

Market access and firm performance: evidence based on GIS analysis of road network & manufacturing plant data of India by Sonnath Sharma *et al.*

Discussant comments by Mark Roberts
Lead Urban Economist, World Bank

16 May 2022

Brief recap of paper

- Central research questions:
 1. Are Indian firms that have better market access more profitable?
 2. Does better market access for a state's firms reduce the dispersion in returns on assets?
- Questions are addressed by combining micro firm data from Annual Survey of Industries for 2001-2015 with a measure of market access (MA)
- MA measure constructed at district level using Open Street Maps data and based on driving distances through road network
- For Q2, state level panel regressions (TWFE) - a measure of dispersion of returns on MA (average across districts) and controls
- **Main finding:** Better market access → ↓ dispersion of returns

Overall impressions

- **Interesting and important research questions:**
 - Lots of previous research on impacts of \uparrow transport connectivity on wages & local incomes, but less on impacts on firm profitability
 - Little previous research on impacts of \uparrow connectivity on efficiency of resource allocation across firms
- **Great underlying data!** Authors do a nice job of assembling a panel of 652,864 enterprises & district-level MA measure based on combination of OSM (\Rightarrow not only capturing major roads) & lights data
- **Important policy implication:** Finding that \uparrow MA \rightarrow \downarrow dispersion of returns strengthens case for investing in roads
- **But** paper is still clearly an early draft \Rightarrow natural room for improvement

Comments, questions, suggestions

- Paper only currently addresses one of its two research questions - no results on whether better MA is associated with improved firm profitability
 - Should be easy to resolve given data - estimate firm-level regressions
- Construction of market access measure:
 - $MAI_i = \sum_j (NTL_j / (d_{ij})^\theta)$
 - $\theta = 2$ is assumed without discussion, but why 2? Large literature on estimating θ that could refer to. Room for robustness checks
 - d is measured as driving *distance*, but why not driving *time*?
 - Consider using major population centers instead of district centroids as points between which d_{ij} is calculated
 - What about international market access? Could measure as, e.g., distance to nearest international port
 - Quality of OSM may be geographically uneven - has this been investigated?

Comments, questions, suggestions

- Why state-level regressions?
 - Aggregates much of the underlying richness in the data...
 - Is this because ASI lacks geographic identifiers below the state???
 - For research Q1, why not firm level regressions? For research Q2, possible to consider district-level regressions also???
- Clarify where the identifying variation is coming from:
 - $ROADDispersion_{s,t} = \beta_1 MAI_{s,t} + (\text{state FE}) + (\text{time FE}) + v_{s,t}$
 - Inclusion of state FE $\Rightarrow \beta_1$ identified based on within-state variation in MAI
 - Given that NTL is measured only for 2014 \Rightarrow variation in MAI must come from temporal (annual) variation in road network coverage
 - Be useful to describe & give an overview of this variation
- Suggest less detail on GIS techniques (which are standard), more on identification strategy, discussion of results & robustness