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.

Motivation
○●○

Methods & Data
○○○

Results
○○○

Conclusion
○

Appendix

Achieved Outcomes vs. Enabling Conditions, 2016-2019
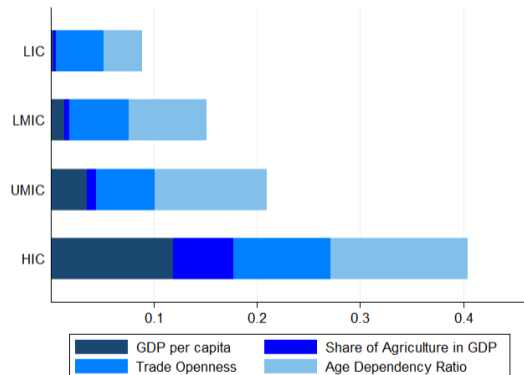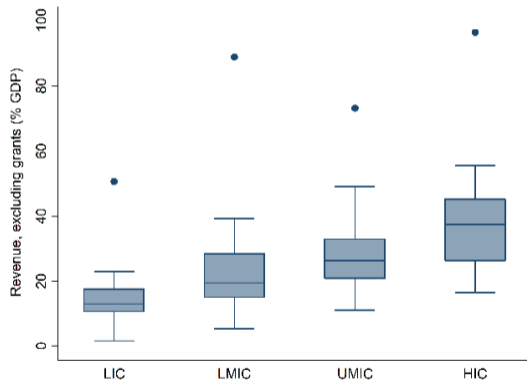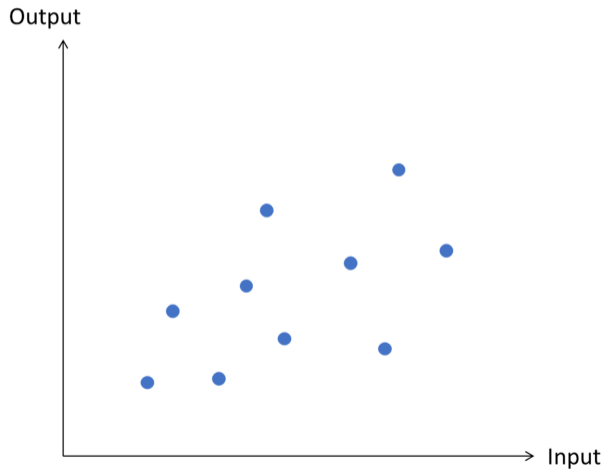
This Paper

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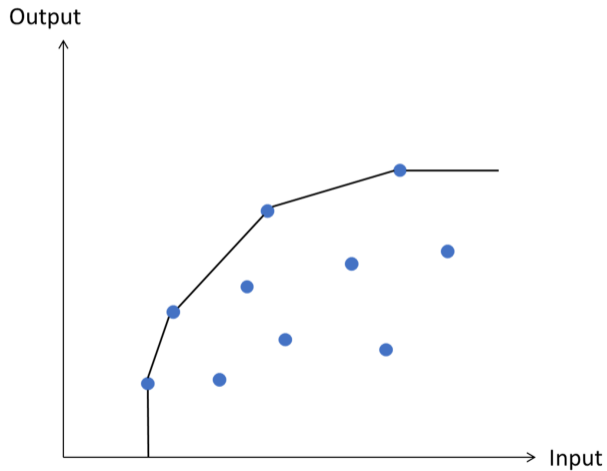
▶ Approach
   ▶ Non-parametric frontier analysis based on data envelopment analysis (DEA).
      ▶ DEA output variable: Government revenue (% of GDP).
      ▶ DEA input variable: Composite index as proxy for enabling economic conditions.

