# Tax structures and size based tax policies: a cross-country database and comparative study

Pierre Bachas (UC Berkeley) & Anders Jensen (LSE)\*

#### Abstract

We present preliminary results on how tax structures vary across countries and income levels. The data collected provides information on tax rates by income type but also on enforcement mechanisms (e.g. Large taxpayer units, auditing strategies) and design of the tax system (e.g. Alternative minimum taxes). An important goal is to cover all levels of economic development, in order to obtain an accurate picture of practices across income levels. We present descriptive results on large taxpayer units, corporate and alternative minimum taxes and the composition of the personal income tax. We observe that developing countries tend to rely on size based taxed policies instead of direct measures of a taxpayer's payment capacity. A story of information constraint evolving with development (Kleven, Kreiner & Saez 2012) can explain these policies.

#### Motivation

Academic researchers have showed renewed interest in how taxation varies with economic development (Gordon & Li 2009; Kleven, Kreiner & Saez 2010). A starting point for this literature is the large difference in tax to gdp ratios between OECD countries (40%) and developing countries (20%). In a dedicated issue of the Journal of Economic Perspectives, Besley & Person (2014) discuss possible explanations with an emphasis on political economy

<sup>\*</sup>This version March 2015. Corresponding authors: Bachas, bachas@berkeley.edu - Jensen, a.jensen1@lse.ac.uk. We are grateful to Goya Razavi for outstanding research assistance. We thank the International Growth Center and STICERD for its generous support.

and tax morale, itself a separate chapter (Luttmer & Singhal 2014). In this project we emphasize the precise tax revenue mechanisms used across countries. Recent within country studies, using micro data, show that tax structures such as Alternative Minimum Taxes (Best et al 2014), corporate tax rates determined by revenue (Bachas 2014) and Large Taxpayer Units (Almunia & Lopez-Rodriguez 2014) have an important impact on tax collection.

#### Objectives

The first goal is to provide a set of descriptive results on the mix of tax instruments and tax enforcement, with a particular focus on variation across income levels. The second goal is to study the determinants of tax policy choice by income level. Our working hypothesis is that the information stock is a constraint for tax-authorities, who revise the optimal tax and enforcement mix as the information stock increases. The observed reliance on size based tax instruments illustrates this concept. For firms effective corporate tax rates can be determined by the firm's turnover instead of profits (the true income generated). In the case of individuals, the wage and capital income structure can incentivize information reporting by (very) high-income taxpayers. Any tax and enforcement policy relying on size distorts firms and individuals' choices. This generates distortions to production which impose a cost to society. However if size is easily observable, then information constrained authorities might find it revenue efficient to use such policies (Best et al 2014). The third goal is to measure the performance in terms of revenue of the specific tax and enforcement instruments which will require precise data on government revenue items.

#### **Preliminary results**

The analysis is based on a sample of 150 countries. In order to show variation across income levels we group countries into five income quintiles (Table 1 in the Appendix displays the countries in each quintile). Most of the analysis is based on cross-quintile variation in outcomes. In what follows, we present a subset of results on four areas: Large taxpayer units, corporate income tax structures, personal income taxes (salary income, capital income, interest and dividends etc.), tax administrations, tax revenue efforts. The database is being expanded towards enforcement and auditing strategies and indirect taxes.

**Large Taxpayer Units** Large taxpayer units (LTUs) are a special unit dedicated to improve enforcement of important taxpayers. Taxpayers belonging to the LTU are required to report additional information to the tax authorities, such as third party audited accounts and detailed financial statements, and are assigned a larger team of tax auditors. LTU taxpayers might benefit from a faster service tailored to the needs of large taxpayers.

Figure 1 shows that LTUs have become a quasi-universal enforcement tool. Starting in the 1960's a few rich countries first implemented LTUs, and encouraged by the World Bank most developing countries adopted them during the 1990's and early 2000's.





Figure 2 displays the criterion determining which taxpayer belongs to the LTU. Almost all LTUs use turnover but some are also based on industry and past tax payments. More seldom they are determined by the number of employees and assets of the firm. Potentially consistent with a story of evolving information constraints, we observe that least-developed countries are much more reliant on a unique turnover threshold rule for determining LTU liable firms. In richer countries, turnover is often combined with other characteristics, including industry

and amounts of held assets. Size as proxied by past tax-liability is an important threshold criteria in middle income countries.



Figure 2: Large taxpayer unit: threshold combinations

Figure 3 plots the number of taxpayers in the LTU and the share of total tax revenue from LTU on income. First LTU revenue represents on average 65% of total tax revenue and more than 75% in some countries. Second, while the number of taxpayers belonging to the LTU increases with income, the LTU share in total tax revenue does not. Therefore tax revenue collected by LTU taxpayer is decreasing by income quintiles. Poorer countries heavily rely on a size-based enforcement tool, which targets a limited number of large firms and very rich individuals.



Figure 3: Large taxpayer unit: payers and total yield

**Corporate income tax** Corporate income tax (CIT) applies to incorporated entities and typically taxes profits determined as revenue minus deductible costs. Figure 4 displays the top marginal tax rate of CIT by income, excluding countries with less than 1M inhabitants. The average top CIT rate is around 25% with rates varying from 10 to 40%. Top CIT rates slightly decrease with a country's income (slope is marginally significant).



