MAKSWELL EU

Use of big data for measuring poverty and well-being

Chair: Fabio Bacchini

Discussion Slides

Natalie Shlomo
University of Manchester

Natalie.Shlomo@Manchester.ac.uk

Three presentations from MAKSWELL partners: Germany, UK and Italy

- **Small area poverty indicators adjusted using local price indexes**: head count ratio estimation accounting for sub-national regional price indices and other indicators derived from big data

- **Alternative spatial data sources for small area estimation in developing countries**: using remote-sensed data as covariates in spatial small area estimation

- **Measuring well-being at local level using big data**: using satellite imagery in combination with other data sources

Commonalities

- All papers demonstrate the potential for using big data in official statistics systems for measuring poverty and well-being at local areas
- Big data sources included in these papers: retail scan data, transaction house prices, online web data (NUMBEO), remote sensing data (night time lights, elevation, accessibility to urban centers), satellite imagery
- Big data sources have the potential to incorporate spatial correlations for improving the estimation of indicators through model-based methods either through the mean or through the error structure of the models

General comments:

- The challenges for considering big data sources as fit-for-purpose in the production of official statistics: availability, continuous, consistent and accessible with known quality measurement
- Requires coordination and cooperation with commercial and public institutions holding these data
- We must continue to invest in research on using new forms of data and disseminate results as experimental statistics
- Integration of big data sources with high-quality survey data is preferable compared to solely relying on these sources of data alone

General comments:

- Cooperation with academic institutions and investment in developing the necessary skill sets to be able to integrate new forms of data in statistical system, i.e. advanced complex model-based estimation techniques
- Need better quality evaluations to meet the standards of official statistics
- All examples shown here demonstrate that although the sources are prone to quality issues and selection bias, they have the advantage that they are all *identifiable*, i.e. they can be attributed to a point in time/ space, and hence their potential use for official statistics

Small area poverty indicators adjusted using local price indexes

- Congratulate the authors on innovation in using differential provincial poverty lines
- Is there a more direct way of measuring poverty rather than through the Head Count Ratio and the complex models used
- To obtain the provincial HCR through the FH model, the estimated provincial poverty line is used. This is itself derived from an hedonic regression model: leads to concern for endogeneity as independent variables are themselves based on models

Alternative spatial data sources for small area estimation in developing countries

- Can remote sensor data be a reliable and fit-for-purpose data source to be used in a national statistical system?
- Descriptive statistics would be helpful to understand which model may be appropriate – what are the correlations and spatial correlations between dependent and independent variables?
- Direct estimates and variances for out-of-sample are imputed. Is there concern for confounding of the results of the small area estimates?

Measuring well-being at local level using big data

- When using satellite imagery, the modifiable areal unit problem (MAUP) can bias results and further research is needed on how to use these sources of data in official statistics
- Satellite imagery prone to particular quality issues compared to other data sources, what are some of the issues that need to be considered?

Discussion