Microsimulation approach

Marco V Sánchez
UN-DESA

Inception workshop on "**Strengthening Macroeconomic and Social Policy Coherence through Integrated Macro-Micro Modelling**", Amman, 28-29 March, 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

Why do we need a microsimulation methodology?

- A typical CGE model is composed of groups of representative households and representative workers
  - Only between-group income distribution
  - Omits within-group income distribution
    - can influence poverty outcomes notably
  - And, even if we have 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?
- How can this methodological limitation be overcome?

Some alternative approaches

- Top-down approaches:
  - CGE 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
- Modelling labour market adjustment:
  - ***Household income generation model***: system of equations that determine occupational choice, returns to labour and human capital, consumer prices and other household (individual) income components (Bourguignon et al.).
    - *parametric approach*
  - ***Occupational shifts proxied by a random selection procedure within a segmented labour-market structure*** (Paes de Barros et al.)
    - *non-parametric approach*

Modelling of the labour market

- The two top-down approaches define total per capita household income 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

Modelling of the labour market (non-parametric 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 skill level

  – Labour type j is defined by sex and skills
  – Segments k are defined based on economic sector and occupational category

Classification of population in working age

		Men		Women	
		Skilled	Unskilled	Skilled	Unskilled
Active	Employed				
	Un-employed				
Inactive					

Classification of employed population (EXAMPLE = 16 labour categories)

		Men		Women	
		Skilled	Unskilled	Skilled	Unskilled
Tradables sector	Wage				
	Non-wage				
Non-tradables sector	Wage				
	Non-wage				

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 and
  - education level, and...
  - … within “segments”, according to random numbers
- Income (YPI) is assigned to all those individuals who, according to λ*, become employed, or change their occupational position and/or level of education
- 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 as for shifts between employed and unemployed (U) for shifts by labour category (O) and sector (S)
- 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 (for YPI and YPC), 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 “segments”, but individuals can move from one “segment” to another under certain restrictions (sex, skilled level, and so on)

- 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.

Real GDP growth and labour-market, poverty and inequality results in a baseline scenario for a developing country

	2008	2010	2012
Real GDP (at factor cost)	2.6	2.2	4.4
Total unemployment rate (%)	6.0	5.9	5.9
Employment (in thousands of workers)	1,958	2,035	2,115
Labour income per worker	239,984	242,083	254,820
Total poverty incidence (% of population)	20.7	19.5	16.5
Extreme poverty incidence (% of population)	4.3	4.1	3.6
Gini coefficient for labour income	0.461	0.456	0.447
Gini coefficient for per-capita household income	0.497	0.49	0.478

Source: CGE model and microsimulation results for Costa Rica.

Real labour income of workers by skill, sex and occupational category for a developing country

	Real labour income per worker (*colones*)			Relative remuneration (W_1) [1/]		
	2008	2010	2012	2008	2010	2012
Unskilled female workers						
Wage earners	167,077	165,631	167,821	0.696	0.684	0.659
Non-wage earners	44,021	44,820	47,644	0.183	0.185	0.187
Unskilled male workers						
Wage earners	243,219	258,632	293,963	1.013	1.068	1.154
Non-wage earners	160,052	166,719	190,108	0.667	0.689	0.746
Skilled female workers						
Wage earners	396,095	388,726	389,430	1.651	1.606	1.528
Non-wage earners	90,935	84,048	81,142	0.379	0.347	0.318
Skilled male workers						
Wage earners	390,806	383,148	384,967	1.628	1.583	1.511
Non-wage earners	165,769	155,686	158,703	0.691	0.643	0.623
Average labour income economy	239,984	242,083	254,820	1.000	1.000	1.000

[1/] Changes in relative remuneration are similar across sector of activity.

Source: Baseline estimates of CGE model for Costa Rica.

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.