Figure 5 shows the type of CIT structure by income. Three types of tax structures are commonly used: a flat CIT (red), a progressive CIT for which the marginal tax rate depends on profits (orange) and an increasing average tax rate depending on the firm's revenue (green). This later system applies lower rates to small and medium enterprises as measured by sales, potentially to foster small business growth but also to limit evasion responses, since SME's reported profit is known to be an elastic tax base even in rich countries. Flat CIT are the most prevalent tax structure across income quintile and are especially used by countries of the bottom quintile. On the other hand progressive CIT are principally used by rich countries. The need to limit tax evasion opportunities in poorer countries and to simplify the tax system might explain these differences.



Figure 5: Determinant of Corporate Income Tax

So far we have only seen limited evidence of differences in corporate tax structures across income levels. The study of Alternative minimum taxes (AMT) provides a different picture. AMT's general principal is to tax a wider base than the profit base. In its most common form, the AMT taxes sales at a low rate (below 2%) and corporate tax liability is the maximum of the profit liability and the AMT liability. Figure 6 shows the share of countries by income quintile with an AMT as part of their corporate tax system. AMTs are very popular in the bottom three quintiles while rare in the top two quintiles. This income pattern highlights the role of AMTs as a policy response to widespread tax evasion and the erosion of the corporate tax base. They also represent an example of a revenue versus production efficiency tradeoff faced by a constrained government. The second part of Figure 6 pools displays the base of the 41 AMT tax systems. Half of the AMTs only tax turnover, while charging a fixed fee, sometimes varying by revenue or sector is also a common structure. Finally a few AMTs use the firms assets as the tax base.



Figure 6: Alternative Minimum Tax: Existence and tax base

**Personal Income Tax** Personal income taxes are taxes liable by individual taxpayers both on earned income and on capital income. Capital income can be derived from real estate, interest, dividends and shares valuation. Personal income can be taxed either under the territorial system or the residential system. Under territorial rule, domestic taxes apply only to the income made within the country. Under residential rule taxes "follow" the resident and attempt to tax all income-sources of the agent across the globe. Therefore non-residents are only taxed on local income made in the country (where the US is a noticeable outlier, where non-resident citizens are also taxed on worldwide income). Figure 7 plots for each income quintile, the share of countries with a residential (red) or a territorial (blue) system. The decrease in the territorial practice is evident and monotonic as income increases. If less developed countries are on average more constrained in the amount of information they have on transactions in the economy, it might be efficient to first focus tax efforts on the domestic transactions and sources of income. As the economy changes, and more sources of information become available the tax authorities, it may be worthwhile to shift to a residential system which tracks citizens' income beyond domestic borders. However, we should note that even in the lowest quintile over 70% of countries have already adopted a residential system.



Figure 7: Residential versus Territorial Treatment of Personal Income

Figure 8 displays the top MTR of earned income over per capita income. We observe large absolute variation in the rates which vary from 10% to 60%. Richer countries apply higher top tax rates: the fitted line implies that doubling per capita income is correlated with a 2% increase in the top MTR. Interestingly, small tax rates on top earners (below 20%) are concentrated middle income countries, not in the low-income group.



Tax authorities may choose to introduce tax wedges between different sources of income, such as a favoring the treatment of capital income over earned income, to encourage taxpayers to reveal information on their capital income. Figure 9 plots, by capital income source type, the top MTR on the specific source versus the top MTR on earned income. We draw four relevant observations. First, the 45 degree line highlighted in grey implies an equal treatment of capital income and earned income. We observe an important share of countries with equal tax treatment. Second, in a majority of countries capital income is favored (below 45 degree line) and very rarely disfavored (above 45 degree line). Third the average tax gap is largest for dividend and interest income. Fourth, all regression lines are positive and significant. This implies that the largest the top MTR the largest the gap across income sources.



Figure 10 shows the average ratio of each capital income source over earned income by income quintile. We find supporting evidence that tax base and tax rate distortions favoring capital income are more prevalent in poorer than in rich countries. The evidence is stronger for interest and dividends than for capital gains. However countries in the first quintile have

the highest average ratio for all capital income types. This evidence is consistent with a story of a decreasing importance of information constraints over the development path. If less developed countries simply wanted to relieve capital income sources, it would suffice to entirely exempt such income from taxation; but we find very constant adoption-rates for all types of capital income across development-quintiles. Less-developed countries thus appear motivated to get high income taxpayers' capital sources "on the radar" to eventually leverage the information stock.



**Tax administrations** While there is a general consensus that features of tax administrations are important in explaining cross-development variation in tax take, there is little empirical evidence on this. Our database collects a set of simple indicators on characteristics of tax-administrations, and establishes, in a first instance, some facts about how these features change over levels of development. In a second phase, we are planning to set up an empirical model that allows us to assess the extent to which such cross-development variation in administrative features explain observed differences in collected tax revenue.

We present here two sets of results, relating to the 'institutional design' of the country tax-administration, and to the human resource performance of the administration. This data on tax administrations is drawn from the USAID collection of studies on tax administrations and tax structures (2013). Figure 11 presents income-quintile mean differences in three dimensions of institutional design. The first indicator, semi-autonomy, takes value 1 if the country has a semi-autonomous revenue administration, whereby operations are more independent (i.e., reporting may be made to a board of provate and public representatives, rather than to the Ministry of Finance); the pay-structure may be independent of civil service salary structures; and, there exists budget flexibility. The second indicator, customs integration, takes value 1 whenever the tax and customs administrations are integrated and operate as a single administrative unit. Finally, the indicator tax organisation, takes value 1 if the internal structure of the administration is organised according to function - audit department, remittance department, etc - or takes value 0 if the structure revolves around tax-sources - personal income tax department, sales tax department. Considering the three panels of Figure 11, there is not much discernible variation over income-quintiles. This is surprising, given the emphasis of the literature on variation in tax administration practices. On the other hand, there are dimensions of tax administrations not covered in these indicators including registration strategies, audit algorithms - which may vary more over development levels; we intend to collect data on these dimensions in future work.



