Measuring State Self-Legibility in the Digital Age
Jonathan Phillips

**Motivation**

- Scott (1998) documents how states use their infrastructural and coercive power to make society more legible.
- Yet, in modern states the apparatus of the state itself is large, decentralized, and hard to control.
- The primary threat to political elites is often *intra-state* contestation, and the legibility challenge is reflexive.
  - If all governments have an incentive to increase self-legibility, where do they achieve it?
  - If not, why does legibility vary across the political terrain?
  - What are the consequences of being rendered legible to the state?

**Measures of Self-Legibility**

- Central government databases seek to register government assets.
- Legibility is high where all government assets are accurately recorded.
- Three strategies to measure legibility:
  - Comparing when assets enter government databases to when they're produced.
  - Measuring the amount of missingness in the data.
  - Cross-validation of two sources.

**When and Where do Assets Become Legible?**

- India's annual DISE census records the year a school is founded.
- A school becomes legible where it is founded at least two years before it enters the data for the first time, and then remains in the data until at least 2011.
- Since 2005, 335,402 schools have become legible.
- The central government has 'discovered' four times as many schools as have been built.

**Legibility: % Of Schools Existing in 2005 and Known in 2011 that were also Known in 2005**

- The areas with the greatest increase in legibility are highly populated but more rugged areas.
- The infrastructure of schools that become legible is only marginally weaker than those already 'on the map'.
- Government allocations fail to improve upon becoming legible.
- But inputs dependent on national government start off weaker and converge faster.

**Do Governments know Less than INGOs about their own assets?**

- In 2011, the Nigeria Government conducted a census of health facilities in the North. In 2012, UNICEF and the WHO did the same.
- 80% of observations in the UNICEF data appear in the government data.
- 89% of observations in the Government data appear in the UNICEF data.
- But the pattern of spatial overlap reflects local conflicts and governance capacity.
- Local partners and capacity are vital for central legibility.

**Potential Applications**

- An objective measure of state capacity/legibility, since knowing about assets is often a requisite for managing and improving them.
- An explanatory variable for changes in public service performance and political engagement.
- A diagnostic tool for improving databases and measurement.

---

Prepared for Harvard Government Department Poster Session.