Microsimulation approach

Marco V Sánchez
UN-DESA

Martín Cicowiez
CEDLAS-ULP
UN-DESA consultant

Second intensive training workshop on "**Strengthening Macro-Micro Modelling Capacities to Assess Development Support Measures and Strategies**", Kampala, 11-14 September, 2012.

MDG 1 is part of our analysis

- MDG 1: half the extreme poverty between 1990 and 2015
- MAMS generates an indicator for all MDGs (1, 2, 4, 5, 7a and 7b)
  - Non-poverty MDG indicators:
    - Logistic function
    - Targeting through scaling up public spending
  - MDG 1 generated as a result of general equilibrium effects affecting:
    - per-capita income/consumption (total), given an elasticity
    - average income/consumption between groups of households, given a fixed within-group inequality for each group of households

CGE models and inequality

- A typical CGE model is composed of groups of representative households
  - Only between-group income distribution
  - Omits within-group income distribution
    - can influence poverty outcomes notably!
  - And, even if we had the detail on within-group income distribution: how do we know which workers are more likely to change position in the labour market?
    - e.g.: if, as a result of a policy simulation, the unemployment rate increases: who is expected to lose her/his job?
- Can this methodological limitation be overcome?

A microsimulation approach can be used: what is it?

- Model that uses information at the level of microeconomic individual agents (individuals, households, firms).
- As such, it permits to evaluate effects of policy or other shocks on those individual agents
  - Often through ex-ante evaluations

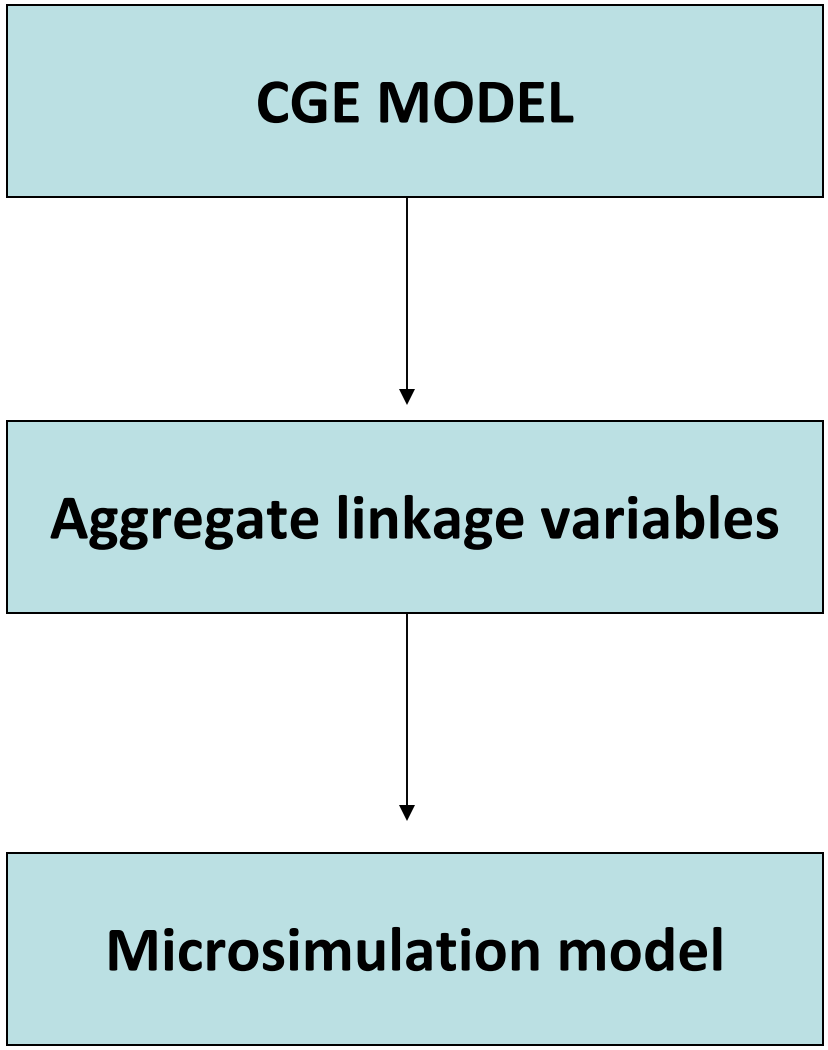
A microsimulation approach can be used: what is it? – cont.