Figure 11: Tax administration: institutional features by income-quintiles Panel A

In contrast to institutional features, there is salient cross-development variation in human resource practices and 'performance'. Figure 12, panel A and B plot respectively the number of tax administrators in the population, and the number of active taxpayers per tax administrator. Both these ratios strongly increase over development, suggesting tax administrations in more developed countries hire more employees who are faced with a larger set of tax agents to administer. On the other hand, the administration budget, relative to taxes collected (Panel C), strongly decreases over development levels. Taken together, these

3

2

25

0

1

3 panels could suggest, as one of many explanations, that human resource practices become more efficient in more developed countries: employees are more numerous and have more tasks at hand (possibly), yet they bring in more tax revenue relative to what they cost. This suggests a large role for understanding correctly the impact of wage-structure career-concerns on performance of tax administrators (Khan et al., 2014).







Panel C



Tax take and tax-mix over development levels As a final part of the project, we have collected a comprehensive and detailed data-base on sources of tax revenue. Importantly, since our tax and administration data covers the full spectrum of development, it was required to complement it with a dataset on tax-revenue variables which is equally comprehensive. Currently, the 'largest' data-sets on tax-revenues across levels of development include up to 72 observations (Kleven, Kreiner, Saez, 2014) and 102 observations (Besley and Persson, 2010). Using the IMF Government Finance Statistics as the unique basis for our data-collection, we generate cross-sectional graphs with up 164 country observations.

The IMF GFS data-set has the advantage of assembling tax-revenue data for a large set of countries, but, importantly, on the basis of a unified, constant set of accounting reporting principles. These are outlined further in the appendix. This is an advantage over other data-bases, such as the ICTD data-base, which pool several data-sources in order to attain a cross-section of countries across the development spectrum. Another edge given by this dataset is the level of detail - allowing e.g. a breakdown by corporate and personal income, by net wealth and inheritance taxes.

There are two main disadvantages with using the IMF GFS data-base. First, the timespan is not comprehensive, and usually provides a country-panel dimension over years between 2003 and 2011. Second, the IMF GFS accounting framework does not attempt to account for resource revenue within each revenue-source. The ICTD database deals with both of these issues. We have therefore adapted the revenue-classification of the ICTD and mapped it into the GFS structure, so that, whenever available, GFS and ICTD deliver the same country-year tax-source observation. In Figure 13, we explore observed differences between the two data-sets: Panel A (B) plots the entire set of country-year observations (the 2010 cross-country section) for which we have both ICTD and GFS data on total tax revenue, and where we divide it by a common series of GDP. The left side graphs shows the same tax take ratio, but using the non-resource ratio in the ICTD data. The main take-away is that ICTD seems to systematically report higher tax take, and that this reporting-wedge is not driven by differences in resource and non-resource revenue shares in the ICTD data. The wedge does seem to diminish once we consider the most current cross-section of 2010, but there remains a systematic over-reporting in the ICTD data.



With this in mind, Figures 14 and 15 document on tax take and total-tax share of the four main revenue-sources which are hypothesized to grow in importance over development levels: personal income tax (PIT), corporate income tax (CIT), sales tax (ST), and property tax (PT). The graphs provides year-country scatterplots for a sample of 141 countries, together with a fitted line on the underlying observations, and a fitted line on the restricted sample in years after 2008. The main revenue-source whose tax-take and share in total taxes decreases, is the set of taxes on international trade transactions. Figure 14 shows that the tax-take of all four revenue-sources does increase across levels of per capita income, with the PIT increase recording the strongest increase. Interestingly, PIT is also by far the tax-source whose reliance in total taxes increases the most; PT also sees a small increase in reliance. On the other hand, CIT and ST remain at constantly perfect levels of total-tax share. This

heterogeneity across revenue-sources in the tax-mix is a motivating point in Jensen (2015), which argues how employee employment growth along the development path can help explain the observed cross-development rising importance of PIT in tax take and tax-mix.



Figure 14: Tax takes across development

Figure 15: Tax mix across development



#### Conclusion

This study attempts sets forth the working hypothesis that limited information is a key constraint which varies across levels of development, and which impacts the design of tax administrations and statutory policies. To address this hypothesis, we have collected a comprehensive database on features of tax administrations, tax structures, and revenue-outcomes across all levels of development. Using this data, the first goal is to provide a set of descriptive results on the mix of tax instruments and tax enforcement, with a particular focus on variation across income levels. The second goal is to study the determinants of tax policy choice by income level. We find higher reliance on size based tax instruments at lower levels of development, which is consistent with the hypothesis of constrained information stock in these countries. In the case of firms, the large taxpayer unit enforcement instrument tagets firm largely on the basis of turnover in developing countries, and focuses enforcement effort on a small segment of firms; in constrast, large tax-payer units in developed countries construct an information set on each firm which goes beyond proxies for size, and consequently expands the enforcement effort on a larger set of firms. In the case of individuals, the wage and capital income structure are observed to strongly incentivize information reporting by (very) high-income taxpayers, and this policy is more wide-spread in developing countries.

