INTERNATIONAL JOURNAL OF MICROSIMULATION (2010) 3(1) 8-23

A Non-Parametric Microsimulation Approach to Assess Changes in Inequality and Poverty

Rob Vos and Marco V. Sánchez
United Nations Department of Economic and Social Affairs, New York, NY 10017, USA; email: vos@un.org; email: sanchez-cantillo@un.org

ABSTRACT: This paper presents a non-parametric microsimulation methodology for assessing the impact of labour market changes and government transfers on income inequality and poverty at the household level. The approach assumes that labour markets are segmented and determines (as part of a randomized process) which individuals are expected to move in or out of employment and which move from one employment segment to another based on either known or counterfactual information of aggregate labour market changes. The methodology assumes that the distribution of earnings of those who become employed in a particular segment resembles that of the individuals observed to be employed in that segment. The approach can be effectively combined in top-down fashion with static or dynamic computable general equilibrium (CGE) models, which typically provide insufficient information about household income distribution. The paper discusses the virtues and limitations of applying this methodology and further explains to practitioners how to implement it as a stand-alone methodology or in combination with a CGE model. It also shows how the methodology can be generalized to also capture the poverty and inequality effects of changes in non-labour incomes, such as government transfers. One great advantage of this method is that it is not very demanding in terms of modelling labour supply and household behaviour as compared with alternative parametric approaches, while at the same time providing a plausible link between changes in overall labour market conditions and the full household income distribution.

1. INTRODUCTION

The empirical literature on income distribution is extensive and rather fragmented. Much focuses on the outcomes of the labour market processes, since these appear to be a crucial determinant of earnings and income inequality. The development of counterfactual microsimulations methods to study the determinants of distributional changes owes much to the seminal work by Mincer (1958), Oaxaca (1973) and Blinder (1973) who focused on the determinants of differentials in wage earnings. These methods are rooted in human capital theory. They are designed to analyze how the income distribution changes depending on the characteristics of individual workers (such as their working experience or level of education), but are less useful in assessing the impact of changes in aggregate labour market conditions (such as the level of unemployment, sectoral labour demand and wages) as further information is required about which workers are most likely to shift position in the labour market in response to changed conditions.

Recently, several microsimulation methods have been developed which draw on an idea originally developed by Orcutt in the 1950s (Orcutt, 1957) and which try to overcome this informational deficit as well as to assess how changes in the labour market affect poverty and income inequality at the household level. One such approach consists of an econometrically estimated household income generation model, as proposed by Bourguignon and others (see *e.g.*, Bourguignon, *et al.*, 2001, 2002a,b). The probabilities and determinants of the model subsequently are used to simulate the impact of changes in labour market conditions, endowments of human capital and returns to these endowments on inequality and poverty at the household level. An alternative counterfactual microsimulation methodology, developed originally by Almeida dos Reis and Paes de Barros (1991), does not explicitly model labour market behaviour, but assumes instead that the way in which shifts in overall labour market conditions (like changes in unemployment, sectoral labour demand or wages) affect occupational conditions of individual workers can be proxied by a random selection procedure in a segmented labour market. This method can be called non-parametric or non-behavioural, because it does not involve econometrically estimated probabilities of the underlying behaviour.

Both methods have been applied in conjunction with economy-wide models, more in particular with computable general equilibrium (CGE) models, which may provide a counterfactual for the simulated impact on labour market conditions of an economic shock or policy change. Such models typically only provide outcomes for employment and wages by rather aggregate labour categories and household groups, though. As a consequence, these models provide too little information about distributional changes to derive robust estimates of the impact of simulated shocks and policies on poverty and income distribution at the household level. This limitation can be overcome by combining the CGE analysis with the type of microsimulation methodologies as just described. Bourguignon *et al.*, (2002a,b) first probed this macro-micro modelling using the parametric or behavioural approach for the case of Indonesia. The use of the non-parametric microsimulation approach in conjunction with CGE model analysis, in turn, was pioneered in a number of Latin American studies and later applied more widely (see Vos *et al.*, 2002).

In this paper we focus on the non-parametric microsimulation methodology, its theoretical foundations and applications, and on how the method works in combination with a dynamic CGE model. In doing so, we will explain the virtues and limitations. The main advantage of the non-parametric microsimulation methodology is that it resolves the 'assignment' problem in the labour market while requiring little actual modelling effort in contrast to alternative parametric approaches. One of its weaknesses is that a certain sequence must be assumed in how different dimensions of aggregate labour market changes (*e.g.* changes in unemployment rates, sectoral labour demand, etc.) impact on the situation of individual workers and their families and this could create a problem of path dependence. As argued in this paper, such path dependence may also affect alternative approaches, but more importantly, it need not be a major concern once one can justify the logic of the followed sequence on the basis of plausible labour market behaviour and as long as the size of certain shifts are not very large.

The remainder of this paper is organized as follows. In section 2, we explain the key theoretical notions behind the non-parametric microsimulations methodology and show how it works in practice. This section also addresses the problem of path dependence and how the methodology can also be used to assess changes in non-labour incomes on household income distribution and poverty. In Appendix A.1 a practitioner's guide is given of the basic steps to be taken.[1] Section 3 explains how the method can be applied in conjunction with an economy-wide model, more in particular a dynamic CGE model. As a further addition to the existing literature, the section addresses how to implement the microsimulations in a dynamic setting and the additional assumptions that need to be made as a consequence. The following section provides a numerical example using outcomes from a CGE model for Costa Rica to analyze the impact of changing labour market conditions and government transfers on inequality and poverty by means of the non-parametric simulation method. Section 5 concludes.

2. NON-PARAMETRIC MICROSIMULATION APPROACH

The basic idea behind the microsimulation methodology is to isolate the effects of key determinants of the changes in poverty and inequality. As mentioned, the methodology presented here was originally developed by Almeida dos Reis and Paes de Barros (1991) to analyze wage inequality and was subsequently generalized in order to analyze income inequality and poverty based on the total per capita household income (see Paes de Barros and Leite, 1998; Paes de Barros, 1999; Frenkel and González, 2002; Ganuza *et al.*, 2002). The approach does not explicitly model labour market behaviour for reasons explained further below. Instead, it assumes that the impact of changes in overall labour market conditions on the employment status and labour incomes of individual workers can be proxied through a random selection procedure in the context of segmented labour markets.

2.1. Basic Notions

The basic intuition behind the non-parametric microsimulation approach referred in the above can be explained through a set of basic identities. Total per capita household income (ypc_h) is defined as:

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

where, n_h is the size of household h, yp_{hi} the labour income of member i of household h, and yq_h the sum of all non-labour incomes of the household, defined as:

$$yq_h = \sum_{i=1}^{n_h} yqp_{hi} + yqt_h \qquad (2)$$

In equation (2), yqp_{hi} equals individual non-labour income of member i of household h and yqt_h equals other household incomes. In the microsimulations, yp_{hi} is altered for some individuals i of household h as a result of changes in the variables that define the labour market structure.

The labour market structure in year t is defined first in terms of rates of labour participation (P_j) and unemployment (U_j) among different groups j of the population in working age. Individuals maybe classified according to personal characteristics such as sex, age and skill. Next, the structure of employment is defined by sector of activity (S) and occupational category (O) and remuneration (W_1), the level of overall remuneration (W_2) as well as the skill composition of the employed population (represented by variable M). The labour-market structure can thus be written as:

$$\lambda = \lambda\,(P, U, S, O, W_1, W_2, M) \qquad (3)$$

To define a labour market structure in year t, the population at working age in that year could be classified, for example, into *four types* of individuals j according to sex and two levels of education (skill versus unskilled), while *four segments* of the labour market k are defined according to occupational category (wage employees versus self-employed or non-wage workers) and sector of economic activity (for example, agriculture versus non-agriculture). The microsimulation methodology is flexible, however, as regards the number and types of categorizations the user wishes to identify.

