which eventually strengthens state institutions (Gadenne, 2017; Martin, 2013).

(ii) ***Hypothesis 2.*** *Changes in the quality of institutions cause a subsequent change in tax capacity*.

- *This hypothesis is supported by several empirical studies*. Several studies (i.e. Ghura, 1998; Gupta, 2007) on the determinants of tax capacity found that improvement in state institutions would raise tax revenue.[3] Improving general state institutions may be associated with improvement in revenue institutions (i.e., better revenue administration), which would then boost tax capacity. Rodrik, Subramanian, and Trebbi (2004) conclude that institutions rule over all economic objectives and therefore suggest that building stronger institutional structures will partly increase countries' capacity to carry out their economic functions which in this case includes tax capacity. Besley and Persson (2014) analyzed the role played by state institutions in generating tax revenue and highlighted institutional structure as one main reason why poor countries are poor which could also explain their weakness in raising tax revenue.
- *Other studies also emphasized that countries with weak institutions are unlikely to have strong motives to build tax capacity*.[4] In such cases, building strong tax capacity is often not in the interest of those who dominate the political institutions. On the contrary, these political

[3] According to recent estimate by IMF (2019), an improvement in the control of corruption (an indicator of institutional quality) by one-third of a standard deviation is associated with an increase of 1.2 percentage points in government revenue as a share of GDP.

[4] These studies include Burgess and Stern (1993); Bird, Martinez-Vazquez, and Torgler (2008); Besley and Persson (2009); Abed and Gupta (2002); Ahmad and Ajaz (2010); Bird (2008); Gupta (2007); Le, Moreno-Dodson, and Bayraktar (2012); Keen (2012); Besley and Persson (2013).

institutions hinder fiscal reforms and tax efforts and have tend to trap many countries in a low equilibrium tax revenue outcome. More generally the literature (Acemoglu et al., 2001; North, 1990) argued that strengthening laws and regulations with effective enforcement and impartial system of governance will support investment and innovation, which provide a conducive environment to increased tax revenue and economic growth.

9. **The two hypotheses would have different policy implications, calling for careful scrutiny**. If Hypothesis 1 (tax capacity→institutions) holds, countries (especially with poor institutions and low tax capacity) should first work on reform measures that will produce immediate short-term rise in tax capacity (Akitoby et al., 2019).[5] Such revenue enhancing efforts would help improve institutions. Under Hypothesis 2 (institutions→tax capacity), however, it is essential to focus on strengthening institutions, to move to a new higher equilibrium tax capacity position. This difference in the policy implications calls for careful scrutiny for the causality between tax capacity and institutions. To embark on such causality analysis, the next section explores some stylized facts with description of data used in the analysis.

III. STYLIZED FACTS

10. **Throughout the paper, the analyses use panel data for 110 non-resource rich countries covering the period from 1996 to 2017**.[6] After examining the entire panel (all countries), the panel is further disaggregated into LICs (23 countries), lower middle-income countries (LMICs) (29 countries), upper middle-income countries (UMICs) (27 countries), and high income countries (HICs) (31 countries).[7] In light of the existence of "*a minimum tax to GDP ratio (a tax tipping point) associated with a significant acceleration in the process of growth and development*" (Gaspar et al., 2016a), the panel is also divided into countries below (18 countries) and above (92 countries) the tax tipping point (estimated at 12¾ percent of GDP). This also helps in explicitly detecting whether countries with high tax capacity could achieve better institutions and/or if better institutions could enable a country to reach at least the tipping point.

11. **Three institutional indicators and one composite index are used as measures of institutions.** The three indicators (government effectiveness, the control of corruption, and the freedom to trade) are selected, given they are more relevant for tax capacity than other 10 governance indicators that had been tested.[8] A composite institutional index is constructed using

[5] These measures include simplifying the tax system; curbing exemptions; reforming indirect taxes on goods and services (e.g., excises); and better managing compliance risks through strengthening taxpayer segmentation (often beginning with strengthening the large taxpayer office (LTO).

[6] See Appendix 1 for detailed explanation of data used in the paper and its sources.

[7] See Table A1.1 (Appendix 1) for the lists of countries in these panels.

[8] These selected institutional variables are expected to have higher impact on tax capacity. For instance, an improvement in corruption, the quality of public and civil service, policy formulation and implementation (government effectiveness), and better access to trade international (freedom to trade) are expected to directly improve revenue institutions (public financial management and tax administration) which will eventually impact on tax capacity.

principal component analysis (PCA) with 13 institutional indicators (including the three selected indicators).[9] In the absence of actual indicator of institutions, these indicators (worldwide governance indices, Fraser institute's economic freedom indices and transparency international corruption perception indices) have been widely used in the literature as a measure of governance and institutional quality (Kurtz and Schrank, 2007; Alonso and Garcimartin, 2013; Law et al, 2013; Goes, 2016).[10]

12. **The absence of a time-series data on revenue-specific institutions unfortunately does not allow a causality test between tax capacity and such institutions.** The cross-sectional performance of the revenue-specific institutions tends to be highly correlated with that of state institutions (Box 1). Given this, state institutions can be used as proxies for revenue-specific institutions, and the causality analysis between tax capacity and state institutions is relevant.

13. **Looking at the evolution of tax capacity and institutions of all income groups, we identify the following stylized facts** (Figure 2 and 3):

(1) *Indicators of tax capacity and institutions are highly correlated (Figure 2)*. In general, countries with higher tax capacity tend to have stronger institutions, for all institution indicators. While there is significant variance among countries, they are still highly correlated. The correlation between tax capacity and government effectiveness indicators, for instance, is 0.62, the correlation for corruption perception indexes is 0.65, the correlation for freedom to trade internationally is 0.52 while correlation for derived index is 0.60.

(2) *There is high heterogeneity by income groups*. HICs tend to collect larger tax revenues (in terms of GDP) with better institutions. For LMICs, and UMICs, the correlations are weaker. Also, some LICs, LMICs, and UMICs tend to collect above the tax tipping point with poor institutional quality.[11]

(3) *Countries with very low tax capacity have weak institutions, without exception (Figure 3)*. The figures on corruption perception, government effectiveness, and the derived overall institutions index indicate that low tax capacity—apparently below the tipping point (12.75 percent of GDP)—is always associated with weak institutions. In terms of the variances of institution indicators, there is a striking difference among countries above and below the tipping point. Should low tax capacity countries always suffer from weak institutions, those countries may not be able to achieve higher growth and development, as identified in Gaspar et al (2016a).

[9] The PCA have been widely applied in economic literature (see Harris, 1997; Bai, 1993; Caudill et al 2000) for aggregating data scattered in many numeric measures. See Appendix 2 for detailed explanation of how the overall institutional index is constructed.

[10] Detailed definitions of these indicators are available on their database described above.

[11] Detailed charts on individual country groups are available on request.

Box 1. General State Institutions and Revenue-Specific Institutions

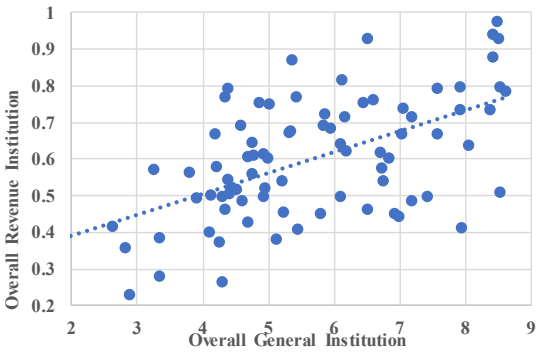
This Box examines the relationship between general state institutions and revenue-specific institutions. For this purpose, the paper uses the IMF's Revenue Administration Fiscal Information Tool (RA-FIT) database, which contains relevant institutional indicators related to tax administration.[1] These are related to risk management, public accountability, third party information, and autonomy. These indicators are averaged, with equal weights, to derive an overall revenue institution index.[2] In light of the lack of historical RA-FIT data, this paper only undertakes a simple cross-section regression and correlation.

Table 1. Cross-Section Regression of the Impact of State Institution on Revenue Institution

Independent variables: general institutions	Dependent variables: revenue institutions
	Overall revenue institution
Government effectiveness	0.43 (6.73***)
Constant	-0.95 (-14.52***)
R-squared	*0.36*
Corruption perception	0.35 (4.58***)
Constant	-1.87 (-6.43***)
R-squared	*0.21*
Freedom to Trade	0.79 (4.70***)
Constant	-2.09 (-6.30***)
R-squared	*0.22*
Overall general institution	0.58 (6.10***)
Constant	-1.53 (-9.35***)
R-squared	*0.32*
No. of Observation	*110*

Note: *, ** and *** indicate the statistical significance at 10%, 5% and 1% level respectively. All variables are in natural logarithms and *t-Statistics* are in parenthesis

Figure 1. Cross-Section Correlation Between Revenue and State Institutions



The cross-section analyses show positive and statistically significant results (Table 1 and Figure 1). They suggest that the performance of general state institutions—using measures for government effectiveness, the control of corruption, freedom to trade, rule of law etc.—and revenue-specific institutions is highly and positively correlated, implying that countries with good state institutions tend to have good revenue institutions as well.

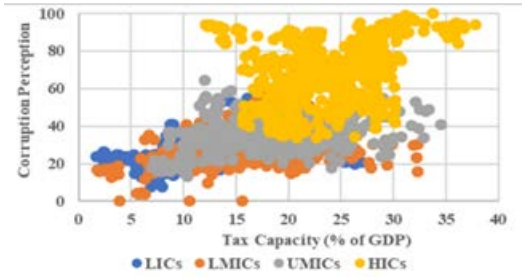
---

[1] The RA-FIT data portal is developed by the IMF to disseminate data gathered using the RA-FIT data collection platform. More details about RA-FIT and access to data is available at: https://data.rafit.org/

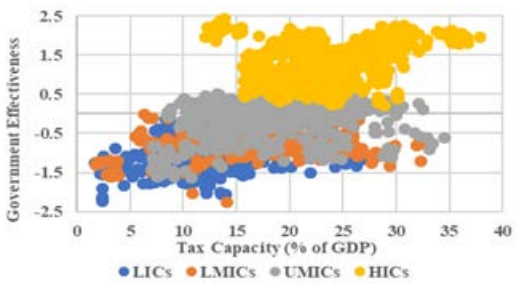
[2] The overall revenue institution index is constructed following IMF (2019). Specific indicators (used from the RA-FIT dataset) include: key compliance risk; automated risk profiling for payments and returns; use of electronic payment methods; formal sets of service delivery standard made public and published; strategic plan prepared and made public; annual business plan prepared and made public; annual report prepared and made public; use of external auditors and formal internal assurance mechanism; enterprise-wide risk policy; existence of code of conduct; provision of e-services - integrated taxpayer accounts, online applications for taxpayers, electronic invoicing system; and use of computer-based information systems for processing data received in respect of employer wage and salary information, financial institutions, other government agencies, international exchange, insurance companies, online trading, assets leasing and VAT invoicing.

