Multidimensional Poverty: One Aim, Three Approaches, Quite a Few Different Results

Julia Karpati and Chris de Neubourg

IARIW-World Bank Conference;
Washington DC, November 2019
Session 3A

Discussed by Roberto Zelli - Sapienza University of Rome

Similarities and Differences of Three Multidimensional approaches

- Consensual deprivation (EU-wide measure of deprivation)
- Multidimensional Poverty Index (MPI) (Global MPI - UNDP)
- Multidimensional Overlapping Deprivation Analysis (MODA) (MODA - UNICEF)

In a policy-oriented perspective, different theoretical framework and targets influence:

- choice of indicators and dimensions;
- unit of analysis vs unit of measurement;
- aggregation and weighting schemes;
- deprivation thresholds.

Indicators and dimensions I

The three representative measures of the three approaches (EU material deprivation, Global MPI, MODA) follows a **"counting" approach**, with significant distinctions.

- Notation. Let us assume to have $i = 1, \ldots, N$ individuals of a population, $d = 1, \ldots, D$ dimensions of deprivation, and, for each dimension d, $j = 1, \ldots, n_d$ attributes of deprivation, w_j^d the weight associated to attribute j .
- Attributes can be continuous or discrete or binary. Indicators or items are binary variables that (eventually) dichotomize the attributes. All the three approaches use binary indicators.
- Indicators or items:

$$y_{i,j}^d = \begin{cases} 1, & \text{if } x_{i,j}^d \leq \lambda^j, \\ 0, & \text{otherwise.} \end{cases}$$

Indicators and dimensions II

- Number of dimensions:
  - D=1 EU material deprivation (MD)
  - D=3 Global MPI
  - D=# depend on the life-stage (max 8) CrossCountry-MODA

choice of the dimensions theoretically-driven.

For each dimension:
Deprivation score of each individual results in the weighted counting value

$$S_i^d = \sum_{j=1}^{n_d} w_j^d \cdot y_{i,j}^d, \quad w_j^d = \frac{1}{n_d}$$

Overall (multidimensional) deprivation score:

$$S_i = \frac{1}{D} \sum_{d=1}^{D} S_i^d$$

Deprivation cut-off λ to establish who is overall deprived and who is not (intermediate approach):

$$MP_i = \begin{cases} 1, & \text{if } S_i \geq \lambda, \\ 0, & \text{otherwise.} \end{cases}$$

Union Approach for each dimension (no need of weighting, constraint on the number of indicators):

$$S_i^d = \begin{cases} 1, & \text{if at least one } y_{i,j}^d = 1, \ i.e. \ \sum_{j=1}^{n_d} y_{i,j}^d > 0 \\ 0, & \text{otherwise.} \end{cases}$$

Overall (multidimensional) deprivation score:

$$S_i = \sum_{d=1}^{D} S_i^d$$

Deprivation cut-off k to establish who is overall deprived and who is not: (CC-MODA union approach, but shows all)

$$MP_i = \begin{cases} 1, & \text{if } S_i \geq k, \\ 0, & \text{otherwise.} \end{cases}$$

$$1 \leq k \leq D \text{ where } \begin{cases} k = 1, & \text{union approach,} \\ 1 < k < D, & \text{intermediate approach,} \\ k = D, & \text{intersection approach.} \end{cases}$$

An individual can be identified as multidimensionally poor if he/she suffers from at least one dimension ("union" approach) or from all the deprivation indicators ("intersection" approach), or if the number of deprivations is above a fixed value, say k, with $1 < k < D$ ("intermediate").
Analysis indicator by indicator and analysis of all possible overlapping across dimensions.

Different weights or equal weights?

Long debate (Decancq and Lugo, ER 2013)
... distinction between:

- weights assigned to each indicator within the same dimension to measure the dimension
- weights assigned to each dimension to identify the overall status of deprivation

- weights to indicators: instrumental for measuring a latent variable (the dimension)
- weights to dimensions: they may reflect normative statement or preferences expressed by the society

Each dimension can be seen as a latent variable: it is not directly observable, but it manifests itself through a set of attributes.
The latent variable can be measured only indirectly using a model.
Different types of *latent variable models* according to the type of observed and of latent variables (Collins and Lanza, 2010):

Observed / Latent	Continous	Discrete
Continous	Factor Analysis	Latent Profile (mixture)
Discrete	Item Response Theory	**Latent Class Models**

Validation

- **trade-off** between the interpretability (that reflects the ease with which the user may understand and properly use and analyse the data) and accuracy (that can be increased by the use of multivariate analysis to identify the data structure).
- Importance of **sensitivity analysis** to assess the robustness of the indexes.
- Ideally, all potential sources of uncertainty should be addressed: selection of individual indicators, data quality, dichotomization, weighting, aggregation method, thresholds, ...

Latent class models: our experience (Dotto, Farcomeni, Pittau, Zelli, JRRS-A 2019

Built a dynamic latent Markov model that classified individuals by their current and inter-temporal deprivation status:

- each individual is classified into latent classes of poverty with a certain probability membership;
- this classification is based on the observed profiles of binary indicators, without resorting to aggregation and weighting of the indicators;
- we derived a weighting scheme in order to evaluate ex-post whether the equal weights assumption adopted by Eurostat was coherent with the classification

When the number of lacking items was three or more, the probability of being poor was extremely high irrespective of which items are missing, so the EU MD works well with a cut-off of $k = 3$.
However, there were specific combinations of two lacking items that lead to high probabilities to be poor, and other ones that lead to low probabilities to be poor. The raw sum cannot distinguish such situations.
Similarly, there could be configurations of three lacking items that not necessarily lead to high probability of being poor.

- A sound theoretical framework is the primary ingredient.
- Nevertheless the statistical analysis could help in thinking about the framework used (a sort of "backward thinking": "does the theoretical derived model provide a good fit to the data? What the lack of fit tells about the conceptual definition of the composite of the indicators chosen for it? What concept would the available indicators good measure of?" Proving an answer to these questions gives robustness and coherence to the index.

No index can be better than the data it uses. But this is an argument for improving the data, not abandoning the index.
(UN, Human Development Report, 1992).

Data Issues

1. One database or at least with a common unique identifier for hh or individuals.
   - a single source cannot have all the information required to measure multi-dimensional poverty
   - ? use of statistical matching techniques (PSM) in order to derive a unique synthetic dataset in which all the variables (coming from different sources) are jointly available.
2. Cross-sectional vs longitudinal data.
   - the vast majority of empirical studies in this domain assesses static multidimensional deprivation, due to data limitations
   - panel data allows distinction between current and persistent multidimensional deprivation and mobility in and out of deprivation
   - ? introduction of a longitudinal component in a survey (rotating panel designs,...)
   - ? in absence of (reliable) panel data, building pseudo(synthetic)-panel can offer a solution to address inter-temporal issues