Now, let earnings of an individual i be a function of personal characteristics, such as sex, age and skills, and his or her position in the labour market.

Let skills, sex, age and other individual attributes be represented by a variable *c*. Individual earnings can then be written as:

$$yp_i = f(\lambda, c_i) \tag{4}$$

In each microsimulation, employment conditions of an individual *i* may change as a result of changes in labour market conditions. Hence:

$$yp_i^* = f(\lambda^*, c_i) \tag{5}$$

where λ^* represents a counterfactual labour market structure. In the application of the microsimulation methodology, the effects can be assessed by altering the variables *P*, *U, S*, *O*, W_1, W_2 and *M* separately or in sequential cumulative form.

Given λ^* and the resulting yp_i^*, the simulated per capita income is defined as:

$$ypc_h^* = \frac{1}{n_h}\left[\sum_{i=1}^{n_h} yp_{hi}^* + yq_h\right] \tag{6}$$

The distribution of per capita household income in year *t* is a function of the above-mentioned labour market variables, as well as the skill, age and sex distribution within households, represented by *c*, and other factors, captured by a parameter *a*:

$$\begin{aligned} D(ypc_i) &= D(P,U,S,O,W_1,W_2,M,c,a) \\ &= F(\lambda,c,a) \end{aligned} \tag{7}$$

Similarly, the distribution of income per capita (*ypc*) in year t^* is defined as:

$$D^*(ypc_i) = F(\lambda^*, c^*, a^*) \tag{8}$$

Assuming "other factors" (*a*) are constant and since a change in the overall level of earnings W_2 does not influence the distribution, the change in inequality in labour earnings between two years (or between a counterfactual and an observed distribution) can be defined as a function of the changes in the remaining six labour market variables and in individual attributes (*c*):

$$\begin{aligned} \Delta D(yp_i) = \; & D(P^*,U^*,S^*,O^*,W_1^*,M^*,c^*) \\ & - D(P,U,S,O,W_1,M,c) \end{aligned}$$

(9)

In a microsimulation, the estimates of the six labour market variables of year *t* would be replaced by those of year *t** (or the counterfactual obtained from, for instance, a CGE model simulation). The simulated change is thus defined as:

$$\begin{aligned} \Delta^{sim}D(yp_i) &\equiv D^{sim}(yp_i) - D(yp_i) \\ &= \sum_{s=1}^{6}[D_s^{sim}(yp_i) - D_{s-1}^{sim}(yp_i)] \end{aligned} \tag{10}$$

where

$$D_1^{sim}(yp_i) = D(P^*,U,S,O,W_1,M,c),$$

$$D_2^{sim}(yp_i) = D(P^*,U^*,S,O,W_1,M,c),$$

$$D_3^{sim}(yp_i) = D(P^*,U^*,S^*,O,W_1,M,c), \;\ldots\ldots\; \text{etc.}$$

In principle, the difference between the observed and simulated change is the result of interaction effects and the fact that *c* may have changed as denoted as follows:[2]

$$\Delta D_{res} = \Delta D(yp_i) - \Delta^{sim}D(yp_i) \tag{11}$$

The interaction effects (ΔD_{res}) are derived here as a residual. In the discussion further below we discuss the importance of these effects when addressing the problem of path dependency in the application of this microsimulation methodology.

2.2. Randomized labour market behaviour

The methodology does not explicitly model labour market behaviour. The latter is approached in a rudimentary way by assuming a certain compartmentalization or segmentation of the labour market and the possibility of individuals to move from one segment to another. Individuals that change segment are assigned a new labour income which is the average of the workers in that segment. In a parametric approach, such as that proposed by Bourguignon and others, for instance, one would model the probability of an individual worker being employed or not, being in a particular occupation and sector of activity and earning a labour income corresponding to the worker's occupation. Those probabilities then can be used to simulate the likelihood of a worker to change position in the labour market when overall conditions change.

A conventional approach to study the distribution of earnings is to apply the human capital model and to model labour participation and earnings functions based on individual characteristics and conditioning factors of the household the individual belongs to. In essence, this is also the approach applied by Bourguignon and others, in the development of their microsimulations approach. This approach to modelling labour market behaviour had been criticized, however, for overemphasizing labour supply factors and inadequately considering the demand side of the labour market (see *e.g.* Hartog, 1985, 1986). Alternative models acknowledge that there is an 'assignment problem' in the economy (see *e.g.* Sattinger, 1993). These labour market models consider that the distribution of earnings cannot be explained by merely considering characteristics of *individuals*, but that the characteristics of *jobs* also need to be taken into account. Assignment models, which consider individual and job characteristics, assume that selection in labour market positioning is the result of voluntary choices by individual workers. Models of labour market segmentation, in contrast, assume the selection is involuntary (see the overview of theories in Atkinson and Bourguignon, 2000:23). This approach goes to another extreme by assuming that access to some segments of the labour market is restricted. It is then labour demand, rather than individual attributes on the supply side that is seen as the major determinant of the labour market position of individuals and hence also of the distribution of earnings. There is no consensus as to which labour market modelling approach provides the best empirical approximation of actual behaviour and of changes in labour earnings.

The non-parametric approach to microsimulations presented here takes a middle ground. It considers both individual characteristics of workers and a certain labour market segmentation, but allows workers to move across segments at the margin; that is, workers are allowed to move from unemployment into employment, from non-wage to wage employment or from agriculture to non-agriculture, or *vice versa*, for instance, depending on changes in aggregate labour market conditions set from both the supply and demand side. Given the complexities in adequately modelling labour market behaviour empirically, it is then assumed that the probability that one rather than another individual moves position may just as well be approximated by a *randomized process*.

In order to simulate which person or worker is affected by a particular change in labour market conditions, the non-parametric microsimulation approach assigns random numbers to individuals grouped by the predefined individual attributes and labour market segments. By repeating the simulations a sufficient number of times—in Monte Carlo fashion—a confidence interval of 95 per cent for the results (such as inequality and poverty indices) can be generated.[3] This randomized process is used to determine: (*i*) which persons at working age change their labour force status (inactivity versus activity; and, if active, employed versus unemployed); (*ii*) who will change from one segment of the labour market to another (sector and/or occupational category); (*iii*) which employed persons obtain a different level of education; and (*iv*) how are new mean labour incomes assigned to individuals.[4] Appendix A.1 spells out each step to be taken when applying the simulation methodology in practice.

2.3. Simulating income inequality and poverty

Following the indicated procedure, a new income distribution is generated in each simulation. From here any desired poverty and inequality index can be calculated. Since random numbers are used, the mean values of the simulated poverty and inequality indices of each iteration have to be calculated. Hence, the assumption is that, on average, the effect of the random changes correctly reflects the impact of the actual changes in the labour market. Analogous to equation (10), the mean simulated change in, for instance, the indices of earnings inequality can be written as:

$$\overline{\Delta^{sim} I(yp_i)} \equiv \sum_{s=1}^{6} [\bar{I}_s^{sim}(yp_i) - \bar{I}_{s-1}^{sim}(yp_i)] \qquad (12)$$

where

$\bar{I}_0^{sim}(yp_i) = I(yp_i)$.