Figure 2. Tax Capacity and Institution by Income Group

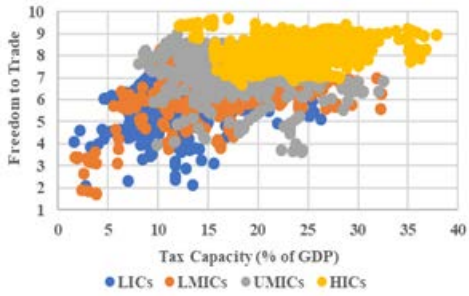
Tax Capacity and Corruption Perception



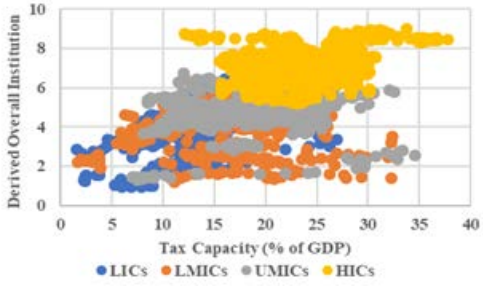
Tax Capacity and Government Effectiveness



Tax Capacity and Freedom to Trade



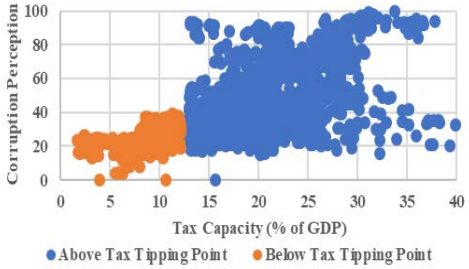
Tax Capacity and Overall Institution Index



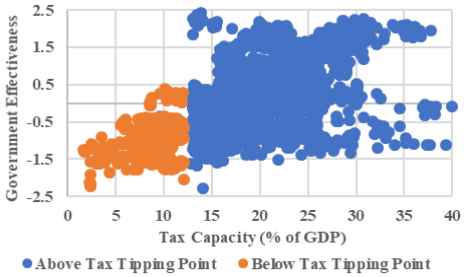
Sources: IMF Database and World Bank Databank

Figure 3. Tax Capacity and Institution of Countries Above/Below Tax Tipping Point

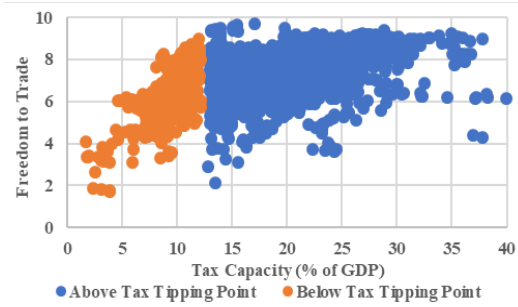
Tax Capacity and Corruption Perception



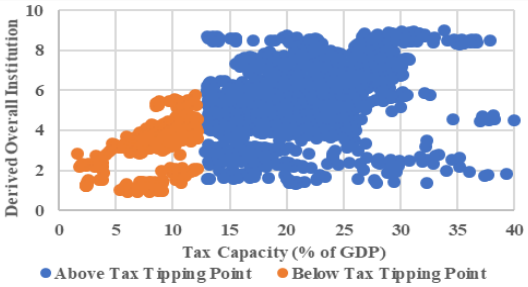
Tax Capacity and Government Effectiveness



Tax Capacity and Freedom to Trade



Tax Capacity and Overall Institution Index



Sources: IMF Database and World Bank Databank

14. **Table 2 shows several notable features reflecting expected variations in tax capacity and state institutions.** The computed standard deviation (across time and cross-sections) are significantly different from each other, suggesting that much of the variations in both tax capacity and state institutions are cross-country. Variations in tax capacity are also significantly larger than state institutions. There is a clear variation across country groups in tax capacity, with LICs and HICs having the lowest cross-section and time variations respectively. State institutions remains highly persistent over time and with not much cross-country variations among country groups. Looking at countries above and below the tax tipping point, the variations suggest that many of the LICs are still operating below the tax threshold and low state institutions.

Table 2. Average Cross-Section and Time Variations (Standard Deviation) in Tax Capacity and Institutions (1996-2017)

Country Groups	Tax Capacity		State Institutions (Index)	
	Time Variations	Cross-Section Variations	Time Variations	Cross-Section Variations
All	2.03	7.50	0.25	1.82
LICs	2.52	3.89	0.28	0.97
LMICs	2.52	7.42	0.24	1.20
UMICs	1.98	5.79	0.28	1.10
HICs	1.26	6.26	0.22	0.97
AboveTP	2.07	6.82	0.26	1.79
BelowTP	1.86	2.90	0.23	1.10

IV. EMPIRICAL APPROACH

15. **There are two distinct features in the empirical approach adopted for this paper:**

- **The empirical approach tests the long-run causality between institutions and tax capacity, based on a panel vector error correction model initially explored by Hurlin and Venet (2001).** Such a technical approach is needed in the absence of sufficiently long time-series (which would be required for conducting standard Granger causality tests).[12] Another panel data technique designed specifically for testing causality is explored in Dumitrescu and Hurlin (2012). This technique, however, is limited to testing short run and bivariate panel causality. Given the highly persistent nature of institutions variables, the plausibility of a short run causality could be questioned. Thus, the paper focuses on the long run causality rather than the short run causality tests often seen in the recent institutions and growth literature (Wilson, 2016; Goes, 2016).
- **Another distinct feature is to use a multivariate (rather than traditional bivariate) approach to the panel causality test.** In addition to the institutions and tax capacity variables, the level of real GDP per capita is also included in the model as a vector. This is

[12] Specifically, the availability of data-series on state institutions is usually limited. Most institutional indicators are dated from 1995.

motivated by the observation that real GDP per capita is found to be a common factor driving both tax capacity and the level of institutions based on the literature that investigate their determinants (Alonso and Garcimartin, 2013; Gupta, 2007).[13] The inclusion of GDP per capita has also been buttressed by empirical studies that have shown the impact of institutions and tax capacity on economic growth (Knack and Keefer, 1995; Mauro, 1995; Engen and Skinner, 1996; Barro and Redlick, 2011; Fatas and Mihov, 2013) and the establishment of bidirectional causality revealed in recent literature (Law et al, 2013; Abdullah and Moley, 2014; Wilson, 2016; Goes, 2016).

Unit Root and Cointegration Tests

16. **To begin the process, the panel vector autoregressive (PVAR) methodology is first examined ((1)-(3))**. The PVAR representation (as laid out in Hurlin and Venet, 2001[14] and Vidangos, 2009) is first examined to reveal the contemporaneous effects among the variables in the system. The econometric framework considers three covariance variables (institutions, tax capacity and GDP per capita) observed for 22 years and 110 countries as follows: [15]

$$Inst_{it} = \sum_{k=1}^{n} \gamma_e Inst_{it-k} + \sum_{k=1}^{n} \delta_e Taxgdp_{it-k} + \sum_{k=1}^{n} \pi_e GDP_{it-k} + \varepsilon_{eit} \qquad (1)$$

$$Taxgdp_{it} = \sum_{k=1}^{n} \gamma_y Inst_{it-k} + \sum_{k=1}^{n} \delta_y Taxgdp_{it-k} + \sum_{k=1}^{n} \pi_y GDP_{it-k} + \varepsilon_{yit} \qquad (2)$$

$$GDP_{it} = \sum_{k=1}^{n} \gamma_z Inst_{it-k} + \sum_{k=1}^{n} \delta_z Taxgdp_{it-k} + \sum_{k=1}^{n} \pi_z GDP_{it-k} + \varepsilon_{zit} \qquad (3)$$

where $Inst_{it}$ are institutional variables (government effectiveness (GE_{it}), corruption perception (CP_{it}), freedom to trade internationally (FTI_{it}), and the composite institutional index (D_Inst_{it})); $Taxgdp_{it}$ is the tax revenue to GDP ratio (tax capacity); GDP_{it} is the gross domestic product per capita; and $\delta, \gamma \,\&\, \pi$ are various coefficients of institutions, tax capacity and GDP per capita respectively.

[13] Openness (measured usually as sum of exports and imports) is also identified as a common factor driving both variables. This is excluded from the vector given that they are component of GDP.

[14] Also applied in Coondoo and Dinda (2002) and Hoffmann et al (2005).

[15] The analysis is also augmented (for robustness check) with error correction mechanism based on Engle and Granger (1987) testing procedure capturing the fixed effects. Detail of the testing procedure and results are presented in Appendix IV. This procedure was also adopted in Apergis and Payne (2009 and 2010) and when testing the long run causality between energy consumption and economic growth in Central and South America.

- **The paper adopts the homogenous causality hypothesis presented in Hurlin and Venet (2001)**.[16] The presence of homogenous causality assumes that all coefficients are identical for all countries and are different from zero. This implies that the estimated PVAR adopted above does not include fixed effects. Vidangos (2009) justified this by comparing results from models with fixed effects and those without fixed effects and concluded that the latter produce more robust results.[17] Previous studies (MaCurdy, 1982; Meghir and Pistaferri, 2004) also confirmed this and have rejected specifications with fixed effects in such variables.

17. **The individual variables are tested for the presence of a unit root, and the equations (1)-(3) are tested for cointegration relationships.** The Levin, Lin, and Chu (2002) (LLC) and Im, Pesaran, and Shin (2003) (IPS) panel unit root tests were adopted in the paper. The optimal lag lengths were assigned based on the modified Akaike's (AIC), Schwarz's (SC) and Hannan Quinn's (HQ) information criterion.[18] If all variables had unit roots (in at least one unit root test), a cointegration test is required. The paper employs the panel cointegration test from the combined Johansen (1988) and Fisher (1932) developed in Maddala and Wu (1999). The Johansen-Fisher panel cointegration test combines tests from individual cross-sections to obtain test statistics for a full panel.

- **The panel unit root tests have the null hypothesis of unit roots**.[19] Table A3.1 (Appendix 3) presents the results of these unit root tests for tax capacity, GDP per capita, and institutions variables across the four income groups as well as country grouping above and below the tax tipping point. The results reveal evidence of a unit root (in levels) in nearly all the panel series and the null hypothesis of unit roots in most of the group cases could not be rejected. The unit root results also suggest that institutions, tax capacity and GDP per capita are I(1) variables.
- **The tests confirm the presence of a cointegrating relationship for the equations (1), (2), and (3).** Table A3.2 (Appendix 3) presents the cointegrating results. The results from the tests performed detects two cointegrating equations across many of the groupings. This means that two long run equations that can be derived from the cointegrating vectors.[20] The

[16] Hurlin and Venet (2001) present four types of causality hypothesis that may emerge from the panel data: (i) homogenous non-causality, (ii) homogenous causality, (iii) heterogenous causality, and (iv) heterogenous non-causality.

[17] According to Vidangos, introducing fixed effects may produce less robust results given that: (i) it would require estimating the models by generalized method of moments (GMM), which is problematic because the available instruments (i.e. institutions) are likely to be very weak; (ii) for variables with high degree of persistence (i.e. institutions), PVAR without fixed effects can be adopted.

[18] The criteria with the smallest lag lengths is used. Optimum lag length ranges between 1 to 3 for all variables. This lag lengths were also used in the panel VAR and VECM estimations.

[19] Wilson (2016) and Gani (2011) have tested for unit roots in categorical series such as governance indicators.

