



New data challenges – estimation techniques

Céline Rouquette, IEA

Paris, October 2 & 3

- In an ideal world, statisticians could collect all the data they need.
- However, this is often not possible:
 - Data collection has a heavy cost – money and staff time to process
 - Conditions are not there to collect data (absence of register of the units, of legal basis...)
 - Topic is sensitive – you cannot collect directly
- And even when you have collected data, you need estimation techniques
 - To extrapolate from a sample survey
 - To estimate the value you need based on administrative data
 - To interpolate time series between two data collections
 - To model what cannot be easily observed from proxies

- Quick estimation of electricity monthly consumption
 - Survey only a handful of major distributors, and extrapolate given their weight in distribution the previous year
- Measure of fuel wood used by households to heat their dwelling
 - Most cannot tell exactly (invoice absent or incomplete) or have no clue
 - Estimate from surface of dwelling and experts knowledge of average consumption?
 - Use a « donor » from the households who know, for those who do not?
- Breakdown of sectoral energy end use by detailed uses (heating, cooling, electric appliances...)
 - Collect information on appliances and their characteristics and extrapolate with average consumption?
 - Use data from smart meters and modelling equations?

- Some fuels are not covered
 - Example: Biomass data are not covered in China energy balances, nor collected and published by India.
 - For China we approximate production by end use in households as reported by a survey by Tsinghua University.
 - For India we used rural population growth as a trend, now we use WHO data reliance on biomass for cooking.
- Energy balances are not published annually
 - Example: for Ethiopia we obtained from the country full balances for 2014 and 2015, but in 2018 for 2016, only an electricity balance.
 - We received the AFREC and UNSD questionnaires though.
 - As there is a one-year lag, and an issue with fiscal year not being calendar year, we used GDP or population growth to extrapolate.

Ideas for discussion

- How can we help our member states to improve their data when basic data are missing?
 - Should we provide a database of good estimation practices?
 - Should we harmonise how we estimate for missing values?
 - How can we share sources of information and good estimations techniques?