Accordingly, the residual change is defined as:

$$\Delta_{res} I = \Delta I(yp_i) - \overline{\Delta^{sim} I(yp_i)} \qquad (13)$$

The main advantage of this non-parametric micro-simulation methodology is that it allows simulating the impact of changes in the labour-market structure on the full income distribution, while staying low in modelling intensity. In fact, it requires relatively little information beyond the micro dataset providing the full income distribution, such as a household survey. The method further allows for a dynamic decomposition of the relative importance of the factors driving changes in the distributions of incomes and in poverty.

2.4. Path dependence

A weakness of the approach is, of course, that it does not make full use of all possibly available information on labour market behaviour. However, as indicated, doubts about the robustness of the parameters derived from existing labour market models may precisely provide one reason for using the more simple non-parametric approach. More critically, perhaps, the simulation results may be path dependent. That is, results might change depending on the choice of the base year when the microsimulation method is applied in a case of analyzing the distributional change between two observed labour market structures. Path dependence could also be a problem in any sequential simulation of changes in labour market conditions. In other words, it could make a difference, given the cumulative effects, whether in an assumed sequence one would first simulate, say, changes in employment by occupational category (O) rather than by sector of employment (S). In the simulation procedure one has to define the sequence upfront; hence, knowing whether the particular sequence in which changes in labour market conditions are imposed matters or not is important.[5]

It can be shown formally that path dependence in terms of the order of sequence shows greater sensitivity when the interaction terms of the changes in the labour market variables are large (see De Jong, 2001 and Vos and De Jong, 2001; and Appendix A.2 for a proof). The interaction terms, in turn, could become large when major shifts in the labour market take place. However, as argued by Ganuza *et al.* (2002), there is also an economic logic to the order in which the various effects need to be analyzed. The sequence as typically defined in applications of the non-parametric microsimulation method reproduces the steps used in many labour market models. It is assumed that agents first decide whether to participate or not (*P*). Then the market defines whether they can find employment or not (*U*). If they found employment, the adjustment process of the labour market defines in which sector (*S*) they will be located and the occupation category of their employment (*O*) (though, in practice, the sector and occupational choice more likely will be a simultaneous process.). Obviously, their decisions as to whether to work or not can be influenced by the relative remuneration, but, once the sector and the occupation in which they obtain employment are defined, the probability of the relative remuneration they will have becomes known. The changes in labour supply and demand in each segment of the labour market are likely to be a factor underlying changes in the remuneration structure for different sex and skill groups, while the change in average level of remuneration is a reflection of the overall performance of the economy. Finally, the skill composition of the workforce is partly a result of demand factors in the labour market, but probably mainly the result of more exogenous changes in endowments. Of course, human capital formation is also likely to be a function of expected earnings in the labour market (see *e.g.* Behrman, 1999). These are arguments for considering changes in the skill composition as the last step in the sequential simulation.

Although even by the above logic the sequence of some of the parameter changes could be reversed (such as between the assignment by sector or occupation category), the proposed sequence would seem plausible in most contexts. Empirical tests with alternative sequences for the urban labour market variables using the living standards measurement study (LSMS) surveys for Ecuador suggest that interaction terms in this particular case are small; hence, aggregate results are not sensitive to the sequence of parameter changes (Vos and de Jong, 2001). The same study for Ecuador also shows that altering the occupational category structure prior to the sector of economic activity structure yields somewhat larger residuals but does not much change the relative importance of the impact of the changes in the labour market variables on outcomes for inequality and poverty.[6]

2.4. Accounting for changes in non-labour incomes

The impact of changes in non-labour incomes (such as pension incomes or other transfers, rental incomes, as well as, with a negative sign, direct taxes) can be simulated by assigning the income change to households that are expected to receive the income. This may require additional information about which household are expected to benefit from the income change. The direct impact on inequality and poverty can then be simulated by making the appropriate adjustments to variable yq_h in equation 1 so as to obtain a new level and distribution of per capita household income ypc_h. Assuming information is available about which type of households are expected to gain (or lose) from the non-labour income change, no randomized assignment procedure needs to be applied. If it further maybe assumed that non-labour income generation does not directly interact with labour market variables, this additional step in the microsimulations procedure will not be path dependent and, hence, can be introduced either at the end or beginning of the microsimulations.

Changes in non-labour incomes may also have economy-wide effects by the way these are generated or by the impact on prices and economic activity. For instance, a cash transfer programme with significant coverage will have fiscal effects and affect production activities depending on the way the programme is financed (*e.g.* via higher taxation or increases in public borrowing). Changes in rental incomes may be associated with, say, changes in interest rates which likely have economy-wide effects through investment and consumption behaviour. Cash transfers received by individuals or families may affect labour participation behaviour (as it may induce them to work less or not to seek work) and, if significant, could alter outcomes in terms of unemployment and wage rates. As a further example, changes in workers' remittances from abroad will affect the balance of payments and may influence the real exchange rate, which may trigger further economy wide effects through its impact on exports, production, real wages, and employment. When using the microsimulation method in combination with an economy-wide framework, such as a CGE model, such general equilibrium effects can be taken into account by first simulating the impact of, say, a government transfer programme through the CGE model and then using the resulting labour market outcomes in an application of the non-parametric microsimulation methodology as outlined above.[7] This way, both the direct and second-order general equilibrium effects of a transfer on the distribution of per capita household incomes can be accounted for. In the illustration of an application of the non-parametric microsimulation method presented in section 4, an example is included of both the direct and indirect effects of a government transfer to households.

3. TOP-DOWN COMBINATION OF A CGE MODEL AND THE NON-PARAMETRIC MICRO-SIMULATION METHOD

The microsimulation approach outlined above may be applied to analyze distributional changes comparing two observed distributions or by imposing a counterfactual labour market structure (*e.g.* resulting after a macroeconomic shock or policy change) simulated through an economy-wide model. In the former case one would "impose" the labour market structure as observed in, say, 2010 on to household or labour force survey data for 2000 to assess the impact of changes in different labour market variables on poverty and inequality. In the latter case, the changes in labour market variables are model driven, for instance as simulated through a CGE model.

CGE models typically have enough detail by sectors and labour categories to provide enough 'structure' to meaningfully apply the microsimulation method. Macroeconomic models mostly only generate counterfactuals for aggregate employment and unemployment and for average, economy-wide wages. This would still leave the assignment problem as to, for example, which workers are more likely to become unemployed (or which unemployed would find a job) when assessing distributional outcomes, but clearly would not reveal any other labour market shifts which result in distributional changes.

The non-parametric microsimulation method has been applied to both outcomes derived from static (see *e.g.* Vos *et al.*, 2002, 2006) and dynamic CGE models (see *e.g.* Sánchez, 2004; Sánchez and Vos, 2005, 2006). In both cases, the modelling approach is called 'top down' in that the results of the CGE model are taken forward to the microsimulations, but that there is no feedback from the outcomes of the latter into the CGE model. A major advantage of doing this in top-down fashion is that the analysis (and 'modelling') of household and labour market behaviour can be done separately from the economy-wide analysis and that there is no need to reconcile household survey data with national accounts and other macroeconomic data. The communication between the two types of models is in the form of information about prices, wages and employment and there is no need to reconcile data on levels.

One reason to combine the two methods is that CGE models typically only provide information about (simulated) changes in income distribution between fairly aggregate labour categories and household groups, hence missing out much of possible within-group changes. The microsimulations can then help to make up for the missing detail. An alternative approach would be to introduce distribution functions into the CGE model. However, this is often done by assuming given (static) distribution functions and does not resolve the 'assignment problem', as discussed above, and because of which the within-group distribution may not be assumed constant. The microsimulation methodology resolves this problem, but under the limiting assumptions as discussed in the previous section.