[20] PVAR cointegration following the Johansen (1988) methodology (with no fixed effects) were also conducted and the conclusion of two cointegrating equations was confirmed across panel groupings.

existence of cointegrating relationships also indicate that there must be causality in at least one direction.

18. **These results call for a dynamic error correction representation to test the causality.** The PVAR discussed above tends to only indicate temporal precedence without any economic support and such precedence may not be a conclusive proof of causality and more importantly a long run causality. In this case, Engle and Granger (1987) suggest that when testing for causality between two non-stationary variables that are cointegrated, then it is necessary to specify a model with a dynamic error correction representation rather than a VAR.

Panel Vector Error Correction Model

19. **As the variables are cointegrated, a panel Vector Error Correction Model (VECM) is used to perform long-run causality tests**. Due to the high persistence in institutions, the plausibility of a short run causality could be questioned. Therefore, the paper took a further step to test the long run causality between institutions and tax capacity with the help of a panel VECM.[21]

20. **Using the panel VECM, a long run causality can be established via the error correction term (ECT)**. This would ensure that a displacement from the equilibrium relation implies a response from one of the variables to attain the equilibrium.[22]

21. **The following panel VECM is estimated**:

$$\Delta Inst_{it} = \alpha_1 + \beta_e ECT_{it-1} + \sum_{k=1}^{n} \gamma_e \Delta Inst_{it-k} + \sum_{k=1}^{n} \delta_e \Delta Taxgdp_{it-k} + \sum_{k=1}^{n} \pi_e \Delta GDP_{it-k} + \varepsilon_{eit} \quad (4)$$

$$\Delta Taxgdp_{it} = \alpha_2 + \beta_y ECT_{it-1} + \sum_{k=1}^{n} \gamma_y \Delta Inst_{it-k} + \sum_{k=1}^{n} \delta_y \Delta Taxgdp_{it-k} + \sum_{k=1}^{n} \pi_y \Delta GDP_{it-k} + \varepsilon_{yit} \quad (5)$$

$$\Delta GDP_{it} = \alpha_3 + \beta_z ECT_{it-1} + \sum_{k=1}^{n} \gamma_z \Delta Inst_{it-k} + \sum_{k=1}^{n} \delta_z \Delta Taxgdp_{it-k} + \sum_{k=1}^{n} \pi_z \Delta GDP_{it-k} + \varepsilon_{zit} \quad (6)$$

$$\beta_e,\ \beta_y\ \&\ \beta_z\ establish\ long\ run\ causation$$

$$\delta_e,\ \gamma_y\ \&,\ \gamma_z\ establish\ short\ run\ causation$$

[21] The panel VECM distinguishes between a long run and short run relationship among variables and identifies sources of causation that cannot be detected by the usual Granger causality test (Oh and Lee, 2004; Asafu-Adjaye, 2000).

[22] Granger causality test establishes a short run causation through the lagged values of the dependent and independent variables. The long run causality on the other hand, depends on the significance of the long run relationship which is tested through the lagged ECT derived from the long run equilibrium relationship.

And the error correction terms are defined as:

$$ECT_{it-1}(eq4) = \alpha_1 + Inst_{it-1} - \sigma_{1e}GDP_{it-1} - \sigma_{2e}Taxgdp_{it-1}$$

$$ECT_{it-1}(eq5) = \alpha_2 + Taxgdp_{it-1} - \sigma_{1y}GDP_{it-1} - \sigma_{3y}Inst_{it-1}$$

$$ECT_{it-1}(eq6) = \alpha_3 + GDP_{it-1} - \sigma_{3z}Inst_{it-1} - \sigma_{2z}Taxgdp_{it-1}$$

where ECT_{it-1} are the lagged error-correction terms derived from the long run cointegrating relationship; the ε_{it} are the error-correction terms assumed to be uncorrelated and random with mean zero; the σ_s are the long run coefficients from cointegrating vectors. The coefficients $\beta\ (e, y, z)$ of the *ECTs* represent the deviation of the dependent variables from the long run equilibrium (dynamic adjustment towards the long run equilibrium path). As assumed in Hurlin and Venet (2001), the autoregressive coefficients γ_e and the slope coefficients δ_e and π_e are constant $\forall k\ \in\ [1, n]$. It is also assumed that γ_e are identical for all individual units (countries), whereas the slope coefficients δ_e and π_e could have an individual dimension. The GDP per capita equation is included to reconfirm the established relationship in the literature between institutions and growth (Acemoglu and Robinson, 2008). The test-statistics of the panel VECM follows a Chi-square distribution and is derived by estimating a system of equation from the panel VECM.

22. **Through the *ECT*, a VECM tend to offer an alternative test of causality by testing the following null hypotheses**:

$$H_0^1: \beta_e = 0, and\ \ H_0^2: \beta_y = 0, and\ \ H_0^3: \beta_z = 0.$$

From testing hypothesis 1 and 2, we have four possible testing results:

1) If Hypothesis H_0^1 is rejected but Hypothesis H_0^2 is accepted, then there exists causality running unidirectionally from tax capacity to institutions.
2) If Hypothesis H_0^1 is accepted but Hypothesis H_0^2 is rejected, then there exists causality running unidirectionally from institutions to tax capacity.
3) If both Hypotheses H_0^1 and H_0^2 are accepted, there is no long run causal relationship between institutions and tax capacity. Therefore, institutions and tax capacity do not respond to deviation in long run equilibrium in period $t-1$.
4) If both Hypotheses from H_0^1 and H_0^2 are rejected, then there exist a feedback (bidirectional) causal relationship between institutions and tax capacity. Therefore, institutions and tax capacity respond to deviation in long run equilibrium in period $t-1$.

Therefore, if the variables $(Inst_{it}, Taxgdp_{it}, and\ \ GDP_{it})$ are cointegrated then it is expected that at least one or all of the *ECTs* should be significantly non-zero.

23. **This paper focuses on the coefficients of the *ECTs* (long run causality)**. Given that institutions are slow-moving, it may not be plausible to detect the response of institutions to

changes in any of the variables in the short run. Therefore, long run causality of the dependent variables is tested by the simple t-test of the *ECT* coefficients.

V. EMPIRICAL RESULTS

A. Panel-Causality Tests

24. **A panel VECM is estimated**. Using the detected two cointegrating equations, one can only identify either Equation 4 and 5; 4 and 6 or 5 and 6 at the same time. Given that equation 4 and 5 are the equation of interest of this paper, four restrictions are imposed to fully identify all cointegrating vectors.[23] The following restrictions were imposed on the cointegrating vector of equation 4 and 5:

$$\tau X_{it-1} = \beta\sigma' X_{it-1} = \begin{bmatrix} 0 & 0 \\ \beta_{2y} & 0 \\ 0 & \beta_{3e} \end{bmatrix} \begin{bmatrix} \sigma_{1y} & 1 & \sigma_{3y} \\ \sigma_{1e} & \sigma_{2e} & 1 \end{bmatrix} \begin{bmatrix} gdp_{it-1} \\ Taxgdp_{it-1} \\ Inst_{it-1} \end{bmatrix} \quad (7)$$

Equation 7 shows the long run part of Equation 4 and 5. The long run coefficients (σ_{2y} and σ_{3e}) for tax capacity and institutions respectively are normalized to 1 since they are the equation of interest. The coefficients of the ECTs (β) represent the rate at which both tax capacity and institutions will adjust back to long run equilibrium due to any shocks from tax capacity, institutions and GDP per capita. However, the restricted ECTs (β) do not represent the specified causality relationships of changes in institutions causing changes in tax capacity and vice-versa. For instance, β_{3y} & $\beta_{2e} = 0$ are the coefficients of the ECTs showing deviations of institutions and tax capacity respectively from their own long run equilibrium while $\beta_{1e} = 0$ is the coefficient of the ECT showing deviation of tax capacity from long run equilibrium of GDP per capita. β_{2y} represent the causality running from institutions to tax capacity, β_{3e} represent the causality running from tax capacity to institutions and $\beta_{1y} = 0$ represent the causality running from institutions to GDP per capita which will also be captured when equation 4 and 6 are identified as cointegrating vectors. [24]

25. **The panel-causality test results are supportive of Hypothesis 4 above, indicating a bidirectional causality between institutions and tax capacity**. Table 3 presents the long run multivariate panel causality test results with the *t*-statistics for the coefficients of the *ECT*s.[25] The results (across all panels) reveal a bidirectional causality between tax capacity and all four institutional variables, rejecting the null hypothesis that tax capacity does not cause institutions

[23] To identify all cointegrating vectors, Enders (2015: 361) suggests imposing at least $n \times n$ restriction(s) on cointegrating vectors, where n is the number of cointegrating equations.

[24] Similar restrictions were imposed on Equation 4 and 6 to identify the relationship between institutions and GDP per capita.

[25] The paper reports the t-statistics (rather than coefficients) since the paper focused on determining causation and not impact. The signs (+ or -) of the coefficients are not interpreted either.

and vice-versa. This means that changes in institutional variables cause changes in tax capacity, and likewise a change in tax capacity causes institutional variables to change.[26]

26. **Such bidirectional causality is identified across all country groups**. All coefficients are statistically significant at least at 10 percent level. The finding of bidirectional causality is not affected by the level of income (e.g., LICs, LMICs, UMICs, HICs). The results on countries above and below the tax tipping point also suggest that, irrespective of the level of tax capacity and institutions a country operates in, institutions and tax capacity cause each other.[27] Admittedly, there may be a limited revenue gain by a marginal improvement in institutions in strong-institution countries (e.g., HICs), while weak-institution countries may be expected to gain more with an extra effort for institution building.

27. **The bidirectional causality is also supportive of the earlier finding (Box 1) that state institutions are a good reflection of revenue-specific institutions.** The results suggest that changes in state institutions can be expected to lead to changes in tax capacity. This would be plausible if the level of state institutions is highly associated with that of revenue-specific institutions (as analyzed in Box 1). It is consistent with the notions in the literature that taxation is central to building state institutions (Brautigam 2008; Gaspar et al., 2016b), and building state institutions is also required in generating tax revenue for state development (Rodrik et al., 2004; Besley and Persson, 2014).

28. **The results further suggest a sequencing of policy and/or institutional reforms for the mutual reinforcement of tax capacity and institutions to generate meaningful results.** For changes in tax capacity to translate into change in state institutions, it is important for policymakers to consider focusing on revenue-specific set of policy and institutional changes that will obviously affect tax revenue. For instance, starting with efficient tax policies and revenue administration changes (i.e. tax rate, tax exemptions and risk management), tax revenue could be boosted over the short-term (Akitoby et al 2019). However, the proceeds from these short-term changes could be used to implement expensive institutional reforms that will lead to general improvement in state institutions.[28]

[26] These results are not consistent with Prichard and Leonard (2010)'s finding (after the 1990s there is no longer any evidence of a positive relationship between taxation and broader state capacity gains). This likely reflects a different empirical approach. Prichard and Leonard (2010) use fixed effects with five-year lags of tax variables and thus focus more on short-term relations, while this paper focus on long-term causality (for the reason mentioned above).

[27] Note that, the paper does not investigate the existence of a tipping point in tax capacity that will accelerate institutional development. But rather, it uses the tipping point threshold to detect whether tax capacity and institutions will reinforce each other when a country operate either below or above tax tipping point.