In continued work, we plan on expanding the database in order to be able to study the impact of limited information on design of audit strategies and statutory indirect tax policies.

## References

- [1] Bachas, Pierre [2015]. "Not(ch) your average tax system: behavioural responses of small Costa Rican firms," mimeo, UC Berkeley.
- [2] Bachas, Pierre and Anders Jensen [2015]. "Tax administrations and size-based policies: a cross country database and study", mimeo, London School of Economics and UC Berkeley.
- [3] Timothy, Besley and Torsten Persson [2009]. "The Origins of State Capacity: Property Rights, Taxation, and Politics," *American Economic Review*, 99(4), 1218-44.
- [4] Besley, Timothy and Torsten Persson [2011]. *Pillars of Prosperity: The Political Economics of Development Clusters*, Princeton: Princeton University Press.
- [5] Besley, Timmothy and Torsten Persson [2012]. "Taxation and Development," in A. J. Auerbach and M. S. Feldstein, eds., Final Draft Chapter for the *Handbook of Public Economics*, 2012.
- [6] Best, Michael C. [2014]. "Salary Misreporting and the Role of Firms in Workers' Responses to Taxes: Evidence From Pakistan," mimeo, Stanford Institute for Economic Policy Research.
- [7] Brockmeyer, Anne, Michael Best, Henrik Kleven, Johannes Spinnewijn, and Mazhar Waseem [2013]. "Production versus Revenue Efficiency with Limited Tax Capacity: Theory and Evidence from Pakistan," mimeo, London School of Economics.
- [8] Carillo, Paul, Dina Pomeranz, and Monica Singhal [2014]. "Tax Me if You Can: Evidence on Firm Misreporting Behavior and Evasion Substitution," mimeo, Kennedy School of Government, Harvard University.
- [9] Chetty, Raj [2009]. "Sufficient Statistics for Welfare Analysis: A Bridge Between Structural and Reduced-Form Methods" Annual Review of Economics 1: 451-488, 2009
- [10] Gordon, Roger H. and Wei Li [2009]. "Tax Structures in Developing Countries: Many Puzzles and a Possible Explanation," *Journal of Public Economics*, 2009, 93 (7-8), 855-866.
- [11] International Tax Dialogue [2010]. "Revenue Administration in Sub-Saharan Africa," ITD Comparative Information Series No.1
- [12] Khwaja, Munawer Sultan, Rajul Awasthi, and Jan Loeprick [2011]. "Risk-Based Tax Audits: Approaches and Country Experiences," *Directions in Development: Finance*, The World Bank.
- [13] Jensen, Anders [2015]. "Rise of employees and growth in tax capacity," mimeo, London School of Economics.
- [14] Kleven, Henrik J. [2014]. "How Can Scandinavians Tax So Much?," Journal of Economic Perspectives, 28(4), 77-98.

- [15] Kleven, Henrik J., Claus T. Kreiner, and Emmanuel Saez [2014]. "Why can Modern Governments Tax so Much? An Agency Model of Firms as Fiscal Intermediaries," NBER Working Papers 15218, National Bureau of Economic Research.
- [16] Kleven, Henrik J., Martin Knudsen, Claus T. Kreiner, Soren L. Pedersen, and Emmanuel Saez [2011]. "Unwilling or Unable to Cheat? Evidence from a Tax Audit Experiment in Denmark," *Econometrica*, 79 (3), 651-692.
- [17] Kleven, Henrik J., and Mazhar Waseem [2013]. "Using Notches to Uncover Optimization Frictions and Structural Elasticities: Theory and Evidence from Pakistan," *Quarterly Journal of Economics* 128, 669-723.
- [18] Kumler, Todd J., Eric A. Verhoogen, and Judith Frias [2012]. "Enlisting Workers in Monitoring Firms: Payroll Tax Compliance in Mexico," Columbia University Department of Economics Discussion Papers, 96.
- [19] Organisation for Economic Development and Co-Operation [2013]. "Tax Administration 2013: Comparative Information on OECD and Other Advanced and Emerging Economies," OECD Publishing.
- [20] Pomeranz, Dina [2012]. "No Taxation without Information: Deterrence and Self-Enforcement in the Value Added Tax," Harvard Business School Working Paper, 13-057.
- [24] Munawer Sultan Khwaja, Rajul Awasthi, and Jan Loeprick [2011] "Risk-Based Tax Audits: Approaches and Country Experiences", The World Bank Group.