A further alternative would be to apply a sequential modelling approach whereby bi-directional link between a CGE and a household and occupational choice model is established and requires obtaining a converging solution between both models. An example of this approach can be found in Savard (2003) and has the advantages of considering feedback effects of poverty and inequality on consumption, production, and labour market adjustment and of ensuring coherence between the micro data and the aggregates of the CGE model. This method is, however, more demanding as it requires maintaining coherence in macro-micro behaviour across the two types of models and convergence cannot be guaranteed in practice. The results for poverty and inequality are also less tractable than in the top-down approach, as presented here.

Application of the top-down approach using a static CGE model and the non-parametric microsimulation method is straightforward. It works the same way as when applying the microsimulations to household survey data at two different points in time. In a purely static framework, the CGE model would be used to generate the counterfactual labour market structure $\lambda*$.[8]

The top-down approach can be also be applied using a dynamic CGE model following the procedure spelled out in Sánchez (2004) and Sánchez and Vos (2005, 2006). This procedure combines dynamic CGE simulations and non-parametric microsimulations based on information from one household survey (typically for the base year of the model). Practitioners may have little other alternative if their CGE model performs counterfactual simulations into the future for years for which no new observed survey data are available. In these applications, the dynamic CGE model provides estimates of the labour market structure λ_t^{*base} for each year t of the baseline period. Policy experiments or experiments reflecting an exogenous shock are typically carried out to modify the baseline, generating simulations for new labour market structures λ_t^{*sim} for each scenario *sim*. The new labour market structures for the baseline and all scenarios are then imposed on the base-year survey data set by running the non-parametric microsimulation procedure.

Following this procedure requires accepting a number of additional restrictive assumptions. One would either need to assume no demographic changes take place during the simulation period or that these can be imposed exogenously (by adjusting variables "*c*" and "*a*" in equation 7) as most dynamic CGE models do not endogenize demographic change. More typically, these assume demographic changes are exogenous and in some cases constant (such as participation rates). The consequence of the latter assumption is participation rates (P) would remain unchanged and hence play no role in distributional outcomes assessed through the microsimulations.

4. AN APPLICATION WITH A CGE MODEL FOR COSTA RICA

4.1. Baseline simulations

In this sub-section we show the application of the non-parametric microsimulation approach in combination with a dynamic-recursive CGE model developed by Cicowiez and Sánchez (2009) and designed to assess the impact of external shocks on income distribution and poverty and the effectiveness of social transfer schemes in protecting the poor against such shocks.

The baseline scenario of this model runs from 2002 until 2012 as we apply this using a dataset for Costa Rica. We combined a "growth calibration" with a set of closure rules to enable replication of the aggregate functioning of the economy as observed for 2002-2009. For 2010-12 the baseline is calibrated to an output growth path suggested by projections of the Central Bank of Costa Rica. According to these projections, Costa Rica's economy is seen to recover quickly from the global economic crisis that erupted in 2008 and to converge to its pre-crisis GDP growth rate of around 4.5 per cent per annum.

The non-parametric microsimulation methodology is first applied to generate new baseline income distributions for every year of the simulation period using the outcomes for the labour market variables embedded in the CGE model. Though the actual base year of the CGE model is 2002, below we present a selection of results for the 2008-2012 period only. The micro dataset used for running the microsimulations was derived from the 2008 Multiple Purpose Household Survey of the National Institute of Statistics and Censuses (INEC) of Costa Rica.

As shown in Table 1, the baseline assumptions of the CGE for Costa Rica are consistent with a relatively rapid output recovery from the 2008-2009 global economic crisis with employment and wages showing positive growth during 2008-2012. While the CGE model allows for underutilization of factors and, hence, unemployment, the rate of unemployment remains rather constant as in 2008 (the base year here), the assumed minimum or 'natural' rate of open unemployment of about 6 per cent was reached and consequently with the growth and employment recovery after the 2009 recession, most labour market adjustment falls on real labour incomes. These fall visibly in 2009, but the average labour income per worker (W_2) would recover in 2010, along with overall employment conditions. Employment by sector (S) and type of occupation (O) does not change much in the baseline growth path, but there are significant shifts in relative wages (W_1), as shown in Table 2.

Table 1 Real GDP growth and labour-market, poverty and inequality results in the baseline scenario for Costa Rica, 2008-2012

	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 [1/]	239,984	242,083	254,820
Total poverty incidence (% of population) [2/]	20.7	19.5	16.5
Extreme poverty incidence (% of population) [2/]	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	0.490	0.478

Source: CGE model and microsimulation results for Costa Rica.
Notes:
[1/] Real monthly labour income in *colones*, excluding social security contributions
[2/] Calculation based on a national poverty line.

Table 2: Real labour income of workers by skill, sex and occupational category, 2008-2012

	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

Source: Baseline estimates of CGE model for Costa Rica.
Notes: [1/] Changes in relative remuneration are similar across sector of activity.

Table 2 shows that the relative remuneration of unskilled workers, especially wage male workers, regardless of their occupational category, increase substantially relative to the base year. In contrast, earnings of skilled workers are projected to decline. Exchange rate depreciation, triggered by the recession at the beginning of the period, is the main cause of this redistribution of labour incomes. In the CGE model, the depreciation would stimulate an export-led recovery in subsequent years mainly benefiting export agriculture and other activities that are relatively intensive in the use of unskilled labour. The increase in demand for this labour category exceeds growth of labour supply (which is exogenous in the CGE model), thus pushing up wages of unskilled workers relative to those of other workers.

The information about the simulated changes in the labour market variables as indicated above was subsequently used to run the microsimulations and obtain poverty and inequality estimates for each year of the baseline period. The results are presented in Table 3. Since, as mentioned the participation rate is assumed constant in the Costa Rica CGE, variable *P* does not change and, hence, in this application the sequential and cumulative effects of changes in the labour market structure start with the impact of changes in the unemployment rates (for *j* groups of individuals), followed by the other effects in the sequence of equation 3. Not surprisingly given the CGE baseline outcomes, changes in the remuneration structure (W_1) exert the stronger impact on changes in poverty. Unskilled workers which are more likely to be poor, see above-average income improvements helping a fair number of them (and their families) climb out of poverty. The increase in the mean remuneration (W_2) also has a visible poverty-reducing impact but weaker than the distributional effect. The changes in unemployment (*U*) and composition of employment by sector (*S*) and occupation (*O*) show as expected no notable effect and also changes in the skill composition of employment are not large enough to affect poverty in any significant way.

A similar pattern is found in the impact of the labour market adjustment on inequality as measured by the Gini coefficient. In the baseline simulation it is principally the change in W_1 which helps reduce the inequality in both the distribution of labour incomes and that of household per-capita income.

Table 3 Sequential and cumulative effects for changes in the labour-market parameters for the baseline scenario for Costa Rica[1/], 2008-2012

	Total poverty incidence (% of population) [2/]	Extreme poverty incidence (% of population) [2/]	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	0.490
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.

Key: Simulations present cumulative effects of changes in: U = unemployment; S = employment by sector; O = employment by occupational category; W_1 = relative remuneration per worker; W_2 = mean real remuneration per worker; M = skill level.

Notes: [1/] Sequential and cumulative effects are presented for changes in the type of labour market variables as explained in the text; [2/] Calculation based on a national poverty line.

The changes in poverty and inequality as obtained through the microsimulations are statistically significant (at 95 per cent confidence) for changes in W_1 and W_2, but not for most other effects for which the changes tend to be within the confidence interval, as can be seen Table A.1 in Appendix A3. As explained above, the confidence intervals were obtained by repeating the randomized assignment process 30 times in Monte Carlo fashion.

