CHAPTER 21

Income Distribution in Computable General Equilibrium Modeling

François Bourguignon*, **Maurizio Bussolo****
*Paris School of Economics
**World Bank

Abstract

Studying the poverty and income distribution effects of macroeconomic policies or shocks requires a methodology that accounts on the one hand for the nature of the policy or shock being studied and their aggregate impact on the economy and, on the other hand, the heterogeneity of their overall effects among individuals or households at the micro level. Standard evaluation methods are of little use here. Methods based on the comparison between individuals exposed and not exposed to the policy or shock are clearly not applicable since, by definition, all individuals are affected by any policy with some macroeconomic dimension. The evaluation methodology thus needs to rely not only on a micro but also on a macro counterfactual within some kind of general equilibrium setting.

This chapter reviews the growing literature that addresses this challenge by combining macro and micro modeling in various ways. It describes existing methodologies and their underpinning links with economic theory and presents some applications.

The chapter is structured according to the nature of the computable general equilibrium (CGE) macro model being used and how the macro and micro components of the methodology are connected together and/or integrated. It thus starts from a simple framework where the two components are handled separately: the CGE model is static and Walrasian and the unidirectional link from the macro to the micro component essentially consists of changes in prices. It then moves to more integrated variants, where results from both the CGE and micro models are looped back and forth, and also to non-Walrasian specifications where both price and aggregate quantity changes link the macro and the micro components. Finally the chapter considers dynamic settings where the static CGE model on the macro side is replaced by a recursive-dynamic model, whereas the micro data base is allowed to change over time through both exogenous demographic and policy-induced endogenous economic transitions. Illustrations of that approach are drawn from the Global Income Distribution Dynamics (GIDD) model recently developed at the World Bank.

Keywords

CGE, macro-micro models, static and dynamic microsimulation, inequality, income distribution

JEL classification codes

C15, C54, C68, D33, D58, I30, O15

Handbook of CGE Modeling - Vol. 1 SET
ISSN 2211-6885, http://dx.doi.org/10.1016/B978-0-444-59568-3.00021-3

21.1 INTRODUCTION

Tools for policy analysis concerned with distributional issues tend to be either micro or macro. Typically, microsimulation models based on representative household samples are used for studying reforms of tax-benefit systems, of the delivery of some public goods like education or healthcare, or changes in the regulation of the labor market (retirement age, minimum wage, etc.). Macro modeling, on the other hand, would typically be used to figure out the potential impact of reforms in tax, trade, finance or monetary (including exchange rate) policies on the sector structure of the economy, the level of employment and wages, and their structure by skill and the returns to capital. From that aggregate information, one can then guess which groups in the population will benefit or lose from the reform being analyzed.

The need for combining these two approaches to policy analysis in analyzing distributional issues is quite obvious. Aggregate outcomes from macro models are insufficient to figure out how individual households might be affected by the policy being analyzed. For instance, knowing that wages will go up by x% or unemployment will go down by y% is not enough to know how many households will be lifted from, or on the contrary fall into poverty. Likewise, seemingly micro-oriented policies like a reform in cash transfers or in the pension system may have important macro effects through labor supply and savings. In turn, these macro effects of micro-oriented policies will react on the wages, employment and returns to wealth affecting the income distribution within the population.

Understanding this link between macro and micro effects is of utmost importance. For micro-oriented policies — essentially redistribution policies — final effects might end up being different from what was intended because of macro—micro feedbacks. For macro-oriented policies, the distributional impact of the measures being considered should inform on the identity of winners and losers, the possibility of compensating the latter, and, more generally, the political economy of the reforms.

Despite such importance and despite progresses both in microsimulation and macro modeling, relatively little has been achieved in combining the two approaches so as to get a comprehensive policy-oriented economic model allowing to take systematically into account the distributional impact of policies. It is true that macro models, even of the most sophisticated dynamic stochastic general equilibrium (DSGE) type, often try to take into account some heterogeneity within the population.[1] However, they do it in a highly abstract way that does not really permit dealing with the full heterogeneity of the micro distribution of income and consumption behavior. Most models explicitly dealing with the income distribution are essentially top-down models where a few macro

[1] See, e.g. Castaneda *et al.* (2003).

variables like changes in the level of wages or returns to wealth are used to scale up or down individual incomes in a micro database without any feedback.[2]

On the other side, microsimulation work most often stops short of figuring out the impact of changes in individual behavior on macro equilibrium and therefore on prices or employment. From that point of view, there is something odd when recalling that the father of microsimulation, G. Orcutt, called for such a micro—macro integration more than 40 years ago. He actually supervised the construction of such a model (DYNASIM) in the mid 1970s,[3] but its performance results were somewhat disappointing and a posterior version of the model returned to a simple top-down macro—micro framework (DYNASIM II).