## Appendix 1 - Sample countries

1st Quintile	2nd Quintile	3rd Quintile	4th Quintile	5th Quintile
Afghanistan	Angola	Argentina	Antigua and Barbuda	UAE
Benin	Albania	Azerbaijan	Bahrain	Australia
Burkina Faso	Armenia	Bulgaria	Barbados	Austria
Bangladesh	Bolivia	Bosnia and Herzegovina	Botswana	Belgium
Central African Republic	Bhutan	Belarus	Chile	Bahamas
Congo, DR	Cote d'Ivoire	Belize	Czech Republic	Canada
Ethiopia	Congo, R	Brazil	Dominica	Switzerland
Ghana	Egypt	China	Estonia	Cyprus
Gambia, The	Georgia	Colombia	Equatorial Guinea	Germany
Kenya	Guatemala	Cabo Verde	Grenada	Denmark
Kyrgyz Republic	Honduras	Costa Rica	Croatia	Spain
Cambodia	Indonesia	Dominican Republic	Hungary	Finland
Laos	India	Algeria	Lebanon	France
Liberia	Kiribati	Fiji	St Lucia	United Kingdom
Madagascar	Sri Lanka	Iran, Islamic Rep.	Lithuania	Greece
Mali	Lesotho	Jordan	Latvia	Ireland
Mozambique	Morocco	Kazakhstan	Mexico	Iceland
Malawi	Moldova	Maldives	Malta	Israel
Niger	Mongolia	Marshall Islands	Mauritius	Italy
Nepal	Nigeria	Macedonia	Oman	Japan
Pakistan	Nicaragua	Malaysia	Palau	Korea
Rwanda	Philippines	Namibia	Poland	Kuwait
Togo	Papua New Guinea	Panama	Portugal	Luxembourg
Tajikistan	Paraguay	Peru	Russia	Netherlands
Tanzania	West Bank Gaza	Romania	Slovak Republic	Norway
Uganda	Senegal	El Salvador	Slovenia	New Zealand
Vietnam	Swaziland	Serbia	Seychelles	Qatar
Yemen	Ukraine	Thailand	Trinidad and Tobago	Singapore
Zambia	Uzbekistan	Tunisia	Turkey	Sweden
Zimbabwe	Vanuatu	South Africa	Uruguay	United States

Table 1: Sample countries by per capita income quintile

## Appendix 2 - Cross-country database construction

## **1** IMF Government Finance Statistics

## 1.1 IMF GFS: institutional details

#### 1.1.1 Standardized classification: details of the revenue

The IMF GFS allows for a very disaggregated set of variables for a set of 169 countries from 2000 to 2013. It contains 70 variables on tax revenue on 4 categories: "Taxes", "Social contributions", "Grants" and "Other revenues". "Taxes" is decomposed in 6 main subsections: "Taxes on income, profits, and capital gains", "taxes on payroll and workforce", "taxes on property", "taxes on goods and services", "taxes on international trade and transactions" and "other taxes". Each again contains one to two level of disaggregation. This points out to the over encompassing nature of the IMF GFS database, which is crucial for our research.

The classification used has a certain degree of consistency across government revenue databases and uses almost the same structure as the OECD "Revenue Statistics" publication, for example. The GFS classification is reproduced in Figure 16.

1	Revenue	122	Other social contributions [GFS]
11	Taxes	1221	Employee contributions [GFS]
111	Taxes on income, profits, and capital gains	1222	Employer contributions [GFS]
1111	Payable by individuals	1223	Imputed contributions [GFS]
1112	Payable by corporations and other enterprises	13	Grants
1113	Other taxes on income, profits, and capital gains	131	From foreign governments
11131	Pavable by general government	1311	Current
11132	Unallocable taxes on income, profits, and capital gains	1312	Capital
112	Taxes on payroll and workforce	132	From international organizations
113	Taxes on property	1321	Current
1131	Recurrent taxes on immovable property	1322	Capital
1132	Recurrent taxes on net wealth	133	From other general government units
1133	Fetate inheritance and oilt taxes	1331	Current
1133	Capital levies	1331	Conital
1130	Other resurrent terren on preparty	44	Other revenue
1130	Other recurrent taxes on property	14	Other revenue
114	Taxes on goods and services	141	Property income [GPS]
1141	General taxes on goods and services	1411	Interest [GFS]
11411	Value-added taxes	14111	From nonresidents
11412	Sales taxes	14112	From residents other than general government
11413	Turnover and other general taxes on goods and services	14113	From other general government units
11414	Taxes on financial and capital transactions	1412	Dividends
1142	Excise	14121	From nonresidents
1143	Profits of fiscal monopolies	14122	From residents
1144	Taxes on specific services	1413	Withdrawals from income of guasi-comprations
1145	Taxes on use of goods and on permission to use	1414	Property income from investment income
44464	Motor unbide tower	1415	Dest
11401	Motor vehicle taxes	1415	rveni.
11452	permission to use goods or perform activities	1416	Reinvested earnings on foreign direct investment
114521	Business and professional licenses	142	Sales of goods and services
114522	Pollution taxes	1421	Sales by market establishments
114523	Radio and television licenses	1422	Administrative fees
114524	Licenses and permits for households	1423	Incidental sales by nonmarket establishments
114525	Other taxes on use of goods and on permission to use goods or perform activities not elsewhere classified	1424	Imputed sales of goods and services
1146	Other taxes on goods and services	143	Fines, penalties, and forfeits
115	Taxes on international trade and transactions	144	Transfors not olsowhere classified
4484	Customs and other impact duties	144	Current broofers and classified
1151	Customs and other import duties	1441	Current transiers not eisewhere classified
1152	Taxes on exports	14411	Subsidies
1153	Profits of export or import monopolies	14412	Other current transfers not elsewhere classifie
1154	Exchange profits	1442	Capital transfers not elsewhere classified
1155	Exchange taxes	145	Premiums, fees, and claims related to nonlife insurance and standardized guarantee schemes
1156	Other taxes on international trade and transactions	1451	Premiums, fees, and current claims
116	Other taxes	14511	Premiums
1161	Pavable solely by business	14512	Fees for standardized guarantee schemes
1162	Payable by other than business or unidentifiable	14513	Current claims
12	Social contributions (GFS)	1452	Capital claims
121	Social security contributions (GES)	1106	seaperal violance
1211	Employee contributions (GES)		
1211	Employee contributions [GFS]		
1212	Employer contributions [GFS]		
1213	Self-employed or nonemployed contributions [GFS]		
1214	Unallocable contributions [GFS]	I	

## Figure 16: IMF GFS Classification [GFS]

### 1.1.2 Levels of Government and consolidation

General Government encompasses three categories: central government, state government and local government. Each of these can then be divided in three subsections: Budgetary, extrabudgetary (entities with a budget not included in the general budget) and social security funds.