4.2. Simulating the impact of cash transfers to the poor

As a further example, we show the direct and indirect (general-equilibrium) effects of an expansion of a cash transfer programme for poor households in Costa Rica. The simulation assumes that from 2011 poor households receive an increment of $25 per child in primary school age.[9] The cash transfer is not made conditional on having a child in the household attending primary school.

The general equilibrium effects of this expanded transfer programme are rather small. At the macro level, as shown in Table 4, the cost of the transfer programme leads to an increase of the fiscal deficit by 0.1 per cent of GDP as compared with the baseline. In the simulated scenario the budget deficit is assumed to be financed entirely through domestic borrowing. Aggregate output is affected as less domestic savings become available for private investment.[10] The impact is very modest, however. As in the baseline, the labour market adjustment principally takes place through real wages. All workers see a decline in real labour incomes, though the impact is somewhat stronger among non-wage, self-employed and family workers. As a result of these general equilibrium effects, average household incomes (as measured in the CGE) fall on balance for non-poor households.

In implementing the microsimulations for this case, non-labour incomes were adjusted first for the value of the increase in cash transfer benefits. Subsequently, the labour market adjustments as simulated through the CGE model were imposed on the micro dataset in the same sequence as this was done for the baseline simulation. Unsurprisingly, most of the impact on poverty is explained by the direct effect of the transfer (Table 5). Moderate poverty falls by 1.6 percentage points in 2011 as compared with the baseline and extreme poverty by 0.8 percentage points. The general equilibrium effects in the form of lower real wages slightly offset the reduction in extreme poverty but show no visible effect on moderate poverty. In 2012, the indirect effects are only slightly larger and the overall finding remains that the transfer helps to reduce poverty despite some adverse side effects on the labour market. The transfer further reduces household income inequality, but only to a small degree. This is a result of the modest size of the transfer.

Table 4 Macro-micro effects of simulating an increase in government cash transfers to poor households for Costa Rica[1/], 2011-2012 (deviation with respect to the baseline) [2/]

	2011	2012
Gross capital formation	-1.0	-1.0
Real GDP at market prices	-0.2	-0.3
Government income	-0.5	-0.6
Government expenditure	0.2	0.1
Fiscal deficit/GDP	0.1	0.1
Unemployment rate	0.1	0.2
Employment	0.0	0.0
Wage-earners, unskilled women	0.2	0.2
Non-wage earners, unskilled women	0.0	0.0
Wage-earners, unskilled men	0.0	0.0
Non-wage earners, unskilled men	0.0	0.0
Wage-earners, skilled women	-0.1	-0.2
Non-wage-earners, skilled women	0.0	0.0
Wage-earners, skilled men	0.1	0.0
Non-wage earners, skilled men	0.0	0.0
Labour income	-0.5	-0.6
Wage-earners, unskilled women	-0.4	-0.5
Non-wage earners, unskilled women	-1.2	-1.3
Wage-earners, unskilled men	-0.4	-0.5
Non-wage earners, unskilled men	-0.8	-0.9
Wage-earners, skilled women	-0.5	-0.6
Non-wage-earners, skilled women	-0.8	-1.0
Wage-earners, skilled men	-0.5	-0.5
Non-wage earners, skilled men	-0.5	-0.7

Source: CGE model and microsimulation results for Costa Rica.
Notes: [1/] The performed simulation is explained in the text; [2/] Percentage deviation for all variables except for fiscal deficit for which results are presented in percentage points of GDP, and poverty and inequality for which absolute changes of the indicator are used.

Table 5 Decomposition of the poverty and inequality effect of simulating an increase in government cash transfers to poor households for Costa Rica, 2011-2012 (Absolute deviation with respect to the baseline)

	2011			2012		
	Direct effect	CGE effect	Total effect	Direct effect	CGE effect	Total effect
Total poverty incidence [1/]	-1.6	0.0	-1.6	-1.5	0.1	-1.4
Extreme poverty incidence [1/]	-0.9	0.1	-0.8	-0.9	0.2	-0.7
Gini - labour income	0.000	0.000	0.000	0.000	0.000	0.000
Gini - per-capita household income	-0.004	0.000	-0.004	-0.003	0.000	-0.004

Source: CGE model and microsimulation results for Costa Rica.
Notes: [1/] Incidence as a percentage of the population. Calculation based on nationally defined poverty lines.

5. CONCLUSIONS

The non-parametric microsimulation methodology explained in this paper allows assessing the impact of labour market changes on income inequality and poverty at the household level, while resolving the assignment problem known from the labour market literature. The approach assumes that labour markets are segmented and determines (as part of a randomized process) which individuals are expected to move in or out of employment and which move from one employment segment to another based on either known or counterfactual information of aggregate labour market changes. The methodology assumes that the distribution of earnings of those who become employed in a particular segment resembles that of the individuals observed to be employed in that segment. There is no direct link to individual attributes as assumed in the parametric approaches that are rooted in human capital theory and model labour market behaviour based on such attributes. The non-parametric approach does account for (observed or simulated) changes in the composition of the labour force by individual endowments, including skill level and sex.

The non-parametric method can be applied even if one has access to only one micro dataset, typically a household or labour force survey. The additional information required is relevant summary information about changes in labour market variables (and/or non-labour incomes) either from observed data or a counterfactual simulated through, for instance, a macroeconomic or a CGE model. Parametric approaches, such as that developed by Bourguignon *et al.* (2001) require availability of the micro datasets from at least two household surveys. A further advantage of the non-parametric approach is that it facilitates a detailed insight in the relative importance of a range of labour market shifts on household level inequality and poverty, including changes in participation rates, unemployment, employment shifts by sector and occupational category, labour and non-labour income changes and individual endowments (such as skill level).

Since these different aspects are introduced in the microsimulation methodology in a particular order, a potential problem of path dependence emerges. Parametric microsimulation methods are equally subject to possible path dependence. However, as shown in Appendix A.2, this problem will be limited as long as labour market shifts are not very large. It was also indicated that path dependence is less of an issue if one accepts the suggested logic of the sequence in which these changes are introduced.

The obvious disadvantage of the non-parametric approach is that it does not explicitly model behaviour in the labour market and instead assumes that individuals change position in the labour market as part of a randomized process in the context of a segmented labour market. The use of a bootstrapping (Monte Carlo) procedure allows to generate confidence intervals and application shows that such intervals typically show narrow margins of error and that only a limited number of repetitions (about 30) tend to be needed to obtain stable confidence intervals. Even so, the approach remains a strong simplification and a fuller specification of labour market behaviour may be preferred. In practice, however, empirical labour market models, especially when applied to developing countries, typically show relatively low explanatory power suggesting much of labour market behaviour is left unexplained. This has been an additional reason to develop the non-parametric approach. If behavioural models leave to much unexplained, then a controlled (*i.e.* within a pre-defined structure), randomized process of labour market behaviour might be as good an approximation of actual behaviour and with the advantage of saving on modelling effort and computational time.

Yet, little comparative analysis has been undertaken to pin down with greater robustness how parametric and non-parametric microsimulations methods compare in terms of outcomes and sensitivity to the assumptions made. Such comparison should be subjected to further research. Based on a by now wide application of the non-parametric method presented here we are confident to conclude that the method is relatively easy to apply and that its outcomes are easy to understand and typically plausible for the labour market outcomes that are simulated. We do not argue that this method is superior to the parametric or behavioural approaches, but it is easy to implement and helps overcome many data constraints to practical applications. The method has also proven to be effective in overcoming insufficient distributional detail present in static and dynamic CGE models to assess the impact of trade, tax reforms or other policy reforms on poverty and income distribution. This combination of CGE modelling and the non-parametric microsimulation method has been applied in a top-down, sequential manner. While this has limitations of its own, it has the advantages of being relatively easy to implement and generating easy-to-understand and tractable outcomes for inequality and poverty.