[28] It is important to note that the specific composition and sequencing of policy and institutional reforms will require country specific analysis, given that there are cases where specific revenue administration institutional reform initiatives are warranted immediately.

Table 3. Long Run Panel Causality Test Results

Region/dependent variable	Source of causation: Error correction term (ECT)					
	GE_{it}	CP_{it}	FTI_{it}	D_Inst_{it}	$TaxGDP_{it}$	GDP_{it}
All countries						
GE_{it}	--	--	--	--	-8.29***	-7.86***
CP_{it}	--	--	--	--	-9.24***	-6.18***
FTI_{it}	--	--	--	--	-14.82***	-14.93***
D_Inst_{it}	--	--	--	--	-2.19*	-0.76
$Taxgdp_{it}$	-10.15***	-8.98***	-10.53***	-8.41***	--	--
GDP_{it}	-12.00***	-3.51***	-13.06***	-3.52***	--	--
LICs						
GE_{it}	--	--	--	--	-4.53***	-3.98***
CP_{it}	--	--	--	--	-4.48***	-4.66***
FTI_{it}	--	--	--	--	-6.44***	-5.69***
D_Inst_{it}	--	--	--	--	-1.85*	-1.82*
$Taxgdp_{it}$	-6.67***	-4.60***	-4.98***	-5.18***	--	--
GDP_{it}	-3.08***	-4.48***	-5.10***	-4.44***	--	--
LMICs						
GE_{it}	--	--	--	--	-5.56***	-5.33***
CP_{it}	--	--	--	--	-5.37***	-5.57***
FTI_{it}	--	--	--	--	-6.68***	-6.55***
D_Inst_{it}	--	--	--	--	-2.10*	-2.14*
$Taxgdp_{it}$	-4.10***	-3.31***	-4.45***	-3.98***	--	--
GDP_{it}	-7.04***	-6.79***	-5.45***	-7.19***	--	--
UMICs						
GE_{it}	--	--	--	--	-2.46**	-4.03***
CP_{it}	--	--	--	--	-4.90***	-5.68***
FTI_{it}	--	--	--	--	-3.54***	-2.48*
D_Inst_{it}	--	--	--	--	-1.70*	-0.87
$Taxgdp_{it}$	-2.92***	-4.44***	-3.53***	-4.32***	--	--
GDP_{it}	-9.34***	-8.89***	-8.38***	-7.19***	--	--
HICs						
GE_{it}	--	--	--	--	-3.90**	-3.45***
CP_{it}	--	--	--	--	-4.33***	-4.51***
FTI_{it}	--	--	--	--	-2.15**	-4.36***
D_Inst_{it}	--	--	--	--	-2.46**	-2.61**
$Taxgdp_{it}$	-3.22***	-3.19***	-1.85*	-2.68**	--	--
GDP_{it}	-5.95***	-8.37***	-8.01***	-5.55***	--	--
Above TP						
GE_{it}	--	--	--	--	-8.47***	-7.78***
CP_{it}	--	--	--	--	-7.53***	-7.32***
FTI_{it}	--	--	--	--	-13.76***	-5.25***
D_Inst_{it}	--	--	--	--	-1.84*	-0.24
$Taxgdp_{it}$	-8.03***	-8.65***	-7.68***	-8.07***	--	--
GDP_{it}	-10.10***	-11.52***	-9.04***	-2.02**	--	--
Below TP						
GE_{it}	--	--	--	--	-3.55***	-3.17***
CP_{it}	--	--	--	--	-5.77***	-6.05***
FTI_{it}	--	--	--	--	-7.32***	-3.95***
D_Inst_{it}	--	--	--	--	-1.75*	-1.74*
$Taxgdp_{it}$	-4.69***	-7.60***	-7.18***	-5.81***	--	--
GDP_{it}	-3.10***	-7.34***	-4.69***	-5.49***	--	--

Note: *, ** and *** indicate the rejection of null hypothesis of no granger-causality and with statistical significance at 10, 5, and 1 percent level respectively.

29. **The analysis also reveals that the inclusion of GDP per capita (in multivariate analysis) improve the robustness of the empirical results**.[29] GDP per capita played an important role in establishing the dual-causality between institutions and tax capacity. Given that GDP per capita remains a common factor driving institutions and tax capacity, an indirect causal relationship can be envisaged. Therefore, the multivariate analysis of including GDP per capita to testing the causality have established a direct causality between institutions and tax capacity.[30]

30. **The robustness of this empirical analysis is tested in four different forms.** First, given the heterogeneity across countries, the panel is disaggregated based on income groups and the tax tipping point.[31] Second, a composite institutional index is constructed using PCA to confirm that many state institutions indicators will produce similar results. Third, the methodology (panel VECM) is also augmented with a fixed effects error correction mechanism based on Engle and Granger (1987) procedure and the results confirm a bidirectional causality (See Appendix IV). Fourth, to ensure there is no omitted variable bias, political instability is included in the model as the fourth variable and the results confirm a bidirectional causality (See Appendix VI). Fifth, considering a possible structural break in the sample period (e.g., Prichard and Leonard, 2010), the results of hypothesis 1 (tax capacity→institutions) are also confirmed with different sample periods (1996-2017).[32]

31. **These results, however, only indicate the direction of causality, not the magnitude of influences between the two variables.** Further analyses are needed to examine the magnitude of such influence.

B. Impulse Response Function

32. **To examine the effects of tax capacity on institutions and of institutions on tax capacity, the paper further explores the model, based on the impulse responses derived from the panel VECM**.[33] Through the dynamic (lag) structure of the panel VECM, a shock to a variable does not only directly affects the variable but is also transmitted to all the other endogenous variables in the system. However, the identification of the shock is based on the

[29] Estimations based on bivariate causality (with only tax capacity and institutions) are presented in Appendix V. The estimations show a mixed and inconsistent results of the direction of causality across all panels.

[30] To buttress the validity of the results, we tested the causality between GDP and institutions. The results (across all panels) reveal a bidirectional causality between GDP and institutional variables, in line with the results of Acemoglu and Robinson, 2008.

[31] There was no justification for robustness test with a different sample size since there has not been any structural break, and institutions are persistent.

[32] Prichard and Leonard (2010)—while identifying a positive relationship between taxation and broader improvements in state capacity—find a significant structural break in the data during the 1990s, after which there is no longer any evidence of a positive relationship between taxation and broader state capacity gains.

[33] While impulse response functions are mostly applicable to VAR models, they are also useful in a VECM in detecting the sign of the interaction between two variables.

short and long run restrictions imposed on the model. Figure 4 and 5 show the results of the impulse responses of tax capacity and institutions to a one percent positive shock in the levels of institutions and tax capacity respectively.

33. **The effect of the shock to the model is permanent.** Given that the variables are nonstationary [I(1)], the impulses generated from the cointegrated VECM, have permanent effects (Lutkepohl, 2005; Gonzalo and Ng, 2001).[34] In general, the impulse responses reveal positive and lasting effects, reaching its peak after two to three years and thereafter remain positive at a constant, increasing or decreasing rate. This indicates that short-term tax policy and or revenue administration reform measures serve as catalyst to achieving higher long-term improvement in tax capacity and institutions.

34. **The results for the full sample indicate positive effects of both shocks (Figure 4).** A one percent positive shock of institutions has a positive effect on tax capacity, while the impacts of the same shock of tax capacity on institutions are also positive. While the long run causality tests confirm bi-directional causality between institutions and tax capacity, the impulse response analyses reveal a different magnitude of the responses. For instance, a 1 percent positive shock to institutions will lead to a permanent increase in tax capacity of about 0.8 percent after three years. Whilst a 1 percent positive shock to tax capacity will lead to a permanent increase in institutions of about 0.3 percent after three years.

35. **A positive shock of the same magnitude will have different implications for actual tax capacity and institutions across countries groups.** For example, a 10 percent improvement in the government effectiveness indicator (one measure of institutions) will lead to an average increase of about 0.1, 0.05, 0.02, and 0.14 index point in the overall institutions index in LICs, LMICs, UMICs, and HICs respectively.[35] Whilst 1 percent improvement in tax capacity implies an average increase in tax revenue of about 0.12 percent of GDP for LICs, 0.17 percent of GDP for LMICs, 0.2 percent of GDP for UMICs and 0.25 percent of GDP for HICs.[36]

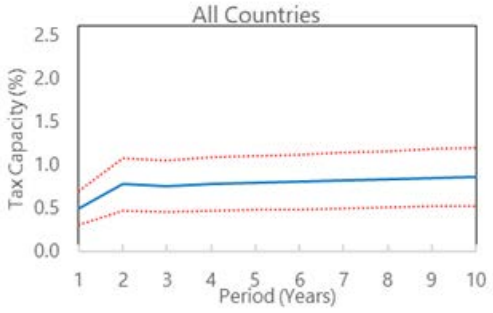
[34] Contrary to VAR (where variables are stationary), the existence of nonstationary variables in the VECM will allow shocks to the variables to have permanent effects and may not return to their initial values even if no further shocks occur (Lutkepohl and Reimers, 1992)

[35] Other things equal, for a typical LICs to move from a poor institutional rating (25th percentile) to a median rating (50th percentile) in the full sample will require an increase in the government effectiveness index of about 150 percent (1.5 index point).

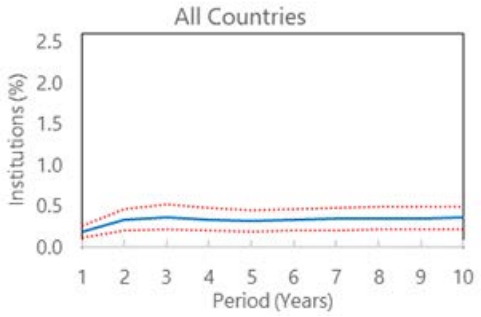
[36] The average annual growth rate in institutions and tax capacity for the full sample of countries over the period 1996 to 2017 is about 0.4 percent and 1.35 percent respectively. See Table A1.2 (Appendix I) for details on other country groups.

