Measuring Untapped Revenue Potential in Developing Countries: Cross-Country Frontier and Panel Data Analysis

Zeljko Bogetic Dominik Naeher Raghavan Narayanan
World Bank University of Göttingen World Bank/EBRD

6 September 2023
Motivation

▶ Assessing countries’ performance in domestic revenue mobilization (DRM)
  ▶ Common indicators focus on achieved outcomes: revenue-to-GDP ratios.
  ▶ Typically suggest that low-income countries perform poorly in DRM.

▶ Shortcomings of these indicators
  ▶ Do not account for differences in fundamental economic conditions and constraints across countries.
  ▶ Underestimate the performance of governments in countries with less-favorable enabling conditions for DRM.
  ▶ Provide little insights on how efficient countries are in DRM for given domestic conditions.
This Paper

▶ Goals
1. Measure the DRM performance of countries when accounting for differences in domestic economic conditions.
2. Estimate each country’s untapped potential for further increasing DRM under current conditions.

▶ Approach
- Non-parametric frontier analysis based on data envelopment analysis (DEA).
  - DEA output variable: Government revenue (% of GDP).
  - DEA input variable: Composite index as proxy for enabling economic conditions.
Graphical Illustration of DEA
Graphical Illustration of DEA
Graphical Illustration of DEA

Efficiency gap
Data Envelopment Analysis

- **Benefits**
  - Efficiency estimates are based on immediate peers, not whole sample.
  - No assumptions on the functional form of the production function.
  - Results are easy to interpret:
    - **DEA efficiency scores**: normalized to range from 0 (inefficient) to 1 (efficient).
    - **Untapped potential**: how much additional revenue should each country be able to achieve given its domestic economic conditions?

- **Limitations**
  - Non-parametric/descriptive nature: no information on root causes of inefficiency
  - No direct welfare implications
Data

- **DEA output variable**: Government revenue, excluding grants (% of GDP)
  - from ICTD/UNU-WIDER Government Revenue Dataset (2020)

- **DEA input variables** (aggregated to 'DEA input index'):
  - GDP per capita (+)
  - Share of agriculture in GDP (-)
  - Trade (% GDP) (+)
  - Age dependency ratio (-)

- **Sample**
  - 118 countries: 25 LIC, 45 LMIC, 48 UMIC
Estimated Frontier for LICs and MICs, 2016-2019
Box Plots of DEA Scores by Income Group and Region
## DEA Results by Income Group and Region

<table>
<thead>
<tr>
<th>Income Group</th>
<th>2016-2019</th>
<th>Change in efficiency since 2012-15</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Rev.</td>
<td>DEA Input Index</td>
</tr>
<tr>
<td>LIC</td>
<td>13.55</td>
<td>0.09</td>
</tr>
<tr>
<td>LMIC</td>
<td>20.77</td>
<td>0.15</td>
</tr>
<tr>
<td>UMIC</td>
<td>26.36</td>
<td>0.21</td>
</tr>
</tbody>
</table>

### Region:

<table>
<thead>
<tr>
<th>Region</th>
<th>2016-2019</th>
<th>Change in efficiency since 2012-15</th>
</tr>
</thead>
<tbody>
<tr>
<td>EAP</td>
<td>22.21</td>
<td>0.19</td>
</tr>
<tr>
<td>ECA</td>
<td>30.46</td>
<td>0.23</td>
</tr>
<tr>
<td>LAC</td>
<td>23.51</td>
<td>0.19</td>
</tr>
<tr>
<td>MENA</td>
<td>22.75</td>
<td>0.21</td>
</tr>
<tr>
<td>SA</td>
<td>17.29</td>
<td>0.17</td>
</tr>
<tr>
<td>SSA</td>
<td>16.52</td>
<td>0.09</td>
</tr>
<tr>
<td>Average</td>
<td>21.52</td>
<td>0.16</td>
</tr>
</tbody>
</table>
Conclusion & Possible Extensions

- New method for measuring countries’ DRM performance
  - Non-parametric approach, easy to interpret
  - Complement insights from existing measures based on achieved outcomes only

- Inform deeper country-level analyses
  - DEA input index: Identify most binding constraint in each country
  - DEA frontier: Identify global country peers

- What drives differences in DRM efficiency across countries?
  - Use DEA scores as dependent variable in regression framework
  - with political/institutional factors as explanatory variables

- Does international support for DRM target countries with high untapped potential?
  - Use DEA scores as explanatory variable in regression framework
  - with foreign aid flows as dependent variable
Summary of Results

Findings

- Globally, revenues are estimated to be at 62% of their potential across LICs and MICs.
- Many LICs (including in SSA) perform close to the efficient frontier for DRM, despite featuring low revenue-to-GDP ratios.
- SA and LAC show the strongest increases in DRM efficiency from 2012 to 2019, indicating that various countries in these regions were catching up with the frontier.
- Looking only at achieved revenue-to-GDP ratios can be misleading for drawing conclusions about DRM efficiency, and for how much more revenue a country can potentially raise, given its economic structure.

Robustness

- Main findings are robust to changes in indicator selection/weighting/aggregation.
- Limited data quality: Quantitative magnitudes should be interpreted with caution.
Parametric vs. Non-parametric Methods

- Different methods for efficiency analysis (Thanasoulis 1993; Sickles and Zelenyuk 2019)
  - Parametric/econometric: Regression analysis (OLS), Stochastic frontier analysis (SFA).
  - Semi-parametric: Neural network analysis.
  - Non-parametric: Data envelopment analysis (DEA), Free disposal hull (FDH).

- Benefits of DEA
  - No assumptions on the functional form of the production function.
  - Efficiency estimates based on immediate peers, not average.
  - Can handle multiple inputs/outputs measured in different units.
  - Quantitative results are easy to interpret.

- Limitations of DEA
  - “Black box”: no information on root causes of inefficiency (causality).
  - No statistical significance tests.
  - Sensitive to variable selection and measurement error.
Economic Fundamentals and Revenues, 2016-2019

- Robustness
- Parametric vs. Non-parametric

Supplementary Materials: 

- DEA Input Variables
- Economic Fundamentals and Revenues, 2016-2019

Graphs showing relationships between revenue (excl. grants as % GDP) and:
- Log GDP per capita
- Share of Agriculture in GDP
- Trade Openness
- Age Dependency Ratio

Dominik Naeher