Notes

[1] A coded version (in STATA) can be obtained from the authors upon request.

[2] In reality there may also be other factors that explain the difference between the actual and simulated change in earnings inequality, *e.g.* demographic changes.

[3] The application of the microsimulation methodology to data sets for Latin American countries suggests that repeating the random selection about 30 times is sufficient to generate stable confidence intervals and repeating it more times would not yield statistically significant differences in the outcomes for labour earnings inequality or indicators for household income inequality and poverty (see Ganuza *et al.*, 2002; Vos *et al.*, 2006).

[4] Mean incomes per decile are calculated (per segment) for employed persons according to the used individual characteristics. These means are assigned to newly employed or to already employed persons who change sector of economic activity, occupational category or skill category.

[5] Please note that path dependence is also an issue in alternative methods. For instance, the parametric approach of Bourguignon *et al.* (2001, 2002a,b) performance simulations for a price effect (such as changes in wage rates), a participation effect (emanating from labour supply behaviour) and a population effect. The outcomes of any sequential decomposition of distributional changes on account of these three effects will be sensitive to the order in which these effects are accounted for.

[6] It should be emphasized, however, that the microsimulation methodology does not allow certain sequences. For instance, the simulation of the impact of an alteration in the remuneration structure assumes full information on both the sector of economic activity and the occupational category. Hence, changes in the remuneration structure cannot be preceded immediately by the simulated changes in either unemployment or participation rates.

[7] Alternative approaches that model labour supply behaviour might capture such effects as part of the behavioural model underpinning the microsimulation method. In the non-parametric approach such effects would need to be captured through the economy-wide model and would require such a model would capture the impact of different types of transfers to households or individuals on labour market behaviour.

[8] Ideally, the CGE model would use data for base year values for labour market and household incomes that are derived from the same survey to be used for the microsimulations to avoid further 'noise' in the results because of discrepancies in the underlying information.

[9] The CGE model for Costa Rica allows assigning transfers by type of household as it distinguishes four groups of households: urban poor and non-poor and rural poor and non-poor. It should be noted that poor households are defined as the bottom 40 per cent of the rural and urban income distribution, respectively.

[10] In this simulation the model assumes an investment-driven macroeconomic closure; that is, savings adjust to meet investment demand. Also, the government account is assumed to adjust this way; that is, the government will try to mobilize private savings through domestic borrowing to finance the budget deficit.

[11] This explanation of the procedure has been adapted from the original specification of Ganuza *et al.* (2002:83-86) and which has been followed in a large number of country case studies. In this appendix, reference to the counterfactual labour market variables is in terms of that for "year *t**", but may be understood equally in an application where the counterfactual is generated through an economy-wide model. The explanation here only refers to changes in labour market variables. See the text for how to include changes in non-labour incomes.

[12] This appendix was adapted from earlier work by De Jong (2001) and Vos and De Jong (2001).

REFERENCES

Almeida dos Reis J G and Paes de Barros R (1991) 'Wage Inequality and the Distribution of Education: A Study of the Evolution of Regional Differences in Inequality in Metropolitan Brazil', *Journal of Development Economics*, 34, 117-143.

Atkinson, A B and Bourguignon F (2000) 'Introduction: Income Distribution and Economics', in Atkinson A B and F Bourguignon (Eds.), *Handbook of Income Distribution, Volume 1*, Amsterdam: North Holland, Elsevier, 1-58.

Behrman, J R (1999) 'Labor Markets in Developing Countries', in Ashenfelter O, Layard R and D Card (Eds.), *Handbook of Labor Economics*, *Volume 3B*, Amsterdam: North Holland, Elsevier, 2859-2939.

Blinder, A S (1973) 'Wage Discrimination: Reduced Form and Structural Estimates', *Journal of Human Resources*, 8, 436:455.

Bourguignon F, Fournier M and Gurgand M (2001) 'Fast Development with a Stable Income

Distribution: Taiwan, 1979-1994', *Review of Income and Wealth*, 47(2), 139-164.
Bourguignon F, Pereira da Silva L A and Stern N (2002a) 'Evaluating the Poverty Impact of Economic Policies: Some Analytical Challenges', Washington D.C.: World Bank.
Bourguignon F, Robilliard A-S and Robinson S (2002b) 'Representative versus Real Households in the Macro-economic Modeling of Inequality', Washington D.C.: IFPRI.
Cicowiez M and Sánchez M V (2009) 'MACEPES: Modelo de Análisis de Choques Exógenos y de Protección Económica y Social', Technical Document prepared for the project *Implications of Macroeconomic Policy, External Shocks and Social Protection Systems for Poverty, Inequality, and Social Vulnerability in Latin America and the Caribbean*, Sub-regional Office of ECLAC in Mexico and United Nations Department of Economic and Social Affairs (UN/DESA), (mimeo, April).
De Jong N (2001) 'Decomposing Changes in Poverty and Inequality: The Case of Urban Panama in the 1990s', Paper prepared for Economic Research Seminar, The Hague: Institute of Social Studies (mimeo).
Frenkel R and González Rozada M (2002) 'Argentina: macroeconomic behaviour, employment and income distribution in the 1990s', in Vos R, Paes de Barros R and L Taylor (Eds.), *Economic Liberalization, Distribution and Poverty. Latin America in the 1990s*, Cheltenham (UK), Northampton (MA): Edward Elgar Publishers, 89-124.
Ganuza E, Paes de Barros R and Vos R (2002) 'Labour market adjustment, poverty, and inequality during liberalization', in: Vos R, L Taylor and R Paes de Barros (Eds.) *Economic Liberalization, Distribution and Poverty. Latin America in the 1990s*, Cheltenham (UK), Northampton (US): Edward Elgar Publishers, 54-88.
Hartog J (1985) 'Earnings Functions. Testing for the Demand Side', *Economics Letters*, 19, 281-285.
Hartog J (1986) 'Earnings Functions. Beyond Human Capital', *Applied Economics*, 18, 1291-1309.
Mincer J (1958) 'Investment in Human Capital and Personal Income Distribution', *Journal of Political Economy*, 66(4), 281-302.
Mincer J (1974) *Schooling, Experience and Earnings*, Columbia University Press: New York.
Oaxaca R L (1973) 'Male-Female Wage Differentials in Urban Labor Markets', *International Economic Review*, 14, 693:709.
Orcutt G H (1957) 'A new type of socio-economic system', *Review of Economics and Statistics*, 80, 1081-1100.
Paes de Barros R (1999) 'Evaluando el Impacto de Cambios en la Estructura Salarial y del Empleo sobre la Distribución de Renta', Rio de Janeiro: IPEA (mimeo).
Paes de Barros R and Leite P (1998) 'O Impacto da Liberalizaçao sobre Distribuiçao de Renda no Brasil', Rio de Janeiro: IPEA (mimeo).
Sánchez M V (2004) *Rising Inequality and Falling Poverty in Costa Rica's Agriculture during Trade Reform: A Macro-micro General Equilibrium Analysis*, Maastricht: Shaker.
Sánchez M V and Vos R (2005) 'Impacto del Tratado de Libre Comercio con Estados Unidos en el crecimiento, la pobreza y la desigualdad en Panamá: Una evaluación ex ante usando un modelo de equilibrio general computable dinámico', Project Report prepared for Panama's government, The Hague, Mexico and Panama City: Institute of Social Studies, Sub-regional Office of ECLAC in Mexico and UNDP.
Sánchez M V and Vos R (2006) 'Impacto del CAFTA en el crecimiento, la pobreza y la desigualdad en Nicaragua', Managua: Ministerio de Fomento a la Industria y el Comercio and UNDP.
Sattinger M (1993) 'Assignment Models of the Distribution of Earnings', *Journal of Economic Literature*, 31(2), 831-880.
Savard L (2003) 'Poverty and Income Distribution in a CGE-Household Micro-Simulation Model: Top-Down/Bottom Up Approach', Cahiers de recherché/Working Paper 03-43, International Development Research Centre (IDRC), Dakar.
Vos R and De Jong N (2001) 'Rising Inequality during Economic Liberalization and Crisis: Macro or Micro Causes in Ecuador's Case? ', Working Paper Series No. 326, The Hague: Institute of Social Studies.
Vos R, Taylor L and Paes de Barros R (Eds.) *Economic liberalisation, distribution and poverty. Latin America in the 1990s*, Cheltenham (UK), Northampton (US): Edward Elgar.
Vos R, Ganuza E, Morley S and Robinson S (2006) *Who Gains from Free Trade? Export-led-Growth, Inequality and Poverty in Latin America*, London: Routledge.

