Introduction to the Ghana Living Standards Survey, Round 5

Training Workshop on the Commitment to Equity Methodology
CEQ Institute and The Ministry of Finance
Accra
February 7-10, 2017
About Household Surveys

• How GSS carries them out
  • draw a random sample of households
    • stratified by populations of interest
    • clustered, for economy
  • ask them a lot of questions
    • see the questionnaire

• Why GSS carries them out
  • useful to establish consumption patterns to establish weights for the consumer price index
  • useful to estimate (or describe) the distribution of income and/or consumption in Ghana
    • poverty
    • inequality
  • keeps researchers busy
About GLSS-5

• How the data are structured
  • Many files (look at subdirectories)
  • More or less follow the structure of the questionnaire
  • Some are “household-level” data
    • one record per household
  • Some are individual-level, and some are community-level

• Skip patterns and missing values
  • look at section 4, part A, for example
  • Stata does not behave well with missing values, so be careful!
About Sampling Weights

• Not every observation in the GLSS is equally valuable
  • urban areas are “over-sampled” because it is easy to interview there
  • there could be other over-sampled groups
    • for example, a small ethnic minority
    • but there are no other over-sampled groups in GLSS data

• Because it is more likely that an urban household is chosen for the sample, it gets less weight than a rural one

• Intuition: the rural households selected must “represent” more households than the urban households selected

• A simple example
About Clustering and Stratification

• Clustering tends to increase the standard errors of things we estimate with survey data
  • households near each other are more similar to each other than they are to other households
  • reduces the variation in the sample

• Stratification can reduce the standard errors, but usually does not in GLSS-type surveys

• Stata has very easy-to-use commands to take this into account
What Does GLSS-5 Have for CEQ?

- Market Income
- Contributory Pensions
- Market Income plus Pensions
- Direct Transfers
- Gross Income
- Non-Taxable Income
- Taxable Income
- Direct Transfers
- Disposable Income
- Direct Taxes
- Net Market Income
- Direct Transfers
- Indirect Subsidies
- Consumable Income
- In-Kind Transfers (Education, Health)
- Copayments, User Fees
- Final Income
How Do We Go About the Analysis?

- We must dig up all the information we can about these income concepts and their components from GLSS
- We must clean it up and save it in one dataset
- We can then apply pre-programmed commands to do much of the statistical work