Motivation
○○○

Methods & Data
●○○

Results
○○○

Conclusion
○

Appendix

Graphical Illustration of DEA

Motivation
○○○

Methods & Data
●○○

Results
○○○

Conclusion
○

Appendix

Graphical Illustration of DEA

Motivation
OOO

Methods & Data
●OO

Results
OOO

Conclusion
O

Appendix

Graphical Illustration of DEA

Output

Efficiency gap

Input

Motivation
○○○

Methods & Data
○●○

Results
○○○

Conclusion
○

Appendix
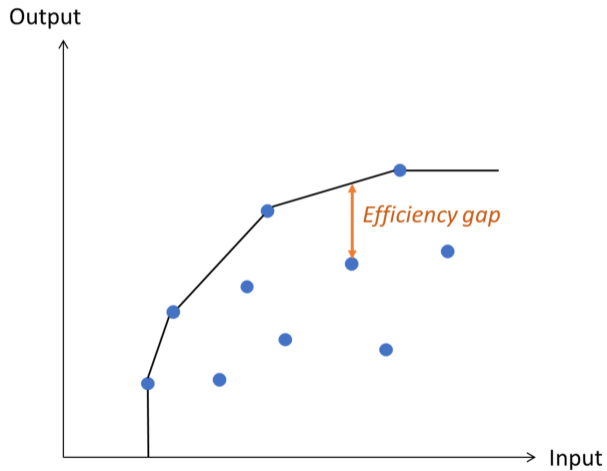
Data Envelopment Analysis

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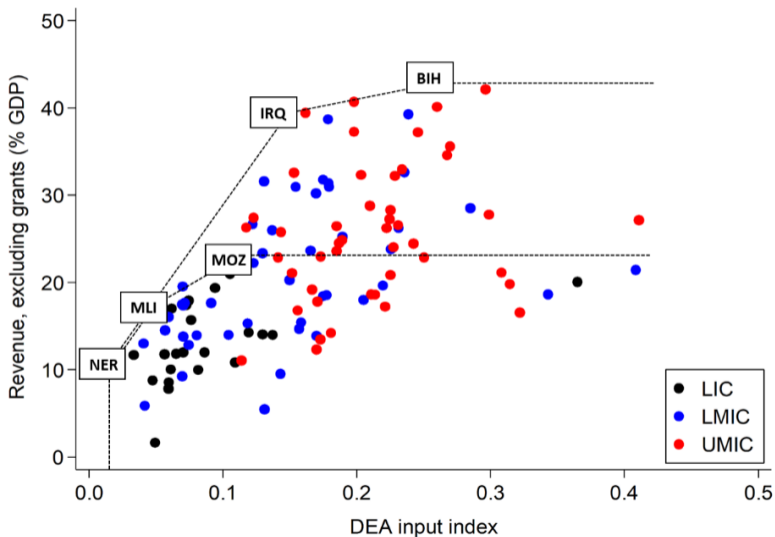
▶ Limitations
   ▶ Non-parametric/descriptive nature: no information on root causes of inefficiency
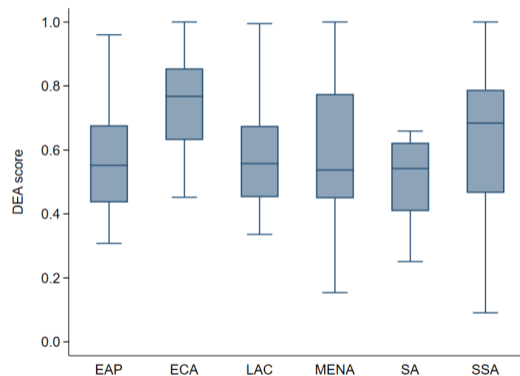   ▶ No direct welfare implications

Motivation
000

Methods & Data
00●

Results
000

Conclusion
0

Appendix

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
  - ▶ Analyze 4-year average values: 2008-2011, 2012-2015, 2016-2019

Motivation
○○○

Methods & Data
○○○

Results
●○○

Conclusion
○

Appendix

Estimated Frontier for LICs and MICs, 2016-2019

Motivation
ooo

Methods & Data
ooo

Results
oeo

Conclusion
o

Appendix

Box Plots of DEA Scores by Income Group and Region

Motivation
○○○

Methods & Data
○○○

Results
○○●

Conclusion
○

Appendix
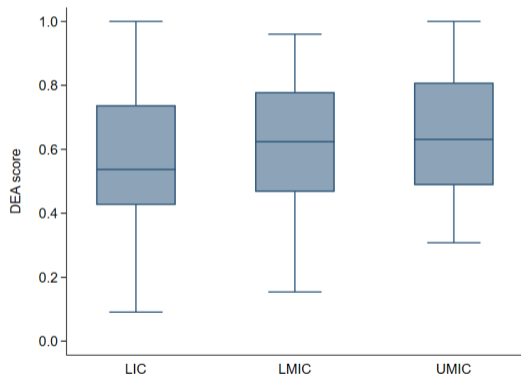
DEA Results by Income Group and Region

	2016-2019			Change in efficiency since 2012-15
	Rev.	DEA Input Index	DEA Score (upper bound)	
Income Group:				
LIC	13.55	0.09	0.58 (0.69)	+0.03
LMIC	20.77	0.15	0.62	+0.02
UMIC	26.36	0.21	0.65	+0.00
Region:				
EAP	22.21	0.19	0.57	+0.02
ECA	30.46	0.23	0.74	-0.01
LAC	23.51	0.19	0.59	+0.03
MENA	22.75	0.21	0.57	-0.03
SA	17.29	0.17	0.51	+0.05
SSA	16.52	0.09	0.64	+0.02
Average	21.52	0.16	0.62	

Motivation
○○○

Methods & Data
○○○

Results
○○○

Conclusion
●

Appendix

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

Dominik Naeher