



How do we Improve the Quality of Tax Data for Research Purposes?

Wilson Prichard

International Center for Tax and Development

Tiangboho Sanogo

University of Auvergne, Clermont-Ferrand, France

University Felix Houphouet-Boigny, Côte d'Ivoire



The Challenge

High quality tax data for research purposes can be a critical tool:

- For tax administration to assess, and improve, their performance
- For researchers to analyze tax performance, and the connections between tax and broad development outcomes, in order to inform policy and practice

However, available tax data has historically been incomplete, and often highly inaccurate, and the ICTD has invested heavily in seeking to address this gap.

Overview

1. The ICTD Government Revenue Dataset
2. Improving the use of data – and next steps
3. Strengthening Sub-National Tax Data



THE ICTD GRD



The ICTD GRD - Motivation

- Initiated in 2010, and launched in 2014, responds to major gaps and analytical inaccuracies in existing international datasets
- Constructed by merging data from all major international datasets and IMF Article IV reports

The ICTD GRD – Data

The ICTD GRD data is notable relative to other sources in five major ways:

1. Much improved data coverage, particularly for developing countries
2. Eliminates important errors and inaccuracies, including related to GDP data
3. More systematic treatment of revenues from natural resources
4. Systematic attention to differences between centralized and federal states
5. Greater transparency than other researcher datasets

The ICTD GRD - Impact

The ICTD GRD has already had a demonstrated impact on research findings and practice:

1. The dataset is now a leading dataset for tax researchers, and expanding
 - Over 5000 downloads over past three months
 - Major research conference at UNU-WIDER
2. Early findings provide evidence that employing improved data yields new and better research results
3. Provides a clearer and more reliable picture of cross-country and within country trends for policy makers

Recent Data Initiatives

1. OECD/ATAF data collection for Africa: Important initiative for deepening data quality in participating countries
2. IMF WoRLD dataset: Major step forward in data transparency
3. IMF work on natural resource revenues: A critical initiative for the future

Sustainability and Adoption

Researchers have long sought to improve tax data – but sustainability, and wide adoption, have been key barriers.

1. Short term: The ICTD has completed a first update of the data, and actively publicized the data, which is now in wide use
2. Long term: We have entered into a partnership with UNU-WIDER who will host, and consistently update, the data, in partnership with the ICTD.

Next Steps – Data Use

Knowing the data being used for research, and using it appropriately

1. Understanding what the numbers capture, and their implications, critical to interpreting results and appropriate analysis
2. Caution in using tax ratios, and using cross-country comparisons appropriately
 - Problems with GDP data can make cross-country comparisons misleading
 - Evaluating tax performance depends on revenue collected, but also understanding the tax base

Next Steps – Resource Revenues

The ICTD GRD attempts to distinguish resource revenues from other tax revenue, but better distinctions are needed

- Clearer definitions
- Better, and publicly available, data
- IMF initiative on this area holds significant potential

Next Steps – Deeper Data

While ICTD GRD tells us how much revenue is collected, understanding of tax systems also requires understanding of what exactly is paid, by who, and relative to what is possible

- How is tax performance relative to potential?
- How fair is the tax system in terms of overall structure?
- How complete is compliance, or do some groups avoid payment? What is the structure of the tax base?



Local Government Dataset in Cote D'Ivoire



Overview

- The Côte d'Ivoire LGR data aims at filling the gap relative to absence of dataset on local governments' revenue;
- This data represents the most complete and most accurate dataset on local government revenue, covering 12 years for 90% of the 196 existing municipalities;
- It is an important tool for statistical analysis and tax policy reforms;
- However, the dataset suffers from two major limitations : Missing data remain and some variables need to be disaggregated;
- There is a need for training local administrations and to set up electronic system for data collection;
- It is valuable to undertake such project in developing countries where not existing.

Outline

1. Motivations
2. Methodology
3. Outputs
4. Limitations
5. Using the data and lessons

Motivations

1. Crucial need to undertake research to deepen processes of state-building:
 - Local revenue dataset allow to undertake research that foster administrative reforms
2. Absence of aggregated, reliable data covering all municipalities over a long period.
 - The existing national sources suffer from limitations : Inconsistency with the definition of variables between municipalities, presence of outliers, significant errors
3. A comprehensive dataset contributes to widen the debate on local governments' responsibilities.

Methodology - Access

Contacting the different sources of data and dealing with constraints of gaining access

- Sending formal letters for clear explanations;
 - Organizing several rounds of meetings;
 - Interviewing eight selected municipalities.
-
- Department in charge of Decentralization and Local Development, Ministry of Interior (Data requested: The administrative accounts for municipalities (local tax and non-tax revenues and expenditures))
 - Ministry of Economy and Finance (Data requested : The national account for local governments grants)
 - Financial services of 8 municipalities (Data requested : Local accounts) (Bouake (Centre); Korhogo (North), Daloa (West); Kanakono (North); Abobo, Yopougon, Plateau(South), Tengrela(North))

Methodology –Data Cleaning

Making suitable for research (Data cleaning)

- Manual data entering
- Dealing with outliers, possible errors of transcription or reporting
- Running some standard statistical tests
- Adding note for clear explanation
- Finding all abrupt changes and treating
- Merging data form the different sources

Outputs (Data)