Figure 4. Impulse Response Analysis Results for All Sample Countries

Response of Tax Capacity to Shock in institutions



Response of Institutions to Shock in Tax Capacity

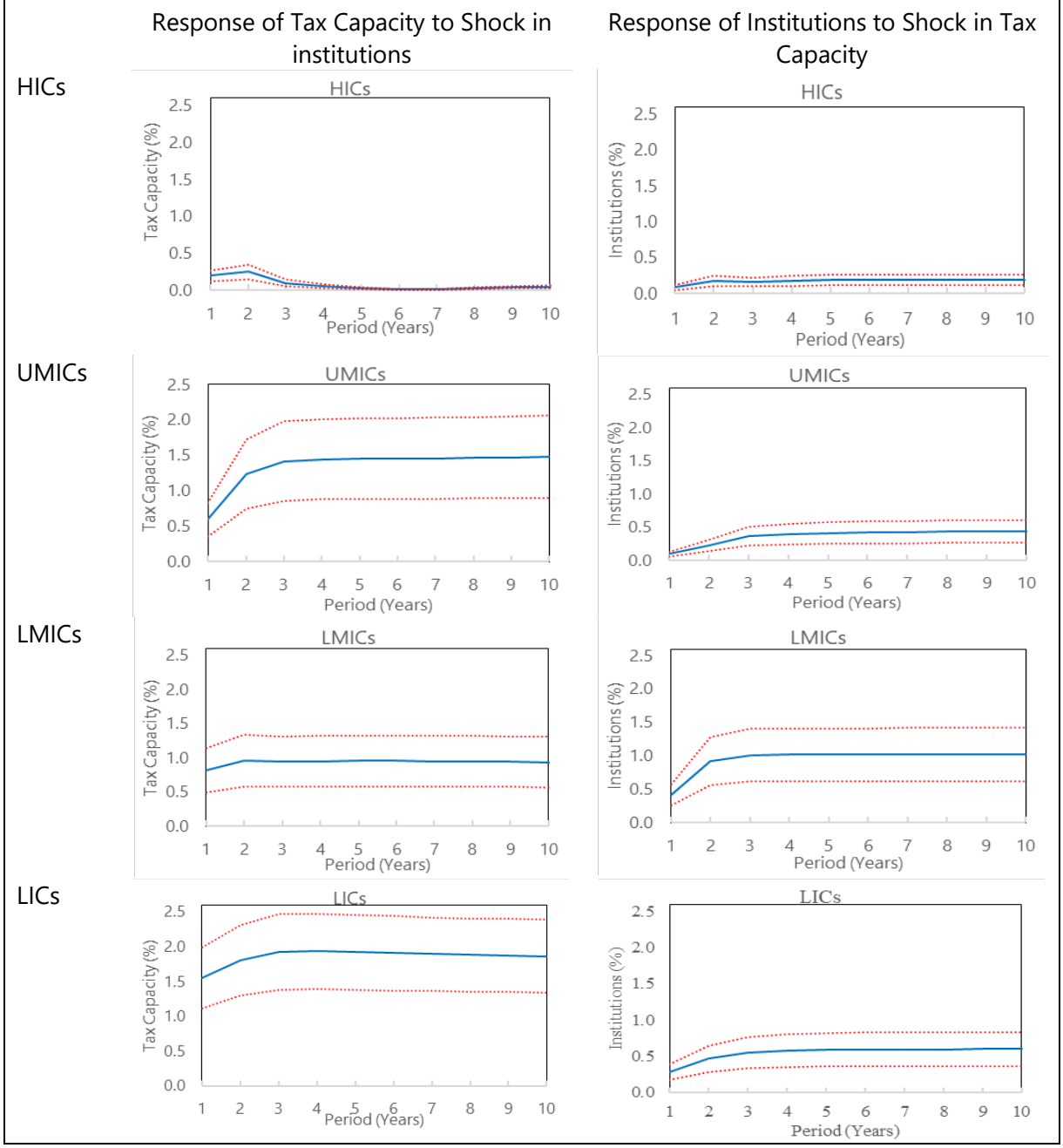


Note: Dotted lines indicate a 95 percent confidence interval

36. **The results by income group, however, reveal high heterogeneity and statistical significance across country groups, with a low impact in HICs (Figure 5).**

- **HICs: Both effects—of a tax capacity shock on institutions and of an institution shock on tax capacity—are limited throughout the period.** These results likely suggest that for countries reaching a certain level of institutions, further improvement in institutions may have limited effects on tax capacity. In the meantime, in countries with already high level of tax capacity, further increase in tax capacity may not necessarily lead to much improvement in institutions.
- **UMICs and LMICs: Larger impacts are revealed from both shocks.** A shock in institutions has a positive effect on tax capacity. In the meantime, a tax capacity shock also has a positive effect on institutions.
- **LICs: Largest response of tax capacity to shock in institutions are revealed, though a shock in tax capacity would have relatively small impacts on institutions**. The responses of tax capacity (to a shock in institutions) are large, suggesting the scope for initial gains through strengthening institutions. This supports the benefits of institution building in these countries. It also implies that, assuming a symmetric impact of the institutions, a slippage in institution building could have significant adverse impacts on tax capacity. In contrast, the effects of a tax capacity shock on institutions are relatively low, in comparison with LMICs.

Figure 5. Impulse Response Analysis Results by Income Group
(With one percent shock)

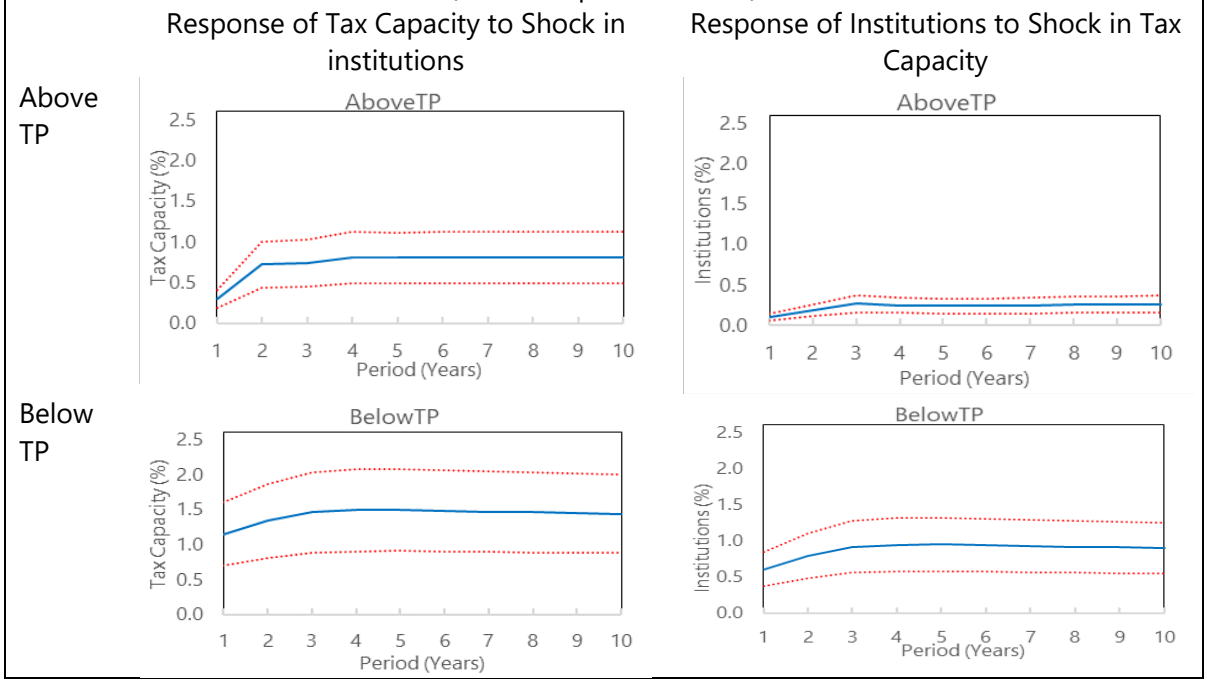


Note: Dotted lines indicate a 95 percent confidence interval

37. **Countries below and above tax tipping point (TP): The impact of both shocks are positive and responses of countries below TP are like LICs (Figure 6)**. In the countries below TP, the impacts of an institution shock on tax capacity is larger than in those above the tipping point. There is similarity between responses of the LICs and below TP countries' tax capacity to shock in institutions and this can be attributed to the fact that most LICs operates at a low tax capacity. This suggest a scope for large gains through strengthening institutions and/or

enhancing tax capacity, which could help develop a virtuous spiral of tax capacity and state institutions. This finding also supports the existence of a tipping point (identified in Gaspar et al., 2016a), implying that the level of tax capacity matters for institutions and thus for growth.

Figure 6. Impulse Response Analysis Results Above and Below Tax Tipping Point
(With one percent shock)



Note: Dotted lines indicate a 95 percent confidence interval

VI. Conclusion and Policy Implications

38. **The paper confirms a long-run bidirectional causality between tax capacity and state institutions across all country groups.** This indicates that tax capacity and institutions are mutually reinforcing, with changes in tax capacity leading to changes in institutional structure and vice versa. It also suggests that building state institutions (even in areas that are not directly related to revenue collections may indirectly help improve tax capacity.

39. **In terms of the causal effects, however, there is high heterogeneity across country groups**. Based on the impulse response analyses, the paper also finds that the causal effects in the HICs are generally low in both directions, while in developing countries, both tax capacity and institution shocks have positive larger and long-run impacts on institutions and tax capacity, respectively.

40. **These findings have important policy implications, particularly in developing countries**:

- **Countries with low tax capacity and weak institutions should do all they could to exit from a vicious cycle**. The existence of the bidirectional causality suggests that countries with

low tax capacity and weak institutions may well be trapped in a vicious cycle, with undeveloped institutions adversely affecting tax capacity, which could further weaken institutions. To exit from such a cycle, starting with short-term tax policy and administrative reforms (e.g., increasing tax rates, removing exemptions, and improving auditing and risk management) may be helpful, as strengthening institutions could often take time. The eventual improvement in institutions could raise the awareness of the citizens or business groups to put more pressure on their governments for enhanced accountability and equitable use of the tax proceeds (Li and Xia, 2008).

- **Further, enhanced efforts on multiple fronts may facilitate a virtuous cycle with enhanced tax capacity improving institutions and better institutions in turn further enhancing tax capacity**. To facilitate a virtuous cycle with higher tax capacity leading to better institutions, tax proceeds should be used efficiently to support the improvement in institutions. Countries with low tax capacity and weak institutions (e.g., LICs, countries operating below the tax revenue tipping point) should strive to develop their institutions and raise tax capacity simultaneously, as they reinforce each other. This can be achieved with capacity building supports by international partners. A Medium-Term Revenue Strategy, involving the coherent and structured reforms of its tax system (tax policy and customs administration, and legal measures), can be a useful guide to enhancing tax capacity on a sustainable basis.

REFERENCES

Abed, G.T., and Gupta, S. eds. (2002) Governance, Corruption, Economic Performance. Washington D.C: International Monetary Fund.

Abdullah, S., and Morley, B. (2014) Environmental taxes and economic growth: Evidence from panel causality tests. Energy Economics, Vol. 42, pp. 27-33.

Acemoglu, D. and Robinson, J. (2008) The role of institutions in growth and development. Commission on Growth and Development Working Paper No. 10.

Acemoglu, D., Johnson, S., & Robinson, J. A. (2001). The colonial origins of comparative development: An empirical investigation. American Economic Review, 91, 1369–1401.

Ahmad, E., and Ajaz, T. (2010) The effect of corruption and governance on tax revenues. Pakistan Development Review 49(4):405–17.

Akitoby, B., Honda, J., Miyamoto, H., Primus, K., and Sy, M. (2019) Case studies in tax revenue mobilization in low-income countries. IMF Working Paper 19/104. Washington DC: International Monetary Fund.

Alonso, J.A., and Garcimartin, C. (2013) The determinants of institutional quality: more on the debate. *Journal of International Development* 25, 206-226.

Apergis, N., and Payne, J.E. (2009) Energy consumption and growth in Central America: evidence from a panel cointegration and error correction model. Energy Economics. 31, 211-216.

Apergis, N., and Payne, J.E. (2010) Energy consumption and growth in South America: evidence from a panel error correction model. Energy Economics. 32, 1421-1426.

