Expert Group Meeting on Data Disaggregation
27-29 June 2016
New York
Expert Group Meeting on Data Disaggregation

New York, 27-29 June, 2016

Olivier Dupriez, Lead Statistician, World Bank
Disaggregation of, and by, income poverty

• % of poor at national and international poverty lines are SDG indicators
  • To be disaggregated by urban/rural, age, sex, and employment status
• Consumption-based measure in most low-income countries
• Source of data: household surveys (+ price data / PPPs)
• Disaggregation of key socio-economic indicators by poverty status (NPL) highly relevant for targeting and monitoring “pro-poor” policies
  • Multi-topic surveys → poverty profiles w/ disaggregation of various indicators
• Not much breakdown by poor/non-poor in time series databases
  • Poverty numbers usually not reported by age/sex/employment status
What the World Bank is doing

• Produce global poverty estimates (at international poverty line)
  • Use microdata when available (ex-post harmonization work), otherwise grouped data
  • Currently not disaggregated (except urban/rural for very few countries)

• Support data collection (“ending data deprivation” 2015-30 survey initiative)
  • Multi-topic surveys for maximum relevance and poverty profiling

• Participate in development and implementation of improved survey methods
  • Survey design, data capture/editing, data curation and dissemination

• Provide tools and guidelines
  • CAPI Survey Solutions, data documentation and dissemination, and others

• Research work
  • E.g., survey-to-survey imputations, poverty mapping, synthetic data, micro-simulation

• Data dissemination (e.g., WDI, Microdata Library)
(Some of) the challenges for disaggregation

• Poverty is measured at the household level, not individual level
  • Not ideal for disaggregation of/by individual-level indicators
  • No plan to change that

• Comparability (within and across countries) and reliability of data
  • Need better methods for measuring consumption / treating outliers / adjusting for non-responses, etc. (being developed)
  • No consensus on definition of “urban”

• Demand for larger samples
  • Issues of data quality, cost of surveys → not always recommended
  • Alternatives: survey-to-survey imputations, complex sampling, use new data sources

• Not all groups of population are / can be covered in “standard” surveys
  • Issues of sample size, sample frames, questionnaire design, cost, risk

• Data availability and accessibility (need access to microdata)
Data quality/comparability – An example

• Non-response may have impact on disaggregated estimates
• Solution (?) in sample calibration
Priorities for next steps

• Research on better measuring income / consumption
  • On survey design, sampling, data capture / editing / analysis
  • Must be a coordinated effort (→ ISWGHS)

• Innovation: develop/test/improve modeling methods and tools
  • Poverty maps/small area estimates, data fusion, synthetic data, micro-simulation

• Funding for data collection in low-income countries
  • Fill data gaps, with better data integration (consistency across data sources)

• More open dissemination of fully-documented microdata
  • Not all disaggregation will/can/should be made by statistics offices
  • Need to promote a global (responsible) data liberation