APPENDIX A.1 The non-parametric micro-simulation methodology step by step[11]

For year *t* (or any observed distribution) an alternative structure of the labour market is defined on the basis of data for year *t** (or a counterfactual distribution). In each iteration of the microsimulations a random number drawn from a normal distribution is assigned to each individual (of a sub-group) of the population in year *t*. This number is used to rank the individuals. The following simulations are considered (separately or sequentially):

1. Change of the *participation rate* (*P*) of each group *j* of the population.

- *Objective*: Determine the indices of poverty and inequality if the participation rates in year *t* were to be equal to those in year *t**.
- *Procedure*: Within each group *j* the persons at working age are in the first place ranked according to labour force status – starting with the economically active – and in the second place on the basis of the random numbers. If for a group *j* the participation rate in year *t** is

at most equal to that in year t, the last economically active individuals of types j with lower participation rates in year t^* than in year t will be reclassified as economically inactive and their labour income is set to zero. If the corresponding participation rate is higher, it will subsequently be determined whether the new economically active persons will be employed. If so, they will be randomly assigned a labour income.

2. Change of the *unemployment rate* (*U*) of economically active persons of type *j*.

- *Objective*: Determine the indices of poverty and inequality if the unemployment rates in year t were to be equal to those in year t^*.
- *Procedure*: Only the economically active population is considered. Within each group j the individuals are in the first place ranked according to employment condition – starting with the employed – and in the second place on the basis of the random numbers. For the types j with higher rates of unemployment in year t^* than in year t, the last employed persons of each type j are reclassified as unemployed and their labour income is set to zero. In case of types j with lower rates of unemployment in year t^*, the newly employed are grouped into deciles on the basis of the random numbers and assigned the mean labour income of the corresponding decile of employed persons in year t.

3. Change of the *sector of activity* (*S*) of wage employees and non-wage workers of type *j*

- *Objective*: Determine the indices of poverty and inequality if the proportion of persons employed by sector in year t were to be the same as in year t^*.
- *Procedure*: Only the employed population is considered. Mean incomes per decile of employed persons of type j in each sector are calculated for both occupational categories. Within each group j the individuals are in the first place ranked according to sector of activity and in the second place on the basis of the random numbers. In groups in which the proportion of persons working in sector 2 is lower in year t^* than in year t, the first persons of sector 2 move to sector 1, if there were two sectors, for example. In groups j in which the proportion of persons in sector 2 is higher in year t^* than in year t, the last persons of sector 1 move to sector 2. Within each group j the persons who change from one sector to the other are classified into deciles on the basis of their random number and their labour income is replaced by the corresponding mean income of the decile of all persons who in year t are actually working in the sector of destination.

4. Change of the *occupational category* (*O*) of employed persons of type *j* in each sector of activity.

- *Objective*: Determine the indices of poverty and inequality if the proportion of wage employees in year t were to be the same as in year t^*.
- *Procedure*: Only the employed population is considered. Mean incomes are calculated per decile of wage employees and non-wage workers of type j in each sector of activity. For both sectors of activity within each group j the individuals are in the first place ranked according to occupational category – starting with the wage employees – and in the second place on the basis of the random numbers. In groups in which the proportion of wage employees is lower in year t^* than in year t, the last wage employees become non-wage workers. In groups in which the proportion of wage employees is higher in year t^* than in year t, the first non-wage workers become wage employees. Within each group j the persons who change from one occupational category to the other are classified into deciles on the basis of their random number and their labour income is replaced by the corresponding mean income of the decile of all persons who in year t are actually working in the occupational category of destination.

5. Change of the *remuneration structure* (W_1)

- *Objective*: Determine the indices of poverty and inequality if the structure of labour incomes by segment k in year t were to be that of year t^*.
- *Procedure*: Only the employed population is considered. Mean labour incomes are calculated for, say, each of the 16 groups jk of employed persons – for 4 types of workers j and 4 segments k defined for 2 sectors and 2 occupational categories, as well as an overall mean, for both year t^* and year t. Subsequently, the following relative mean incomes are calculated for year t^*:

$$s_{jk} = \frac{\overline{yp_{jk}^{t*}}}{\overline{yp^{t*}}} \tag{A.1}$$

The mean labour income in year t of each group jk is multiplied by the corresponding s_{jk} in order to obtain a new mean labour income for each group jk in prices of year t:

$$\overline{yp}^{*}_{jk} = \frac{\overline{yp_{jk}^{t*}}}{\overline{yp^{t*}}} \cdot \overline{yp^{t}} \tag{A.2}$$

In turn, the new mean incomes of the groups jk are expressed as a proportion of the corresponding mean in year t, and subsequently the year t labour income of each individual i in group jk is multiplied by the proportion for the group:

$$yp_{jki}^{t*} = \frac{\overline{yp}^{*}_{jk}}{\overline{yp_{jk}^{t}}} \cdot yp_{jki}^{t} \tag{A.3}$$

As a final step, all the individual incomes are multiplied by an adjustment factor, so as to keep the overall mean income constant.

6. Change of the *level of remuneration* (W_2).

- *Objective*: Determine the indices of poverty and inequality if the level of real incomes of year *t* were to be that of year *t**.
- *Procedure*: Only the employed population is considered. New labour incomes are calculated by multiplying the year *t* labour income of each income recipient by the ratio of mean income in year *t** (in currency of year *t*) to that in year *t*:

$$yp_{jki}^{t**} = \frac{\overline{yp^{t*}}}{\overline{yp^{t}}} \cdot yp_{jki}^{t} \qquad \text{(A.4)}$$

7. Change of the *level of skill* (*M*) of employed individual *j* in segment *k*.

- *Objective*: Determine the indices of poverty and inequality if the proportion of skilled workers in year *t* were to be same as in year *t**.
- *Procedure*: Only the employed population is considered. Mean incomes are calculated per decile of employed individual *j* in each segment *k*. Individuals within each group defined by sex and age, for example, and segment are in the first place classified according to skill – starting with the unskilled workers – and in the second place on the basis of the random numbers. In groups in which the proportion of skilled workers is higher in year *t** than in year *t*, the last unskilled workers are reclassified as skilled workers. In case of groups with lower proportions of skilled workers year *t**, the first skilled workers move to the category of unskilled workers. Within each group *j* the persons who change from unskilled to skilled are classified into deciles on the basis of their random number and their labour income is replaced by the mean income of the corresponding decile of all persons who are actually skilled in year *t*. In the reverse case, the actual year *t* incomes are replaced by that of the corresponding decile of unskilled workers.