Asafu-Adjaye, J. (2000) The relationship between energy consumption, energy prices and economic growth: time series evidence from Asian developing countries. Energy Economics 22: 615-625.

Bai, J. (1993) Inferential theory for factor models of large dimensions. *Econometrica* 71, 135–171.

Barro, R.J., and Redlick, C.J. (2011) Macroeconomic effects from government purchases and taxes. The Quarterly Journal of Economics, Vol. 126, No. 1, pp. 51-102.

Besley, T., and Persson, T. (2014) Why do developing countries tax so little? Journal of Economic Perspectives. Vol. 28, No. 4, pp. 99-120.

Besley, T. and T. Persson, 2013, "Taxation and Development," *Handbook of Public Economics*, Volume 5. Edited by A. J. Auerbach, R. Chetty, M. Feldstein and E. Saez, North Holland.

Besley, T., and Persson, T. (2009) The origins of state capacity: property rights, taxation, and policy. *American Economic Review* Vol. 99, No. 4, pp. 1218–44.

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. Vol. 38 No. 1, pp.55-71.

Bird, R. (2008) Tax challenges facing developing countries. Rotman Working Papers Series No. 12, Rotman Institute for International Business, University of Toronto.

Bornhorst, F., Gupta, S. Thornton, J. (2008) Natural resource endowments, governance, and the domestic revenue effort: evidence from a panel of countries. IMF Working Paper 08/170. Washington DC: International Monetary Fund.

Brautigam, D. (2008) Introduction: Taxation and state building in developing countries. In: Brautigam D, Fjeldstad O-H and Moore M (eds) Taxation and State Building in Developing Countries: Capacity and Consent. Cambridge: Cambridge University Press.

Burgess, R. and Stern, N. (1993) Taxation and Development. *Journal of Economic Literature*. Vol.31, No. 2, pp.762-830.

Caudill, S. B., Zanella, F. C. & Mixon, F. G. (2000) Is economic freedom one dimension? a factor analysis of some common measures of economic freedom. *Journal of Economic Development* 25, 17–40.

Coondoo, D. and Dinda S. (2002). Causality between income and emission: a country group-specific econometric analysis. *Ecological Economics* 40: 351–367.

Dumitrescu, E, & Hurlin, C, (2012) Testing for granger non-causality in heterogeneous panels. *Economic Modelling* 29, 1450–60.

Enders, W. (2015). Applied Econometric Time Series, 3rd edn. New York: John Wiley & Sons.

Engen, E.M., and Skinner, J. (1996) Taxation and economic growth. NBER Working Paper, 5826. Cambridge: National Bureau of Economic Research.

Engle, R.E., and Granger, C.W.J. (1987) Cointegration and error-correction: representation, estimation, and testing. *Econometrica*. 55, 251-276.

Fatas, A., and Mihov, I. (2013) Policy volatility, institutions, and economic growth. *The Review of Economics and Statistics,* 95(2): 362–376.

Fisher, RA, (1932) Statistical methods for research workers. 4th Edition, Oliver & Boyd, Edinburgh.

Fjeldstad, O., and Moore, M. (2008) Tax reform and state-building in a globalized world. In: Brautigam D, Fjeldstad O-H and Moore M (eds) Taxation and State Building in Developing Countries: Capacity and Consent. Cambridge: Cambridge University Press.

Gadenne, L. (2017) Tax Me, but Spend Wisely? Sources of Public Finance and Government Accountability. *American Economic Journal: Applied Economics.* 9(1): 274–314.

Gani, A. (2011) Governance and growth in developing countries. *Journal of Economic Issues*. 45:1, 19-40.

Gaspar, V., Amaglobeli, D., Garcia-Escribano, M., Prady, D. and Soto, M. (2019) Fiscal policy and development: human, social, and physical investment for the SDGs. IMF Staff Discussion Note 19/03. Washington DC: International Monetary Fund.

Gaspar, V., Jaramillo, L. and Wingender, P. (2016a) Tax capacity and growth: is there a tipping point? IMF Working Paper 16/234. Washington DC: International Monetary Fund.

Gaspar, V., Jaramillo, L. and Wingender, P. (2016b) Political institutions, state building, and tax capacity: crossing the tipping point. IMF Working Paper 16/233. Washington DC: International Monetary Fund.

Ghura, D. (1998) Tax revenue in sub-Saharan Africa: effects of economic policies and corruption. IMF Working Paper 98/135. Washington DC: International Monetary Fund.

Goes, C., (2016) Institutions and growth: a GMM/IV panel VAR approach. Economic Letters. 138, 85-91.

Gonzalo, J. and Ng, S. (2001) A systematic framework for analyzing the dynamic effects of permanent and transitory shocks. Journal of Economic Dynamics & Control, 25, 1527-1546.

Gupta, A.S. (2007) Determinants of tax revenue efforts in developing countries. IMF Working Paper 07/184. Washington DC: International Monetary Fund.

Harris, D. (1997) Principal component analysis of cointegrated time series. *Econometric Theory 13, 529–557*.

Hurlin C, Venet B. (2001) Granger causality tests in panel data models with fixed coefficients. Working Paper. EURIsCO, Universit´e Paris IX Dauphin.

Im, KS, Pesaran, MH, and Shin, Y, (2003) Testing for unit roots in heterogeneous panels. *Journal of Econometrics* 115, 53–74.

International Monetary Fund, (2012) Macroeconomic policy frameworks for Resource-rich developing countries. Available at: https://www.imf.org/external/np/pp/eng/2012/082412.pdf

_____, (2019) Curbing corruption. Fiscal Monitor Chapter 2. International Monetary Fund, Washington DC.

Johansen, S, (1988) Statistical analysis of cointegration vectors. *Journal of Economic Dynamic and Control* 12, 231–54.

Kaufmann, D., Kraay, A. and Zoido-Lobaton, P. (1999). Aggregating Governance Indicators. World Bank Working Paper No. 2195.

Keen, M. (2012) Taxation and Development-Again. IMF Working Paper 12/220. Washington DC: International Monetary Fund.

Knack, S. and Keefer, P. (1995) Institutions and economic performance: institutional measures cross-country tests using alternative. Economics and Politics, Vol. 7, No. 3, pp. 207-227.

Kurtz, M.J., & Schrank, A. (2007). Growth and governance: models, measures, and mechanisms. *Journal of Politics*, 69, 538–554.

Law, S.H., Lim T.C., and Ismail N.W. (2013) Institutions and economic development: A Granger causality analysis of panel data evidence. *Economic System*, 37, 610-624.

Le, T. M., Moreno-Dodson, B. and Bayraktar, N. (2012) Tax capacity and tax effort: extended cross-country analysis from 1994 to 2009. World Bank Policy Research Working Paper No. 6252. Washington, DC: World Bank.

Levin, A, Lin, CF, & Chu, CSJ, (2002) Unit root test in panel data: asymptotic and finite sample properties. *Journal of Econometrics* 108, 1–24.

Li, S., & Xia, J. (2008). The roles and performance of state firms and nonstate firms in China's economic transition. *World Development*, 36, 39–54.

Lutkepohl, H. (2005). New Introduction to Multiple Time Series Analysis. Berlin, Heidelberg: Springer Berlin Heidelberg.

Lutkepohl, H., and Reimers, H. (1992). Impulse response analysis of cointegrated systems. Journal of Economic Dynamics and Control, 16(1), pp. 53-78.

MaCurdy, T.E. (1982) The use of time series processes to model the error structure of earnings in a longitudinal data analysis. Journal of Econometrics, 18, 83-114.

Maddala, GS, and Wu, S, (1999) A comparative study of unit root tests with panel data and a new simple test. *Oxford Bulletin of Economics and Statistics*, Special Issue, 631–652.

Meghir, C. and Pistaferri, L. (2004) Income variance dynamics and heterogeneity. Econometrica, 72(1), 1-32.

Martin, L. (2013) Taxation and accountability: experimental evidence for taxation's effect on citizen behavior. Yale University. Available at: http://cega.berkeley.edu/assets/cega_events/55/Martin_WGAPE-WB.pdf.

Mauro, P. (1995) Corruption and growth. The Quarterly Journal of Economics, Vol. 110, No. 3, pp. 681-712.

North, D.C. (1990) Institutions, Institutional Change, and Economic Performance. New York: Cambridge University Press.

Oh, W. and Lee, K. (2004) Causal relationship between energy consumption and GDP revisited: the case of Korea 1970-1999. Energy Economics 26: 51-59.

Prichard, W and Leonard, D.K. (2010) Does reliance on tax revenue build state capacity in sub-Saharan Africa? International Review of Administrative Sciences. Vol. 76, No.4, pp. 653-675.

Rodrik, D., Subramanian, A., and Trebbi, F. (2004) 'Institutions rule: The primacy of institutions over geography and integration in economic development', *Journal of Economic Growth, Vol.9, pp.131-165*.

Smith, L.I. (2002) 'A tutorial on principal components analysis', available at: http://www.cs.otago.ac.nz/cosc453/student_tutorials/principal_components.pdf.

Vidangos, I. (2009) Fluctuations in individual labor income: a panel VAR analysis. Federal Reserve Board Working Paper series. Available at: https://www.federalreserve.gov/pubs/feds/2009/200909/200909pap.pdf

Wilson, R. (2016) Does governance cause growth? evidence from China. *World Development*. Vol. 79, pp.138-151.

APPENDIX

Appendix I. Data Summary and Sources

This paper uses the data from: IMF (World Economic Outlook database for tax revenue data), the World Bank (World Development Indicators for GDP and governance data), Transparency International database for corruption perception index, and the Fraser Institute database for economic freedom index. An unbalance panel estimation technique is adopted and GDP per capita was measured in real terms (2010 prices) United States dollars, converted at market exchange rates. The analysis in this paper is carried out on the premise that the level of income matters and the income groups a country belongs to is important. Therefore, the entire panel datasets were divided based on income groups following the 2016 World Bank Atlas method.

Inferences from disaggregated income grouping suggest further scrutiny of the panel. There exist some variations in the income groups especially on tax capacity and that could have effect on the causality results. Given this, the entire panel is divided into two other groups, namely; countries below the so-called tax tipping point and countries above the tax tipping point.

Non-resource rich countries are selected to exclude resource-related tax revenue which could affect the robustness of the results. The selection of resource-rich countries follows IMF (2012) criteria which is to have on average either natural resource revenue or exports of at least 20 percent of total fiscal revenue and exports, respectively.

