Small Area Estimation of Poverty

Discussion of a presentation by Prof. J.N.K. Rao

By Peter Lanjouw (World Bank)

Outline of comments

- Some remarks on EB vs ELL approach
 - Importance of data configuration
 - Some issues common to EB and ELL
 - Area Homogeneity Assumption
 - Handling "intrahousehold correlation"
- Validation of ELL approach using Brazil data
 - Area homogeneity assumption
 - Estimates versus "truth"
 - Sensitivity to assumed application of location effect

Data Configuration

- Origins of ELL approach
 - Outgrowth from a debate/challenge from the Statistical Office of Ecuador (INEC) in mid 1990s.
 - Focus from the start was on developing country imperatives:
 - Designed to "fit" with the data configuration common in these settings.
 - Living Standards Measurement Surveys are an important source of data in developing countries.
 - Such surveys typically have relatively small samples, and follow a complex design.
 - LSMS for Ecuador: 463 sampled clusters (EA's) out of roughly 2300 nationwide.
 - Data collected in 53 out of 915 parroquias (domains).
 - No possibility of linking survey households to census.
- Refinements offered by the EB approach thus apply to roughly 5% of domains of interest in Ecuador
- In some countries, e.g. India, survey is stratified at the domain level (the district).
 - This would appear to suit EB approach better. However, in such cases, ELL would apply a domain level fixed effect.
 - Not clear whether "test" of ELL in Molina-Rao applies this fixed-effect specification.

Issues Common to EB and ELL

- "Area Homogeneity Assumption"
- Both EB and ELL method assume that conditional distribution of **y** given **x** in small area A is the same as in larger region, R.
- Homogeneity assumption can fail when:
 - Slope coefficients are not the same in A as in R
 - There exists an unobserved small-area effect
- We examine this assumption empirically in Brazil (see below).

Handling "location effect"

- In EB approach intra-household correlation of residuals is assessed at the domain, not cluster, level.
 - When there are few domains with sample data this can be a problem
 - Recall, only 53 domains in Ecuador contain sample data
 - With separate, stratum-level, models this leaves very few domains per model from which to draw a location-effect.
- In ELL, the correlation is assessed at the cluster level.
 - Many more clusters in the sample data.
 - However, key question now, is at what level to apply intrahousehold correlation effect in the simulation stage

Tarozzi and Deaton (2009) Critique

Census Structure



 Conservative Option: apply location effect (estimated at cluster level) at domain level

Testing the Poverty Mapping Methodology: Brazil

- Elbers, Lanjouw and Leite (2008) consider Minas Gerais, Brazil
- Brazil Census collects income data
 - Thin round (87.5%) collects single-question measure of household income
 - Thick round (12.5%) collects more detailed info.
 - Neither are judged reliable for an 'official' poverty map.
- We focus on Minas Gerais (for computational ease)
 - 606,000 households in 12.5% sample (out of 4.8m)
 - 12.5% sample covers all 853 municipalities in Minas Gerais

Minas Gerais: Brazil within Brazil

• Per Capita Income



Testing the Poverty Mapping Methodology: Brazil

- We draw 41 synthetic surveys from Census sample
 - 21 mimic sample design of POF 2,800 households
 - 13 households per cluster/EA
 - 241 EA's in about 151 Municipalities
 - 20 mimic sample design of PNAD 12,000 households
 - 16 households per cluster/EA
 - 779 EA's in 123 municipalities
- We produce 41 poverty maps for Minas Gerais
 - We estimate location effect at EA level
 - We apply location effect at Municipality level
 - Tarozzi and Deaton's conservative approach

Testing Area Homogeneity in Brazil

- Differences in returns
 - Apply one model in full census sample (specified in one PNAD sample)
 - Re-estimate model separately in each municipality (again in PNAD sample)
 - Compare predicted municipality-level income



Testing the Methodology in Brazil

• Municipal level Poverty Estimates versus "Truth"



Testing the Methodology in Brazil

• Municipal level Poverty Estimates versus "Truth"



• Overly Precise estimates?



• Overly Precise estimates?



Are poverty estimates usable?



Proportion of municipalities with significant HCR change versus Confidence Interval

• What if "standard" approach had been applied? (cluster effect to EA)



Proportion of municipalities with significant HCR change versus Confidence Interval