Due to changes in the participation rate and the unemployment rate in the sequential simulation it is possible that persons become classified as employed, but that there is no information concerning occupational category for these persons. For this reason, in the part of the sequential simulations in which the employment structure according to sector of activity is changed, mean proportions of persons employed in, say, the non-agricultural sector in year *t** are used (instead of different proportions for, say, wage employees and non-wage workers separately) in cases of lack of information concerning the occupational category.

APPENDIX A.2 DOES PATH DEPENDENCE MATTER?[12]

The simulated change in earnings inequality in the sequential simulation is defined as:

$$\overline{\Delta^{sim} I(yp_i)} \equiv \sum_{s=1}^{6} [\bar{I}_s^{sim}(yp_i) - \bar{I}_{s-1}^{sim}(yp_i)] \qquad \text{(A.5)}$$

where:

$$\bar{I}_1^{sim}(yp_i) = \bar{I}(P^*, U, S, O, W_1, M, c),$$

$$\bar{I}_2^{sim}(yp_i) = \bar{I}(P^*, U^*, S, O, W_1, M, c),$$

$$\bar{I}_3^{sim}(yp_i) = \bar{I}(P^*, U^*, S^*, O, W_1, M, c), \ldots.., \text{etc.}$$

and

$$\bar{I}_0^{sim}(yp_i) = I(yp_i) = I(P, U, S, O, W_1, M, c)$$

In equation A.5, $\bar{I}$ is the mean value of the summary statistic, $\overline{\Delta I}$ the mean simulated change in the summary statistic, while *P*, *U*, *S*, *O*, W_1 and *M* are labour market variables as defined in the text and *c* refers to the sex and age distribution among members of each household.

The total simulated change in inequality can be decomposed in changes due to alteration of participation rates, changes due to alteration of unemployment rates given the simulated incomes as a result of altering the participation rates, and a residual *R*.

$$\overline{\Delta^{sim} I(yp_i)} \equiv \sum_{s=1}^{2} [\bar{I}_s^{sim}(yp_i) - \bar{I}_{s-1}^{sim}(yp_i)] + R \qquad \text{(A.6)}$$

Changes in the distribution as a consequence of shifts in the participation (*P*) and unemployment (*U*) rates, respectively, can be defined as:

$$\begin{aligned}\overline{\Delta_P^{sim} I(yp_i)} &= \bar{I}_1^{sim}(yp_i) - I(yp_i) \\ &= \bar{I}(P^*, U, S, O, W_1, M, c) - I(yp_i)\end{aligned},$$

and

$$\overline{\Delta_U^{sim} I(yp_i)} = \bar{I}(P, U^*, S, O, W_1, M, c) - I(yp_i),$$

and the cumulative sequential change of a change in the participation and unemployment rates as:

$$\overline{\Delta_U^{sim} I(yp_i^P)} = \overline{I_U^{sim}(yp_i^P) - I(yp_i^P)},$$

so that we can write:

$$\overline{\Delta^{sim} I(yp_i)} \equiv \overline{\Delta_P^{sim} I(yp_i)} + \overline{\Delta_U^{sim} I(yp_i^P)} + R \qquad \text{(A.7)}$$

The second component in this expression is the distribution that results from altering the unemployment rates, given the simulated incomes as a result of altering the participation rates. Hence it includes an interaction term:

$$\overline{\Delta^{sim} I(yp_i)} \equiv \overline{\Delta_P^{sim} I(yp_i)} + \overline{\Delta_U^{sim} I(yp_i)} + \overline{\Delta_U^{sim} I(\Delta yp_i)} + R \tag{A.8}$$

If the sequence of altering *P* and *U* were reversed, the result would become:

$$\overline{\Delta^{sim} I(yp_i)} \equiv \overline{\Delta_U^{sim} I(yp_i)} + \overline{\Delta_P^{sim} I(yp_i^U)} + R \tag{A.9a}$$

or

$$\overline{\Delta^{sim} I(yp_i)} \equiv \overline{\Delta_P^{sim} I(yp_i)} + \overline{\Delta_U^{sim} I(yp_i)} + \overline{\Delta_U^{sim} I(\Delta yp_i)} + R \tag{A.9b}$$

For the relative importance of altering respectively parameters *P* and *U* to be more or less the same in both sequences, it should hold that the Marginal effects of altering *P* or *U* should be similar, irrespective of the sequence:

$$\overline{\Delta_P^{sim} I(yp_i^U)} \cong \overline{\Delta_P^{sim} I(yp_i)} \tag{A.10a}$$

and

$$\overline{\Delta_U^{sim} I(yp_i^P)} \cong \overline{\Delta_U^{sim} I(yp_i)} \tag{A.10b}$$

This means that the interaction effects should be relatively small.

Similar notation can be used for the (simulated) changes in the poverty and inequality indices of per capita income. In the simulations, the effect of altering characteristics of individuals in the household and other factors will appear in the residuals – *i.e*. the difference between the observed change in the indices and changes simulated by imposing an alternative labour market structure λ*.

APPENDIX A.3: ADDITIONAL MICROSIMULATION RESULTS

Table A.1 Sequential and cumulative effects and confidence intervals for changes in the labour-market parameters for the baseline scenario for Costa Rica[1/], 2012

Variable	Obs.	Mean	Standard Error	[95% Confidence Interval]	
U					
Total poverty incidence [2/]	30	20.45249	0.00685	20.43850	20.46649
Extreme poverty incidence [2/]	30	4.21342	0.00788	4.19730	4.22954
Gini - labour income	30	0.46117	0.00003	0.46111	0.46123
Gini - per-capita household income	30	0.49661	0.00003	0.49655	0.49668
U+S					
Total poverty incidence [2/]	30	20.48056	0.00658	20.46711	20.49402
Extreme poverty incidence [2/]	30	4.22287	0.00857	4.20533	4.24040
Gini - labour income	30	0.46119	0.00005	0.46109	0.46129
Gini - per-capita household income	30	0.49662	0.00005	0.49652	0.49672
U+S+O					
Total poverty incidence [2/]	30	20.44773	0.00622	20.43501	20.46044
Extreme poverty incidence [2/]	30	4.20806	0.00825	4.19119	4.22492
Gini - labour income	30	0.46098	0.00005	0.46088	0.46108
Gini - per-capita household income	30	0.49647	0.00005	0.49636	0.49657
$U+S+O+W_1$					
Total poverty incidence [2/]	30	18.12911	0.01060	18.10742	18.15079
Extreme poverty incidence [2/]	30	3.76363	0.00698	3.74935	3.77790
Gini - labour income	30	0.44659	0.00005	0.44648	0.44669
Gini - per-capita household income	30	0.47856	0.00005	0.47845	0.47867
$U+S+O+W_1+W_2$					
Total poverty incidence [2/]	30	16.56747	0.01072	16.54556	16.58939
Extreme poverty incidence [2/]	30	3.58093	0.00593	3.56881	3.59305
Gini - labour income	30	0.44667	0.00005	0.44656	0.44677
Gini - per-capita household income	30	0.47851	0.00005	0.47840	0.47862
$U+S+O+W_1+W_2+M$					
Total poverty incidence [2/]	30	16.52524	0.01902	16.48635	16.56413
Extreme poverty incidence [2/]	30	3.59980	0.01135	3.57659	3.62302
Gini - labour income	30	0.44682	0.00007	0.44668	0.44696
Gini - per-capita household income	30	0.47790	0.00009	0.47771	0.47809

Source: CGE model and microsimulation results for Costa Rica.
Notes: [1/] Sequential and cumulative effects are presented for changes in the labour market variables as explained in the text; [2/] The poverty incidence is expressed as a percentage of the total population and was estimated using nationally defined poverty lines.