# 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.

ICTD

# 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

ICTD

# 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

ICTD

# 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

ICTD

# 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 understand 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.

ICTD

# Outline

1. Motivations

2. Methodology

3. Outputs

4. Limitations

5. Using the data and lessons

ICTD

# 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)

ICTD

# 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** | **Shared Tax** |
| - Property Tax | - Property Tax |
| - Patent and Licenses | - Patent and Licenses |
| - Synthetis Tax | - Synthetis Tax |
| | - Vignettes |
| **Own Revenue** | |
| - Flat rate tax | **Own Revenue** |
| - Civil Registration | - Flat rate tax |
| - Markets fees | - Civil Registration |
| - Lease fees | - Markets fees |
| - Fees for the use of public facilities | - Lease fees |
| - Other own revenue | - Fees for the use of public facilities |
| **Other revenue** | - Other own revenue |
| - Vignettes | |
| - Grants from central government | **Other revenue** |
| - External funds | - Grants from central government |
| - Ordinary Reserve Fonds | - External funds |
| - Miscellaneous revenue | - Ordinary Reserve Fonds |
| - Loans | - Miscellaneous revenue |
| | - Loans |

ICTD

# Outputs (Data)

# 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.

ICTD

# Thank you for your attention!

Comments and suggestions are
highly appreciated