Table A1.1: List of Countries

High-Income Countries			
Austria	Greece	Poland	Uruguay
Barbados	Hungary	Portugal	
Belgium	Ireland	Singapore	
Cyprus	Israel	Slovak Republic	
Czech Republic	Italy	Slovenia	
Denmark	Japan	Spain	
Estonia	Luxembourg	Sweden	
Finland	Malta	Switzerland	
France	Netherlands	United Kingdom	
Germany	New Zealand	United States	

Upper-Middle-Income Countries		
Argentina	Gabon	Romania
Belarus	Guyana	Serbia
Bosnia and Herzegov	Jamaica	South Africa
Brazil	Lebanon	Thailand
Bulgaria	Mauritius	Turkey
China	Montenegro	Turkmenistan
Colombia	Namibia	Tuvalu
Costa Rica	Panama	
Croatia	Paraguay	
Dominican Republic	Peru	

Lower-Middle-Income Countries		
Armenia	India	Philippines
Bangladesh	Jordan	Sao Tome and Principe
Cabo Verde	Kenya	Solomon Islands
Djibouti	Kiribati	Tajikistan
Egypt, Arab Re	Kyrgyz Republic	Tunisia
El Salvador	Lesotho	Ukraine
Georgia	Morocco	Uzbekistan
Ghana	Myanmar	Vanuatu
Guatemala	Nicaragua	Vietnam
Honduras	Pakistan	

Low-Income Countries		
Afghanistan	Haiti	Togo
Benin	Liberia	Uganda
Burkina Faso	Madagascar	Zimbabwe
Burundi	Malawi	
Central African Republic	Mozambique	
Comoros	Nepal	
Eritrea	Rwanda	
Ethiopia	Senegal	
Gambia, The	Sierra Leone	
Guinea-Bissau	Tanzania	

Table A1.2: Average Annual Growth Rates in Tax Capacity and Institutions

Country Groups	Tax Capacity			State Institutions (Index)			Pool Observation
	Mean(%)	Median(%)	Standard Deviation(%)	Mean(%)	Median(%)	Standard Deviation(%)	
All	1.35	0.74	12.21	0.42	0.23	3.54	2,310
LICs	3.40	1.86	23.14	0.30	0.34	3.95	483
LMICs	1.40	0.83	8.92	0.72	0.26	4.75	609
UMICs	0.93	1.07	6.81	0.52	0.39	3.13	567
HICs	0.16	0.24	3.61	0.16	0.06	1.77	651
AboveTP	1.13	0.62	11.57	0.41	0.21	3.10	1,932
BelowTP	2.46	1.81	15.02	0.46	0.40	5.26	378

Note: Values are in percentage change and not percent of GDP.

Appendix II. Deriving Overall Institutional Index

The index was derived using the principal component analysis (PCA) which is a process of transforming high-dimension sets of indicators into new indices that capture information on a different dimension and are mutually uncorrelated (Smith, 2002). The first set of indicators used in the PCA is the average governance indicators which includes corruption control, political stability, government effectiveness, voice and accountability, rule of law, and regulatory quality. These governance indices range from -2.5 to +2.5, with -2.5 representing the worst governance and +2.5 the best governance. The second set of indicators is the average indices for all the economic freedom elements which includes the size of government, legal system and property rights, sound money, freedom to trade internationally, and regulations. The economic freedom indices range from 0 to 10, with 0 representing the least free and 10 the freest. The third set of indicators is the widely-used corruption perception index which range from 0 to 100, with 0 representing the most corrupt and 100 the least corrupt.

To have a compatible set of indicators for the PCA, the three sets needs to be on the same scoring range. The range of 0 to 10 was selected as the common range among the three indicators which is the same as the average economic freedom index. Therefore, to get average governance index to this range, we added +2.5 to all series in the panel and thereafter, multiplied by 2. With regards to corruption perception, we simply divide all series in the panel by 10.

Deriving an aggregated index for institutions, the first eigenvectors (loading matrix) from the PCA were used as the required weights for the three institutional indicators. However, the institutional index was derived from the following linear combination:

$$Inst\_Index = \chi_1 Gi + \chi_2 Efi + \chi_3 Cpi \qquad \text{(A1)}$$

Where χ_1, χ_2, χ_3 are the eigenvectors (weights) from the PCA and Gi, Efi, Cpi are average governance indicators, average economic freedom index, and the corruption perception index respectively. From the PCA estimation, about 35 percent of the weights is attributed to corruption perception index, 34 percent to governance indicators and the remaining 31 percent to economic freedom index.

Appendix III. Unit Root and Cointegration Tests

Table A3.1: Panel unit roots test results

Region / variable	Levin, Lin and Chu (LLC)		Im, Pesaran and Shin (IPS)		Critical Values		
	Levels	First differences	Levels	First differences		LLC	IPS
All countries							
GE_{it}	-1.14	-35.58	0.07	-23.54	1%	-2.35	-3.84
CP_{it}	0.45	-39.76	-1.18	-26.55	5%	-1.68	-3.24
FTI_{it} $\}Inst_{it}$	-0.91	-35.73	-3.55	-26.21	10%	-1.32	-2.92
D_Inst_{it}	2.69	-34.46	-0.63	-22.05			
$Taxgdp_{it}$	4.41	-38.16	-0.12	-23.65			
GDP_{it}	27.51	-13.31	4.01	-19.76			
LICs							
GE_{it}	-2.75	-14.93	0.85	-10.18	1%	-2.41	-4.12
CP_{it}	0.02	-16.89	0.87	-11.65	5%	-1.74	-3.34
FTI_{it} $\}Inst_{it}$	3.42	-15.44	-3.96	-10.36	10%	-1.38	-2.97
D_Inst_{it}	1.77	-14.00	-0.11	-9.93			
$Taxgdp_{it}$	4.69	-15.18	2.82	-10.88			
GDP_{it}	14.14	-9.82	2.13	-11.73			
LMICs							
GE_{it}	-0.30	-17.59	-0.06	-9.81	1%	-2.38	-4.12
CP_{it}	3.14	-23.52	-1.08	-16.91	5%	-1.71	-3.34
FTI_{it} $\}Inst_{it}$	3.03	-18.49	-4.61	-13.73	10%	-1.35	-2.97
D_Inst_{it}	3.17	-20.53	-1.75	-14.55			
$Taxgdp_{it}$	2.49	-14.76	-0.39	-10.22			
GDP_{it}	15.39	-3.95	4.45	-8.65			
UMICs							
GE_{it}	0.02	-18.12	-0.70	-11.95	1%	-2.38	-4.12
CP_{it}	2.61	-21.57	-1.30	-13.64	5%	-1.71	-3.34
FTI_{it} $\}Inst_{it}$	1.91	-16.71	-1.32	-12.30	10%	-1.35	-2.97
D_Inst_{it}	2.67	-14.63	0.71	-7.81			
$Taxgdp_{it}$	3.70	-15.28	-1.27	-11.79			
GDP_{it}	13.50	-6.65	3.72	-9.39			
HICs							
GE_{it}	-0.77	-20.24	0.12	-15.00	1%	-2.38	-3.97
CP_{it}	-1.21	-17.52	-0.70	-10.87	5%	-1.71	-3.27
FTI_{it} $\}Inst_{it}$	-4.34	-21.04	1.22	-17.19	10%	-1.35	-2.93
D_Inst_{it}	0.63	-19.31	-0.05	-11.69			
$Taxgdp_{it}$	0.45	-20.98	-1.07	-14.32			
GDP_{it}	14.59	-11.75	-2.13	-9.95			
Above TP							
GE_{it}	-0.96	-33.23	-0.44	-22.14	1%	-2.36	-3.90
CP_{it}	0.37	-35.28	-1.41	-23.17	5%	-1.69	-3.27
FTI_{it} $\}Inst_{it}$	-2.08	-33.52	-2.42	-24.94	10%	-1.33	-2.93
D_Inst_{it}	2.38	-31.61	-0.81	-20.10			
$Taxgdp_{it}$	3.39	-30.24	-0.95	-21.76			
GDP_{it}	24.69	-13.60	2.57	-17.72			
Below TP							
GE_{it}	-1.31	-12.97	1.20	-8.14	1%	-2.42	-4.35
CP_{it}	0.55	-18.43	0.26	-13.32	5%	-1.74	-3.43
FTI_{it} $\}Inst_{it}$	3.22	-12.62	-3.39	-8.38	10%	-1.38	-3.01
D_Inst_{it}	1.98	-13.77	0.29	-9.05			
$Taxgdp_{it}$	3.60	-14.00	1.89	-9.26			
GDP_{it}	12.95	-3.35	4.06	-8.78			

Source: Author's calculations

Note: All three tests were used to examine the null hypothesis of unit roots. To reject the null hypothesis, t-statistics needs to be less than (more negative) the critical values.

Table A3.2: Panel Cointegration test results

Country /variables	Cointegration rank /1	Trace statistics	5% Critical value	Maximum eigenvalue statistics	5% Critical value
All countries					
GDP, Taxgdp & GE	$r = 0$	166.93	29.80	112.16	21.13
	$r \leq 1$	54.77	15.49	54.42	14.26
GDP, Taxgdp & CP	$r = 0$	190.74	29.80	126.26	21.13
	$r \leq 1$	64.47	15.49	64.34	14.26
GDP, Taxgdp & FTI	$r = 0$	296.97	29.80	195.22	21.13
	$r \leq 1$	101.75	15.49	101.75	14.26
GDP, Taxgdp & D_Inst	$r = 0$	121.72	29.80	98.43	21.13
	$r \leq 1$	23.29	15.49	23.17	14.26
LICs					
GDP, Taxgdp & GE	$r = 0$	82.54	35.19	45.39	22.30
	$r \leq 1$	37.15	20.26	26.93	15.89
GDP, Taxgdp & CP	$r = 0$	80.99	35.19	46.87	22.30
	$r \leq 1$	34.12	20.26	21.05	15.89
GDP, Taxgdp & FTI	$r = 0$	116.98	35.19	61.79	22.30
	$r \leq 1$	55.18	20.26	30.54	15.89
GDP, Taxgdp & D_Inst	$r = 0$	60.06	35.19	35.60	22.30
	$r \leq 1$	24.46	20.26	21.77	15.89
LMICs					
GDP, Taxgdp & GE	$r = 0$	63.62	24.28	39.68	17.80
	$r \leq 1$	23.94	12.32	14.83	11.22
GDP, Taxgdp & CP	$r = 0$	89.66	24.28	50.71	17.80
	$r \leq 1$	38.95	12.32	33.33	11.22
GDP, Taxgdp & FTI	$r = 0$	97.23	24.28	50.60	17.80
	$r \leq 1$	46.63	12.32	34.23	11.22
GDP, Taxgdp & D_Inst	$r = 0$	74.85	24.28	56.38	17.80
	$r \leq 1$	18.47	12.32	14.08	11.22
UMICs					
GDP, Taxgdp & GE	$r = 0$	123.23	35.19	82.94	22.30
	$r \leq 1$	40.28	20.26	36.18	15.89
GDP, Taxgdp & CP	$r = 0$	135.24	35.19	86.00	22.30
	$r \leq 1$	49.24	20.26	42.21	15.89
GDP, Taxgdp & FTI	$r = 0$	120.22	35.19	66.70	22.30
	$r \leq 1$	53.53	20.26	47.94	15.89
GDP, Taxgdp & D_Inst	$r = 0$	82.10	35.19	59.00	22.30
	$r \leq 1$	23.10	20.26	22.51	15.89
HICs					
GDP, Taxgdp & GE	$r = 0$	91.41	35.19	76.64	22.30
	$r \leq 1$	25.77	20.26	20.85	15.89
GDP, Taxgdp & CP	$r = 0$	96.53	35.19	71.32	22.30
	$r \leq 1$	25.21	20.26	20.69	15.89
GDP, Taxgdp & FTI	$r = 0$	121.47	35.19	96.06	22.30
	$r \leq 1$	25.41	20.26	21.23	15.89
GDP, Taxgdp & D_Inst	$r = 0$	88.82	35.19	72.50	22.30
	$r \leq 1$	26.35	20.26	21.56	15.89
Above TP					
GDP, Taxgdp & GE	$r = 0$	135.30	29.80	83.29	21.13
	$r \leq 1$	52.01	15.49	50.47	14.26
GDP, Taxgdp & CP	$r = 0$	130.48	29.80	81.20	21.13
	$r \leq 1$	49.28	15.49	47.84	14.26
GDP, Taxgdp & FTI	$r = 0$	261.11	29.80	188.18	21.13
	$r \leq 1$	72.94	15.49	67.16	14.26
GDP, Taxgdp & D_Inst	$r = 0$	79.52	29.80	70.88	21.13
	$r \leq 1$	18.64	15.49	18.53	14.26
Below TP					
GDP, Taxgdp & GE	$r = 0$	115.37	35.19	92.01	22.30
	$r \leq 1$	23.35	20.26	18.40	15.89
GDP, Taxgdp & CP	$r = 0$	128.95	35.19	84.11	22.30
	$r \leq 1$	44.84	20.26	40.20	15.89
GDP, Taxgdp & FTI	$r = 0$	128.31	35.19	85.80	22.30
	$r \leq 1$	42.51	20.26	38.75	15.89
GDP, Taxgdp & D_Inst	$r = 0$	101.82	35.19	107.84	22.30
	$r \leq 1$	28.40	20.26	32.50	15.89

Source: Author's calculations

/1 Reports the significant cointegrating ranks. The null hypothesis of no cointegration is rejected when the statistics is greater than the critical values.