- Typically it requires micro-data from a household survey covering:
  - socio-economic characteristics of individuals
  - labour-market status and labour incomes
  - household spending
- Changes in budget constraints are simulated
- Works at the partial equilibrium level
- It implies no simultaneous modelling of prices, wages, or macro processes

Top-down macro-micro approach

- CGE (macro) simulation results taken and applied to the full distribution as given by a micro data set (i.e., the household survey)
  - Assumption: there are no further feedback effects
- Macro modelling: provides simulation results on production, employment, wages, etc.
- Micro modelling: it permits to transform macro modelling results into results given by micro-data, using the household survey

Top-down modelling approach



Alternative approaches

- Parametric
  - generally imply a system of equations that determine occupational choice, returns to labour and human capital, consumer prices and other household (individual) income components
- Non-parametric
  - generally imply seeking individuals with similar characteristics to simulate certain change (for example, a change in labour income for an individual that moves from unemployment to employment)
  - occupational shifts may be proxied by a random selection procedure within a segmented labour-market structure (e.g. Paes de Barros et al. and subsequent extensions to link with CGE models)

Per capita household income

- Generally, the top-down approach is used to, after simulated (counterfactual) changes, generate a simulated per capita household income defined as follows:

$$ypc_h = \frac{1}{n_h}\left[\sum_{i=1}^{n_h} yp_{hi} + yq_h\right]$$

where,

- n_h = size of household h
- yp_{hi} = labour income of member i of household h
- yq_h = sum of all non-labour incomes of the household

What if poverty is calculated based on consumption?

- ¿Use poverty incidence indicators based on income anyway?
- Alternative 1: in order to match base-year poverty incidence as measured from both income and consumption, there may be two options:
  - recalculate per capita income to match it with per capita consumption
  - recalculate the poverty lines
- Alternative 2: use consumption that changes when income changes

Non-parametric microsimulations: where do we start?

- The method analyses effects of a change in the labour market structure on poverty and inequality.
- It is a counterfactual analysis: what would poverty and inequality indicators look like had the labour market structure be different to the observed one?
  - For example, one could use the labour market structure as recorded in MAMS instead of the observed labour market structure as recorded in a given household survey.

Non-parametric microsimulations: where do we start? – cont.

- The economically active population (EAP) is split up into j groups according to
  - sex (2)
  - level of skills
  - (individual's characteristics)
- The occupied population is split up into k groups according to
  - sector of employment
  - occupational category
  - (individual's characteristics)
- The counterfactual structure of the labour market can be defined arbitrarily or as a result of a macro-simulation.

Classification of population in working age

Partici-pation	Employ-ment status	Men		Women	
		Skilled	Unskilled	Skilled	Unskilled
Active	Employed				
	Un-employed				
Inactive					

Classification of employed population (EXAMPLE = 16 labour categories)

Employed		Men		Women	
Sector	Occupational category	Skilled	Unskilled	Skilled	Unskilled
Tradable sector	Wage				
	Non-wage				
Non-tradable sector	Wage				
	Non-wage				

Modelling of the labour market (Paes de Barros and others approach)

- The labour market structure λ is a function of the following parameters:

$$\lambda = \lambda\,(P, U, S, O, W_1, W_2, M)$$

  - P - participation rates for labour type j
  - U - unemployment rate for labour type j
  - S - employment structure by production sector
  - O - employment structure by occupational category
  - W_1 – remuneration structure by sector
  - W_2 – overall average remuneration
  - M - composition of employment by individual's characteristics (e.g., skill level)

How does it work?

- A random number is assigned to each person at working age
- Population at working age is ordered according to:
  - activity condition (active vs inactive),
  - employment condition (employed vs unemployed)
  - economic sector
  - occupational category
  - education level
  - random numbers

How does it work? – cont.

New labour market structure λ^*

- Individuals become active/inactive, employed/unemployed, change their occupational position and/or level of education
  - How? See example in next slide.
- Income (YPI) is assigned to all those individuals who, according to λ^*, become employed, or change their occupational position and/or level of education
  - How? A non-parametric (random) process is also employed to assigned counterfactual incomes
- Income of all those individuals that become unemployed or inactive are set equal to zero

Example: effect of a change in the unemployment rate of skilled men workers (N=100)

		Simulation 1		Simulation 2		
	N	Un-employment rate falls to 6%	Simu-lated	Un-employment rate increases to 12%	Simu-lated	
Employed	90	Unchanged	90	↓ ↓ ↓ ↓ ↓ The last 2 employed become unemployed	88	Employed
		↑ ↑ ↑ ↑ ↑ The first 4 unemployed become employed	4		2	
Un-employed	10		6	Unchanged	10	Un-employed

How does it work? - cont.