The important work of the IMF GFS has been to consolidate general and central government, as applicable. Consolidating, essentially means controlling for the fact that there might be flows within entities that might be reported twice, and would hence lead to an upward bias in the data. This is done by considering a set of entities as one. It eliminates also the possibility of inconsistent data due to different administrative features in countries over time.

"To relate government aggregates to the economy as a whole (e.g., revenue, expense, or debt to GDP ratios), it is better to eliminate the internal movement of economic value and include only those flows and stock positions that actually cross the boundaries with other sectors or nonresidents." (IMF GFS Manual, 2001) The GFS levels of government is documented in Figure 17.



Figure 17: Levels of Government [GFS]

#### 1.1.3 Coverage: by income-group

Although the GFS is comprehensive in terms of the variables it allows to consider in a analysis, it nevertheless has gaps in its coverage. Moreover, it does not cover pre-2000 data in a consistent manner, since there has been a change in the classification.

Figure 18 displays the coverage in the IMF GFS and the ICTD databases, over income levels and years. The ICTD has a broader coverage than the IMF database. The gap in coverage between the two competing databases becomes smaller in recent years, as the IMF GFS coverage has significantly improved.

		ICTD	IMF Fiscal Affairs	Article IV (Central Gov't) <sup>35</sup>	GFS (Central Gov't <sup>36</sup>	WDI (Central Gov't
Low-	Total revenue	578	411	388	366	303
income	Total tax	576	513	552	367	303
	Tax on income, profit and capital gains	475	513	409	367	303
	Tax on goods and services	496	510	382	367	303
Lower middle- income	Total revenue	856	709	463	591	445
	Total tax	856	693	638	598	445
	Tax on income, profit and capital gains	655	650	400	598	445
	Tax on goods and services	664	625	315	598	445
Upper	Total revenue	883	793	633	434	312
middle-	Total tax	916	770	705	431	312
income	Tax on income, profit and capital gains	770	746	532	430	312
	Tax on goods and services	792	721	395	430	312
All developing countries	Total revenue	2317	1913	1484	1391	1060
	Total tax	2348	1976	1895	1396	1060
	Tax on income, profit and capital gains	1900	1909	1341	1395	1060
	Tax on goods and services	1952	1856	1092	1395	1060
High- income OECD	Total revenue	542	607	282	452	333
	Total tax	587	605	214	461	333
	Tax on income, profit and capital gains	568	605	205	460	333
	Tax on goods and services	586	603	150	460	333
High-	Total revenue	456	351	376	227	188
income non-OECD	Total tax	428	353	256	234	188
	Tax on income, profit and capital gains	351	322	150	234	188
	Tax on goods and services	357	309	56	234	188
All	Total revenue	3465	3022	2250	2070	1581
countries	Total tax	3513	3071	2506	2091	1581
	Tax on income, profit and capital gains	2956	3040	1801	2089	1581
	Tax on goods and services	3035	3025	1397	2089	1581

Figure 18: Data availability by source and country income group 1990-2010 [ICTD]

### 1.1.4 Accounting for social security

The IMF GFS in its original classification used to include the private sector side of social security in the "Tax" category and the public sector side of social security in the "Non-Tax" category. This has however been changed in the 2001 classification, where social security funds are now a separate category, allowing it to be completely distinct from other government revenue.

The IMF GFS takes into account social security by considering which government unit it is managed by. This allows for consistency and comparability across countries.

## 1.2 The value-added of the IMF GFS

To review, there are two main reasons to use the IMF GFS database: consolidation and level of disaggregation. Emphasis should be put on the importance in the level of detail available

in the database. Compared to any other database, such as the newly released ICTD, the IMF GFS has this huge advantage which we are here fully exploiting.

Essentially, the IMF GFS, has often one and sometimes two more levels of details than the ICTD database. Of particular interest are the availability of revenue from taxes on "immovable property taxes", "net wealth", "estate, inheritance, and gift taxes" and, "property income" revenue such as "dividends" and "interest" which are further broken down into "nonresidents" and "residents" categories. For all of these categories, specific variables have been collected on countries' legislation which justifies the use of the IMF GFS to gather revenue variable at such a disaggregated level.

As can be see from Figure 19, there are much more variables available in the IMF GFS. Variables included in both the IMF GFS and ICTD are highlighted in orange.

