

Open data and impact of natural resources: potential and limitations

Fernando Aragon

Simon Fraser University (BC, Canada)

June 2015

The take-away

- Open data has contributed to research on the impact of natural resources
- But, we need to go beyond open data to answer some of the remaining big questions

What is out there?

There are several source of open data that can be used to assess impact of natural resources

1. Country level data

- Economic data: Penn Tables and many others
- Subjective measures of institutions and quality of government: Polity IV, ICRG, Transparency Intl.
- Conflict: UCDP/PRIO, COW
- Replication datasets from previous studies

What is out there?

2. Micro-datasets

- National Censuses
- Household surveys: LSMS, DHS
- Subnational level data: county yearbooks, municipality data
- Conflict: ACLED, PRIO datasets
- Mineral production: USGS and others
- Opinion polls: Latinobarometer, etc.
- Satellite imagery: weather, air quality, economic activity.

What is out there?

3. Not-so-open datasets

- Censuses micro-data
- Mining sites location and production: SNL Metals and Mining Database, Intelligence Mine (Infomine), ASPO dataset (oil)
- Public records under Freedom of Information legislation (in some countries)

What have we learnt?

Country level studies

1. Natural resources + bad institutions = negative outcomes (resource curse)
2. Problem bigger for some type of resources (point source-resources)
3. Volatility may also contribute to negative effects but Dutch disease may not be as important as we thought.

What have we learnt?

Micro (within-country) studies

1. Not much crowding out of manufacturing
 - Necessary condition for Dutch disease story
 - Agglomeration economies? Local manufacturing?
 - Pro-cyclical manufacturing

2. Impacts can be positive or negative: US vs Brazil
 - Is this due to institutions?

3. How rents are distributed seems to matter
 - Fiscal windfalls vs local demand shocks
 - Importance of supply-chain linkages

What have we learnt?

Micro (within-country) studies

4. Spillovers from resource extraction may be economically important
 - Environmental pollution, agglomeration economies

5. Significant effect on other relevant economic outcomes (beyond income and employment)
 - Productivity, education, health, inequality, etc.

Open data potential

- Open data has significantly contributed to this research program.
- Open data can be used to
 - Replicate/validate and deepen these results
 - Examine related questions
 - For example: role of agglomeration economies, effects on other economic outcomes,
 - Disseminate information among stakeholders

But (existing) open data are not enough...

1. Limited geographical and temporal scope

- Few developing countries with enough open datasets: for instance, Peru, Brazil, India, Africa?.
- Low spatial resolution: impacts usually at very low geographical level: villages, towns, counties.
- Micro-datasets available for only few years: difficult to exploit time variation, or examine long-term effects

But (existing) open data are not enough...

2. Several data not yet open or not systematically collected
 - Financial information: payments and transfers to local communities, government, land owners, etc.
 - Mining companies activities: agreements with communities, compensation and other payments, local procurement, salaries, local development projects, etc.
 - Environmental data: quality of air and water
 - Local institutions and political conditions: electoral outcomes, corruption, local govts. capacities and resources, etc.

But (existing) open data are not enough...

3. Limited sources of (good) variation

- Not all research projects constrained by lack of (open) data
- Instead, lack of exogenous sources of variation!
- Not many natural experiments, limited ability of DiD / IV strategies.
- We may need to think seriously about experiments.

Why does it matter?

- Several important questions hard to answer with existing open data
 1. Which specific policies and reforms should be implemented to increase positive effect of resource windfalls?
 2. How to enhance supply-chain linkages?
 3. How should resource rents be distributed? Transfer to local governments, contract with communities, payments to individuals, etc.?

Why does it matter?

4. How to minimize local conflict?
5. How effective are environmental and industry regulations in reducing negative spillovers?

What can we do?

- Pushing for more open data
- Collect more data
- Experiments