Introduction to Building Micro-Simulation Model for Policy Analysis in Stata

Objective of the course

Micro-simulation modelling is an essential analytical tool for assessing the distributional and welfare effects of macroeconomic and fiscal policies using household-level survey data. Unlike macro-structural frameworks such as DSGE and CGE models—which rely on representative agents and often mask heterogeneity—static micro-simulation models capture differences across households, labour groups, and income strata. This introductory course provides a rigorous, hands-on foundation in building STATA-based micro-simulation models, covering the full empirical pipeline from survey database construction and data transformation to behavioural rule specification and counterfactual policy simulations. Participants will generate policy-relevant indicators including Foster–Greer–Thorbecke (FGT) poverty indices, Gini coefficients, consumption-based welfare metrics, and fiscal incidence measures under simulated shocks such as subsidy reforms, carbon taxation, cash transfers, and climate adaptation investments. The course equips policy analysts, particularly in Ministries of Finance with practical skills to integrate household-level micro-simulation frameworks into broader macro-fiscal modelling architectures for analysing policy transmission channels to poverty, inequality, and household welfare. The course is delivered through guided lectures, live software demonstrations, and applied laboratory sessions using Stata.

- **Module 1: Foundations of Micro-Simulation Modelling and STATA Environment**
  - Conceptual Framework of Micro-Simulation Models – Static vs Dynamic Micro-Simulation, Behavioural vs Accounting-Based Models, and Policy Applications in Fiscal and Climate Analysis
  - Introduction to STATA Interface and Workflow – Command syntax and execution, Structure of DO-files and reproducible scripts, and Project folder architecture for modelling
  - Data Management Fundamentals – Importing data into STATA (.dta, .csv, .xls), Examining dataset structure (describe, summarize, codebook), and Data types and storage formats
  - Data Manipulation Techniques – Renaming variables (rename), Recoding categorical variables (recode), Data replacement (replace) and Missing value decoding (mvdecode)
  - Generating New Variables – Arithmetic transformations (gen), Group statistics (egen) and Conditional variable construction
- **Module 2: Empirical Database Construction from Household Surveys**
  - Structure of Household Survey Questionnaires – Consumption modules, Labour and employment modules and Income and transfer modules
  - Household Survey Compilation – Merging survey modules, Construction of household identifiers and Sampling weights and survey design
  - Building Analytical Micro-Datasets – Aggregation of expenditure components, Construction of per-capita consumption and Adult equivalence scaling
  - Harmonization of Survey Variables – Consistency across waves, Price normalization and Regional adjustments
- **Module 3: Data Transformation for Micro-Simulation Analysis**
  - Data Inspection and Cleaning – Distributional diagnostics, Skewness and kurtosis testing and Identification of missing observations

- Treatment of Outliers – Detection via boxplots and Z-scores, Winsorization, Trimming, Replacement strategies and Re-estimation checks
- Functional Transformations – Logarithmic transformations, Deflation and normalization, and Construction of equivalized welfare aggregates
- Preparation of Baseline Welfare Distribution – Income vs Consumption aggregates, Poverty line specification and Household stratification

- **Module 4: Empirical Micro-Simulation Modelling**
  - Design of Policy Simulation Frameworks – Counterfactual construction, and Static behavioural rules
  - Simulation of Policy Shocks – Tax reforms, Transfer policies, Energy subsidy removal and Climate-related fiscal instruments
  - Welfare and Distributional Analysis – Foster-Greer-Thorbecke (FGT) Poverty Measures, Poverty headcount ratio, Poverty gap index, and Squared poverty gap.
  - Inequality Metrics – Gini coefficient, Theil index, and Lorenz curves
- **Module 5: Policy Applications and Interpretation of Simulation Results**
  - Fiscal Incidence Analysis – Direct vs indirect effects and Redistribution impacts
  - Evaluation of Policy Trade-offs – Growth vs equity, and Poverty vs inequality outcomes
  - Sensitivity and Robustness Analysis – Parameter variation, and Alternative poverty lines
  - Reporting and Policy Communication – Interpretation of simulation outputs, Visualization of welfare impacts and Preparation of policy briefs

Mode of delivery

The course is delivered through an intensive, hands-on approach in which participants systematically build a standard Micro-simulation model, over the duration of the training. The programme is highly sequential, requiring full attendance at all sessions, as each step builds directly on the previous one. Participants work in teams to enhance peer learning and methodological coherence, while each participant is required to have an individual laptop and mouse for practical implementation. By the end of the course, each team presents policy simulation results generated from their model. For the online delivery option, participants are required to submit all assignments within the stipulated timelines to ensure completion within the scheduled period, as extensions may incur additional facilitation costs. To apply for this course, fill in the form below or send an email to apply@macrosolve.net or macrosolveinfo@gmail.com.