1Total Gov't Revenue				
	11Taxes			
		111Taxesonincomeprofit	=	
			1111Individuals	_
			1112Unallocable	_
		112Taxesonpayrollandwo		
		113Taxesonproperty	11210	
			1131Recurrenttaxesonimm 1132Recurrenttaxesonnet	
			1133Estateinheritancea	
			1134Taxesonfinancialand	
			1136Otherrecurrenttaxes	
		114Taxesongoodsandserv	_	
			1141Generaltaxesongoods	_
				11411Valueaddedtaxes 11412Salestaxes
			1142Excises	11413Turnoverothergene
			1143Profitsoffiscalmono	
			1144Taxesonspecificserv	
			1145Taxesonuseotgoods	11451Motorvehiclestaxes
		115Taxesonintltradean	1146Othertaxesongoodsa	114520tile/
			1151Customsandotherimpo	
			1152Taxesonexports	
			1153Profitsofexportsor	
			1154Exchangeprofits 1155Exchangetaxes	
			1156Othertaxesoninterna	
		116Othertaxes	_	
	12Social Contribution	121Socialsecuritycontrib		
		12190clubecul tycolicity	1211Employeecontributions	
			1212Employercontributions	
			1213Selfemployedornonem 1214Unallocablecontributi	
		1220 thersocial contributi		
			1221Employeecontributions	
			1222Employercontributions 1223Imputedcontributions	
	13Grants	_		
		131Fromforeigngovernment	12110	
			1312Capital	
		132Frominternationalorga		
			1321Current	
		133Fromothergeneralgove	1322000101	
		1331Current		
	14Other revenue	1332Capital		
	140ther revenue	1411Interest		
		1412Dividends		
		1413Withdrawalsfromincom		
		1415Rent		
		141Propertyincome		
		1421Salesofmarketestabl 1422Administrativefees		
		1423 Incidental sales by no		
		1424 Imputed sales of goods		
		1425alesorgoodsandserv 143Finespenaltiesandfo		
		1441Current		
		1442Capital		
		14Otherrevenue		

## 2 Alternative data-set on government revenues: ICTD

### 2.1 Surveying available datasets

Ultimately, the ICTD GRD has drawn on six cross-country datasets: IMF Government Finance Statistics (GFS), World Bank World Development Indicators (WDI), OECD Tax Statistics, OECD Revenue Statistics in Latin America dataset, CEPAL Tax Statistics, AEO African Fiscal Performance. The ICTD dataset provides one major advantage, in that it has attempted to separate resource and non-resource revenues in all reported revenue country series. Accounting for natural resource revenue, means here to either record resource revenue as non-tax revenue or to separate out from the non-resource component of tax revenue, for countries with significant resource wealth. The main ICTD cost of this methodoloy is that they are not able to report revenue-sources at a level of disaggregation that the IMF GFS database covers.

Three other datasets developed by researchers were surveyed but not included in the final dataset: Keen and Mansour (2009), which is based on IMF Article IV reports, the Michigan Ross School of Business World Tax Database (WTD) and the Oxford Latin America Economic History Database (OxLAD).

#### [FIGURE 5 AROUND HERE]

## 2.2 ICTD data-collection: IMF Article IV reports and other sources

The data-sources used for the ICTD database are displayed in Figure 20. One major contribution of the ICTD dataset has been to systematically compile the data from IMF Article IV. Although the classification in those articles is not as precise from an accounting perspective as the IMF GFS, it allows filling in gaps in regions not covered by international databases.



Figure 20: Sources for revenue data [ICTD]

Note: CEPAL gives a partial account of its sources when they are national statistics offices; its sources otherwise are not indicated (but may be WEO and/or GFS). The source for Michigan Ross WTD is not indicated but clearly derives from GFS. For OxLAD the source is identified as GFS/IFS for 1970 onwards.

## 2.3 ICTD: Standardizing the underlying GDP series

In order to get consistent data from merging those international dataset, one needs to make sure the underlying GDP series are standardized, eliminating discontinuities and ensuring comparability. This has been undertaken by using mainly the WEO GDP series in LCU, which has the advantage of recording regularly GDP rebasing. When data is missing from those series, alternative sources are used, most often the IMF IFS, only in cases it matches closely the WEO GDP in overlapping and nearby years. There are two issues that arise when using Article IV: all the data is found in terms of GDP, so it has been converted back to LCU before getting ratios by using the constructed GDP series. Moreover, currency changes have also been taken into account.

## 2.4 ICTD: Data merging and creation of the 'first choice' dataset

The first step here has been to create a collated dataset, encompassing all sources. The merging was then undertaken using a formula based on four factors: "(a) the level of disaggregation of the data (more is better); (b) the length and completeness of the time series; (c) the level of overlap and consistency when combining multiple sources for a single country; and (d) a ranking of data sources, with IMF Article IV and GFS data given priority with the exception of OECD countries, where OECD data was given priority."

However such method has proven less robust than expected and some manual data cleaning was undertaken in order to construct the ICTD database.

## 2.5 ICTD: Ensuring compatibility between unitary and federal states

As mentioned in "Levels of Government" it is important to account for differences between central government and general government data. General government data has been used whenever possible. However, when this has not been the case a choice had to be made. In centralised countries (generally low-income), the focus has been to gather central government data whereas in countries that have substantial sub-national revenue, the focus has been to gather general government data.

## 2.6 ICTD: Problem of disentangling goods and services taxes and trade taxes

Inescapable definitional inconsistencies has led to aim for consistency within countries, though not across. Econometric solutions are twofold here. Using within country econometric estimators to control for individual heterogeneity and using less disaggregated data such as indirect tax revenue.