REVENUE STRUCTURE 2001-2009	REVENUE STRUCTURE 2010-2012
Shared Tax <ul style="list-style-type: none"> - Property Tax - Patent and Licenses - Synthetis Tax Own Revenue <ul style="list-style-type: none"> - Flat rate tax - Civil Registration - Markets fees - Lease fees - Fees for the use of public facilities - Other own revenue Other revenue <ul style="list-style-type: none"> - Vignettes - Grants from central government - External funds - Ordinary Reserve Fonds - Miscellaneous revenue - Loans 	Shared Tax <ul style="list-style-type: none"> - Property Tax - Patent and Licenses - Synthetis Tax - Vignettes Own Revenue <ul style="list-style-type: none"> - Flat rate tax - Civil Registration - Markets fees - Lease fees - Fees for the use of public facilities - Other own revenue Other revenue <ul style="list-style-type: none"> - Grants from central government - External funds - Ordinary Reserve Fonds - Miscellaneous revenue - Loans

Outputs (Data)

transfers[18] 272231										
	municipality	year	region	localisation	population	grants	ownrevenue	sharedrevenue	otherrevenue	totalrevenue
1	Abobo	2001	ABIDJAN	SOUTH	691790	.	667733	362560	4103	1034396
2	Abobo	2002	ABIDJAN	SOUTH	713185	.	818616	188130	1437	1008183
3	Abobo	2003	ABIDJAN	SOUTH	735243	.	967351	509506	11994	1488851
4	Abobo	2004	ABIDJAN	SOUTH	757962	.	894268	347127	3094	1244489
5	Abobo	2005	ABIDJAN	SOUTH	781425	.	895390	598782	751131	2245303
6	Abobo	2006	ABIDJAN	SOUTH	805593	.	734971	566596	274788	1576355
7	Abobo	2007	ABIDJAN	SOUTH	830508	.	736696	759197	305879	1801772
8	Abobo	2008	ABIDJAN	SOUTH	856194	.	795787	689356	279700	1764843
9	Abobo	2009	ABIDJAN	SOUTH	882674	.	754816	1133987	91786	1980589
10	Abengourou	2001	INDENIE	EAST	86180.6	39072	215452	97953	19991	372468
11	Abengourou	2002	INDENIE	EAST	88845.9	50000	244329	96659	14727	405715
12	Abengourou	2003	INDENIE	EAST	91593.7	50000	267689	131258	15723	464670
13	Abengourou	2004	INDENIE	EAST	94426.5	55000	314584	149985	26174	545743
14	Abengourou	2005	INDENIE	EAST	97346.9	52500	279067	127644	37419	496630
15	Abengourou	2006	INDENIE	EAST	100358	55710	271840	169710	51408	548568
16	Abengourou	2007	INDENIE	EAST	103462	50000	248693	140578	31346	470617
17	Abengourou	2008	INDENIE	EAST	106661	50000	243094	137327	18292	448713
18	Abengourou	2009	INDENIE	EAST	109960	50000	246020	222231	16552	534803
19	Aboisso	2001	SUD-COMOE	SOUTH	40975.7	41842	83844	23917	48967	198570
20	Aboisso	2002	SUD-COMOE	SOUTH	42243	42842	76144	117924	22803	259713
21	Aboisso	2003	SUD-COMOE	SOUTH	43549.5	42842	94496	27335	29158	193831
22	Aboisso	2004	SUD-COMOE	SOUTH	44896.4	48842	99061	61374	55226	264503
23	Aboisso	2005	SUD-COMOE	SOUTH	46284.9	46400	89752	47636	113453	297241
24	Aboisso	2006	SUD-COMOE	SOUTH	47716.4
25	Aboisso	2007	SUD-COMOE	SOUTH	49192.2	44000	94520	109802	88642	336964
26	Aboisso	2008	SUD-COMOE	SOUTH	50713.6	44000	113655	109720	29601	296976
27	Aboisso	2009	SUD-COMOE	SOUTH	52282	44000	114070	167291	66345	391706
28	Adiake	2001	SUD-COMOE	SOUTH	29470.4	39071.56	18770.99	31099.79	8792.28	97734.62

Variables

Filter variables here

Variable	Label
<input checked="" type="checkbox"/> municipality	Municipality
<input checked="" type="checkbox"/> year	year
<input checked="" type="checkbox"/> region	region
<input checked="" type="checkbox"/> localisation	geographical loc
<input checked="" type="checkbox"/> population	estimated popul
<input checked="" type="checkbox"/> grants	Equalisation gran
<input checked="" type="checkbox"/> ownrevenue	Local own revenu
<input checked="" type="checkbox"/> sharedrevenue	Sharing tax_Reve
<input checked="" type="checkbox"/> otherrevenue	None tax revenue
<input checked="" type="checkbox"/> totalrevenue	Local total reven
<input checked="" type="checkbox"/> codicity	codicity
<input checked="" type="checkbox"/> code	codicity
<input checked="" type="checkbox"/> grant_part	share of equalisa
<input checked="" type="checkbox"/> ownrev_part	share of own tax
<input checked="" type="checkbox"/> sharedtax_part	share of tax shar
<input checked="" type="checkbox"/> otherrev_part	share of none tax
<input checked="" type="checkbox"/> transfers	

Properties

Variables

Name	transfers
Label	
Type	float
Format	%0.0g
Value Label	

Abobo

Limitations

1. Missing data remain
2. Outliers due to discretionary decisions form central government
3. Need to disaggregate some variables

Using the Data and Lessons

Example

- Sanogo and Moummi (2015). Local government taxation and access to basic services in Côte d'Ivoire, CERDI, Working paper.

Lessons

- Need for training local administrations and to set up electronic system for data collection;
- Need to update the existing data (from 2012-2015);
- Need to extend the data (1990 to 2000);
- Valuable to undertake such work in developing countries where not existing.



Thank you for your attention!

Comments and suggestions are
highly appreciated