- Same procedure is applied for other shifts:
  - active vs inactive (P)
  - employed vs unemployed (U)
  - employment by sector (S), occupational category (O) and skills (M)
- Wages
  - To simulate changes in W_1 all YPIs within each of the 16 labour categories are multiplied by an adjustment factor, maintaining the overall average wage/labour income level fixed
  - To simulate changes in W_2 all YPIs are multiplied by an adjustment factor such that the overall average labour income level is adjusted in accordance with the average wage increase derived from the counterfactual scenario
- Based on the simulated YPIs the new total per capita household incomes (YPC) are computed obtaining a new, counterfactual income distribution
- New inequality indicators using alternative measures (Gini, Theil, etc.), and poverty indicators (for alternative poverty lines) are computed

Key assumptions

- We do not need a full model of the labour market
  - there are only individuals changing their j or k groups

- A randomized process is applied to simulate the effects of changes in the labour-market structure
  - It assumes that, on average, the effect of the random changes correctly reflects the impact of the actual changes in the labour market

- Because of the introduction of a process of random assignation, the micro-simulations are repeated a large number of times in Monte Carlo fashion ➔ this allows constructing 95 per cent confidence intervals for the indices of inequality and poverty

In summary:

- From CGE model, changes in the labour market structure are applied (individually or sequentially) to micro data, affecting the overall income distribution:

$$\lambda^* = \lambda^*(P^*, U^*, S^*, O^*, W^*_1, W^*_2, M^*)$$

- Who moves? Determined through a random process which generates a new income distribution

- Micro-simulations are repeated many times in Monte Carlo fashion to compute confidence intervals for inequality and poverty indicators that are statistically significant

Advantages vs disadvantages

Advantages:

- Enables to analyse the impact of a wide range of labour-market parameters, individually or sequentially
- Shows separate and combined effects of each type of labour market shift (e.g. unemployment change, wage change, etc.) on poverty and inequality outcomes
- It does not demand econometric estimation

Possible disadvantages:

- Behaviour is not modelled
- Results in sequential application may depend on the order in which the sequence of labour-market parameter changes is applied ("path dependence")

In a dynamic setting

- A number of additional, restrictive assumptions are required as observed survey data may only be available for the base year and perhaps a few years beyond that, but certainly not for the forward simulation period.
- CGE outcomes (deviations from base year for any given simulation year) are imposed on base year household survey data
  - beyond the base year and for lack of additional modelling of demographic shifts and labour participation, it is assumed that no changes in the population structure (such as migration or population ageing) take place during the simulation period.
  - hence, only one household survey is used, to which labour market structures for *t* periods are imposed
  - obvious limitation of the methodology, but justifiable to the extent that the CGE model does not consider such demographic changes either.

Sequential and cumulative effects for changes in the labour-market parameters for the baseline scenario

	Total poverty incidence (% of population)	Extreme poverty incidence (% of population)	Gini coefficient for labour income	Gini coefficient for per-capita household income
2008				
U	20.7	4.3	0.461	0.497
$U+S$	20.7	4.3	0.461	0.497
$U+S+O$	20.7	4.3	0.461	0.497
$U+S+O+W_1$	20.7	4.3	0.461	0.497
$U+S+O+W_1+W_2$	20.7	4.3	0.461	0.497
$U+S+O+W_1+W_2+M$	20.7	4.3	0.461	0.497
2010				
U	20.6	4.3	0.461	0.497
$U+S$	20.6	4.3	0.461	0.497
$U+S+O$	20.6	4.3	0.461	0.497
$U+S+O+W_1$	19.8	4.1	0.456	0.491
$U+S+O+W_1+W_2$	19.6	4.1	0.456	0.491
$U+S+O+W_1+W_2+M$	19.5	4.1	0.456	0.49
2012				
U	20.5	4.2	0.461	0.497
$U+S$	20.5	4.2	0.461	0.497
$U+S+O$	20.4	4.2	0.461	0.496
$U+S+O+W_1$	18.1	3.8	0.447	0.479
$U+S+O+W_1+W_2$	16.6	3.6	0.447	0.479
$U+S+O+W_1+W_2+M$	16.5	3.6	0.447	0.478

Source: CGE model and microsimulation results for Costa Rica.