## 2.7 ICTD: Problems accounting for resource revenue

First, when the IMF article IV does not exist or is not in the public domain, the distinction is not possible. Moreover, differences across countries in the definition of resource revenue can lead to inconsistencies. However, using Article IV throughout instead of relying on countrylevel data can at least partly resolve this issue. Furthermore, Article IV does not consistently account for resource revenues in countries for which that revenue makes up for less than 1-2% of GDP. Lastly, with rising mineral prices and improved taxation of the mining sector in at least some low-income countries, there is a particular need for careful attention to the potential for distortions to the data post-2010.

## 2.8 ICTD: Inherent limitations of available data and the value of transparency

As explained earlier with the manual cleaning of the data, a degree of subjectivity is inevitable in choosing which source is more credible in a given instance. However, full transparency about the data choices is offered. However, some researchers might still feel that data merging is suboptimal because of definitional differences. It is also important to recognize that data collection in some countries, for example in Africa, is often imperfect. "For this reason, there is a strong case for avoiding policy and research conclusions that are driven by very small trends in the data." Over and above the difficulties of getting good data points, a problem arise when comparing countries by using GDP measures. As described previously, the fact that rebasing can be irregular, can lead to underestimation of GDP and hence overestimating of tax to GDP ratios, which is for example what happened in Ghana. Hence short-term and within countries analysis is more robust to such changes than long term cross-country analyses.

## 3 Data collection on taxation practices

## 3.1 Individual and Capital Income Taxation

### 3.1.1 Coverage and Variables of interest

The database covers 157 countries. Variables of interest comprises 8 major categories: personal income, capital gains, self employment, payroll, VAT, dividend, interest and taxes on capital (such as inheritance, wealth and property taxes). To increase precision a specific taxes on capital gains have been split in two categories, allowing to account for real estate specific capital gains taxes. Most countries have been covered and data is widely available. The exception is taxes on capital where data is more sparse.

### 3.1.2 Sources, Quality and Consistency

All data points have been gathered from three public documents released by consulting firm Ernst and Young (EY). These sources have mainly been complementing each other across variables and not within, as to ensure comparability - this is not the case for taxes on capital where two different sources where used, issued by the same firm.

For personal and capital income tax, the source used is the Worldwide personal tax guide 2013-2014 from EY. Second, for VAT, the source used is also from EY: the Worldwide VAT GST and sales tax guide 2014. Third, the international estate and inheritance tax guide was used in addition to the first source to gather data for taxes on capital.

The documents were analysed carefully to gather data with precision. Great care was taken to note any exceptions, or details relevant to the data collected, which are included in the comments. Having in mind consistency, only documents from the same firm was used throughout the process. This cannot be a guarantee of consistency but lessens the risk of getting incomparable data.

## 3.1.3 Challenges

Although sources where used in a manner that would maximise consistency, a problem that arises in the use that is made of different tax legislation in practice. As seen from comments in the database, often restrictions applies or exceptions are made which might affect the tax rate that is used in practice. Very little can be done against this, apart from consistently flagging those exceptions in the comments.

## 3.2 Large Taxpayers Unit (LTU)

## 3.2.1 Coverage and Variables of interest

The database comprises 160 countries. Data is collected on 5 major features of LTUs: existence, threshold characteristics, capacity and yield. The availability of the data is very heterogenous across countries. All variables have been recorded in LCU, as far as this has been possible given the sources available.

## 3.2.2 Sources, Quality and Consistency

The use of a limited amount of sources has not been possible in this case. The process required meticulous research online in different languages (English, French, Spanish) in order to gather a satisfactory amount of data. Data came from a very diverse set of sources, especially for developing countries where data could be found in press releases or newspaper articles. In more developed countries, regional databases were available which eased the process. This is the case for example in South America with the State of Tax Administration in Latin America 2006-2010 compiled by CIAT or in Asia where the ADB published in 2014 A Comparative Analysis of Tax Administration in Asia and the Pacific. In Africa, Revenue Administration in Sub-Saharan African, has also been used. These sources were complemented by the OECD Tax Administration 2013 report. Documents from international organisations such as the World Bank and the IMF have also been used.

Due to such a wide range of sources being used, consistency cannot be guaranteed. However, when possible checks across sources for a given countries have been undertaken to take the most accurate data. If differences existed, data arising from international organisations (the OECD being given most priority) and in more recent years have been given more weight and always supersede the data found in other probably less credible sources such as newspaper article.

In all cases, all data points have been precisely documented to ensure full transparency.

## 3.2.3 Challenges

Ensuring consistency is a task that is in this context very hard, if not impossible. Moreover, since some data points within countries have been collected from different years, any changes in reforms within a country during those years might impact the relevance of the data collected.

## 3.3 Corporate Income Tax

## 3.3.1 Coverage and variables of interest

The dataset on CIT contains 182 countries out of which we have been able to code existence of an Alternative Minimum Tax in 120 countries. We focus on understanding the rates and structure of the tax system, loss carry forward and carry backs, as well as any special tax regime that applies to specific sectors. AMT data describes the alternative base and rate as well the rules determining if a firm belongs to the AMT schedule instead of the regular schedule.

## 3.3.2 Quality and consistency

The data combined Ernst and Young global corporate tax guide, KPMG country tax profiles and Deloitte's tax guides and highlights. Since those variables are mainly statutory variables they were very consistent across the different guides used. Special regimes and Alternative regimes were sometimes present only in one or two of the tax guides.

## 3.3.3 Challenges

The key challenges relate to special tax regimes that might be missed due to non coverage by the guides. In addition a desired section on investment and R&D subsidies and tax incentives was started but not included due to the difficulty in comparing sources and classifying the relevant information.