Appendix IV. Engle-Granger Procedure for Causality Testing

For robustness check, the VECM procedure is augmented with the following Engle-Granger four step procedure –similar approach was also adopted in Apergis and Payne (2009 and 2010)– in determining if two or more variables are cointegrated:

i. Test for stationarity among the variables. The same procedure of testing panel unit root in the study is adopted.

ii. Estimate the long run equilibrium relationship among the variables. A panel OLS fixed-effect model is estimated simultaneously for each variable.

iii. Estimate the error correction model. Using the residual from (ii), an error correction model of Equation 4 and 5 above is estimated. According to Enders (2015), other than the ECT, all procedures developed for a VECM (Equation 4 and 5) is applicable to the system represented by the Engle-Granger error correction model.

iv. Assess the ECT. The ECT is assess in the same manner as the VECM-ECT.

Table A4 present the long run causality test. The results are broadly in line with the results presented in Table 4 above.

Table A4: Engle-Granger Long Run Panel Causality Test Results

Region/dependent variable	Source of causation: Error correction term (ECT)				
	GE_{it}	CP_{it}	FTI_{it}	D_Inst_{it}	$TaxGDP_{i}$
All countries					
GE_{it}	--	--	--	--	-18.27***
CP_{it}	--	--	--	--	-16.29***
FTI_{it}	--	--	--	--	-19.35***
D_Inst_{it}	--	--	--	--	-14.11*
$Taxgdp_{it}$	-16.00***	-15.94***	-17.00***	-16.13***	--
LICs					
GE_{it}	--	--	--	--	-7.55***
CP_{it}	--	--	--	--	-4.62***
FTI_{it}	--	--	--	--	-7.96***
D_Inst_{it}	--	--	--	--	-4.69***
$Taxgdp_{it}$	-7.01***	-4.39***	-6.78***	-6.88***	--
LMICs					
GE_{it}	--	--	--	--	-9.15***
CP_{it}	--	--	--	--	-8.66***
FTI_{it}	--	--	--	--	-11.96***
D_Inst_{it}	--	--	--	--	-9.04***
$Taxgdp_{it}$	-7.47***	-8.02***	-8.19***	-7.69***	--
UMICs					
GE_{it}	--	--	--	--	-9.08***
CP_{it}	--	--	--	--	-8.25***
FTI_{it}	--	--	--	--	-2.62***
D_Inst_{it}	--	--	--	--	-7.02***
$Taxgdp_{it}$	-9.84***	-10.05***	-2.84***	-9.59***	--
HICs					
GE_{it}	--	--	--	--	-7.45***
CP_{it}	--	--	--	--	-8.57***
FTI_{it}	--	--	--	--	-8.46***
D_Inst_{it}	--	--	--	--	-7.15***
$Taxgdp_{it}$	-10.87***	-11.05***	-10.71***	-11.05***	--
Above TP					
GE_{it}	--	--	--	--	-14.62***
CP_{it}	--	--	--	--	-14.10***
FTI_{it}	--	--	--	--	-14.92***
D_Inst_{it}	--	--	--	--	-11.00***
$Taxgdp_{it}$	-14.81***	-14.72***	-14.96***	-14.68***	--
Below TP					
GE_{it}	--	--	--	--	-4.59***
CP_{it}	--	--	--	--	-4.58***
FTI_{it}	--	--	--	--	-9.65***
D_Inst_{it}	--	--	--	--	6.41***
$Taxgdp_{it}$	-5.90***	-5.84***	-5.49***	-5.96***	--

Note: *, ** and *** indicate the rejection of null hypothesis of no granger-causality and with statistical significance at 10, 5, and 1 percent level respectively.

Appendix V. Long Run Bivariate Panel Causality Analysis

Table A5: Bivariate Panel Causality Test Results

Region/dependent variable	Source of causation: Error correction term (ECT)				
	GE_{it}	CP_{it}	FTI_{it}	D_Inst_{it}	$TaxGDP_{it}$
All countries					
GE_{it}, CP_{it}, FTI_{it}, D_Inst_{it} ($Inst_{it}$)	--	--	--	--	-2.23**
	--	--	--	--	-3.61***
	--	--	--	--	-14.17***
	--	--	--	--	-1.21
$Taxgdp_{it}$	-11.10***	-10.03***	-1.00	-10.35***	--
LICs					
GE_{it}, CP_{it}, FTI_{it}, D_Inst_{it} ($Inst_{it}$)	--	--	--	--	-0.45
	--	--	--	--	-4.85***
	--	--	--	--	-7.00***
	--	--	--	--	-1.12
$Taxgdp_{it}$	-6.89***	-2.84***	-0.87	-5.86***	--
LMICs					
GE_{it}, CP_{it}, FTI_{it}, D_Inst_{it} ($Inst_{it}$)	--	--	--	--	-2.61***
	--	--	--	--	-6.25***
	--	--	--	--	-10.40***
	--	--	--	--	-1.93*
$Taxgdp_{it}$	4.81***	-0.74	-2.74***	-5.44***	--
UMICs					
GE_{it}, CP_{it}, FTI_{it}, D_Inst_{it} ($Inst_{it}$)	--	--	--	--	-3.34**
	--	--	--	--	-5.00***
	--	--	--	--	-6.27***
	--	--	--	--	-1.2
$Taxgdp_{it}$	-4.46***	-3.77***	-4.62***	-4.59***	--
HICs					
GE_{it}, CP_{it}, FTI_{it}, D_Inst_{it} ($Inst_{it}$)	--	--	--	--	-2.04**
	--	--	--	--	-4.33***
	--	--	--	--	-6.28***
	--	--	--	--	-3.31***
$Taxgdp_{it}$	-0.92	-1.33	-1.45	-2.50***	--
Above TP					
GE_{it}, CP_{it}, FTI_{it}, D_Inst_{it} ($Inst_{it}$)	--	--	--	--	-2.92***
	--	--	--	--	-1.37
	--	--	--	--	-12.90***
	--	--	--	--	-0.54
$Taxgdp_{it}$	-9.00***	-9.05***	-2.35***	-9.21***	--
Below TP					
GE_{it}, CP_{it}, FTI_{it}, D_Inst_{it} ($Inst_{it}$)	--	--	--	--	-1
	--	--	--	--	-6.44***
	--	--	--	--	-0.39
	--	--	--	--	-1.26
$Taxgdp_{it}$	-6.78***	-2.25**	-6.22**	-6.00***	--

Note: *, ** and *** indicate the rejection of null hypothesis of no granger-causality and with statistical significance at 10, 5, and 1 percent level respectively.

Appendix VI. Long Run Panel Causality Test (Controlling for Political Stability)

Table A6: Four-Variable Panel Causality Test (Controlling for Political Stability)

Region/dependent variable	Source of causation: Error correction term (ECT)				
	GE_{it}	CP_{it}	FTI_{it}	D_Inst_{it}	$TaxGDP_{it}$
All countries					
GE_{it}, CP_{it}, FTI_{it}, D_Inst_{it} ($Inst_{it}$)	--	--	--	--	-8.86***
	--	--	--	--	-10.45***
	--	--	--	--	-15.23***
	--	--	--	--	-2.42*
$Taxgdp_{it}$	-9.83***	-8.58***	-9.11***	-8.85***	--
LICs					
GE_{it}, CP_{it}, FTI_{it}, D_Inst_{it} ($Inst_{it}$)	--	--	--	--	-4.69***
	--	--	--	--	-5.79***
	--	--	--	--	-6.34***
	--	--	--	--	-2.30**
$Taxgdp_{it}$	-6.22***	-5.40***	-4.97***	-5.58***	--
LMICs					
GE_{it}, CP_{it}, FTI_{it}, D_Inst_{it} ($Inst_{it}$)	--	--	--	--	-5.22***
	--	--	--	--	-5.84***
	--	--	--	--	-6.83***
	--	--	--	--	-2.27*
$Taxgdp_{it}$	-3.90***	-3.16***	-4.26***	-4.18***	--
UMICs					
GE_{it}, CP_{it}, FTI_{it}, D_Inst_{it} ($Inst_{it}$)	--	--	--	--	-3.36***
	--	--	--	--	-6.16***
	--	--	--	--	-6.16***
	--	--	--	--	-1.82*
$Taxgdp_{it}$	-5.19***	-5.04***	-5.42***	-5.49***	--
HICs					
GE_{it}, CP_{it}, FTI_{it}, D_Inst_{it} ($Inst_{it}$)	--	--	--	--	-3.43***
	--	--	--	--	-4.80***
	--	--	--	--	-5.52***
	--	--	--	--	-3.05***
$Taxgdp_{it}$	-2.44***	-3.36***	-3.35***	-2.82**	--
Above TP					
GE_{it}, CP_{it}, FTI_{it}, D_Inst_{it} ($Inst_{it}$)	--	--	--	--	-8.89***
	--	--	--	--	-8.53***
	--	--	--	--	-13.62***
	--	--	--	--	-1.68*
$Taxgdp_{it}$	-8.21***	-8.36***	-7.45***	-8.47***	--
Below TP					
GE_{it}, CP_{it}, FTI_{it}, D_Inst_{it} ($Inst_{it}$)	--	--	--	--	-3.52***
	--	--	--	--	-6.19***
	--	--	--	--	-7.95***
	--	--	--	--	-1.69*
$Taxgdp_{it}$	-4.97***	-6.91***	-6.92***	-7.47***	--

Note: *, ** and *** indicate the rejection of null hypothesis of no granger-causality and with statistical significance at 10, 5, and 1 percent level respectively. The table excludes political stability and GDP per capita interactions with tax capacity and institutions.