Several approaches have been proposed over the last 20 years or so to go beyond this simple top-down approach, and to develop an integrated macro—micro framework allowing us to study simultaneously the distributional and macro impact of policy reforms. It should be clear from the outset that the task of developing such macro—micro framework is daunting. In fact, there are challenges from various fronts. The usual problem of data availability and quality is, in the case of a macro—micro model, compounded because data from national accounts and from household surveys — the two most common sources of macro and micro data — need to be reconciled. Additionally, the theoretical pedigrees of micro and macro models are often at odds. Probably the most well-known issue is that of the 'aggregation problem'. This problem exists whenever the aggregate agents' behavior, such as aggregate private demand, cannot be 'treated as if it were the outcome of the decision of a single maximizing consumer' (Deaton and Muellbauer, 1980, p. 148). The conditions for aggregation are quite severe and break down easily in presence of discontinuities and non linearities. In these cases, macro models with representative agents cannot be substitutes for micro models. Consider, for example, the case of participation in the labor market. Theory says that participation should increase with the wage rate. At the micro level, this can be obtained as increasing hours for an already employed individual or as the entrance in the market of someone who was not working. The effects, in terms of income distribution, may be very different in these two cases. However, at the macro level, it is impossible to model these two alternatives with an aggregate representative worker (even if the number of representative workers is greatly increased). And comparable cases where aggregation conditions do not hold arise in almost all markets with imperfection and asymmetric information: goods, credit, education and others.

Notwithstanding these issues and the fact that definitive solutions to these challenges are not yet available, empirical research has advanced in its ability to describe with some

2 In the CGE modeling literature, the first model to do so, although in a parametric way rather than with a micro database, was Adelman and Robinson (1978).

3 See Orcutt *et al.* (1976).

accuracy the likely impact of policy reforms on the distribution of income or, equivalently, the identity of the losers and winners. This chapter reviews various available approaches,[4] analyzing their strengths and weaknesses. It illustrates several of the methodological issues involved in bridging micro and macro modeling within a recursive-dynamic framework by discussing the specification of, and some simulations performed with the Global Income Distribution Model (GIDD) of the World Bank. The chapter ends with considerations on possible next steps in improving existing tools.

The chapter is organized along methodological lines and, more specifically, it is structured according to the characteristics of the macro framework underlying the whole macro–micro model. Although the macro modeling framework may be of different types, the focus will be on computable general equilibrium (CGE) models of different types. Section 21.2 is devoted to the macro–micro link when the underlying CGE model is *static and essentially Walrasian*. The point of departure will be the widely used 'top-down approach' where the two parts of the model are handled separately, the link between them consisting of price changes computed in the macro, or CGE part, at the top and feeding a microsimulation model based on a standard households survey at the bottom. Then, bottom-up and more or less integrated versions of the macro–micro modeling framework aimed at taking into account feedback effects will be discussed. Section 21.3 concentrates on extensions of the preceding approaches when the macroeconomic framework is *not Walrasian* anymore, and the basic top-down link between the macro and the micro parts includes not only changes in prices, but also changes in quantities, typically employment rationing or some sectoral allocation of total employment, e.g. across sectors or urban–rural areas.

Sections 21.4 and 21.5 extend the preceding instruments to a simple *dynamic framework*. The static CGE model on the macro side is replaced by a recursive-dynamic model, whereas the structure of the household sample used in the static approach is allowed to change over time through exogenous demographic, and possibly endogenous economic transitions, some of them policy induced. An illustration of this approach, certainly the most complete and the most attractive way of modeling micro–macro links at this stage, although not devoid of major difficulties, is provided in Section 21.5, which is devoted to a presentation of the GIDD model developed over the last 5 years or so at the World Bank.

It should be noted that the various methods illustrated in this chapter have been applied to study the distributional effects a wide range of issues such as trade policy reform, financial crises, reforms of the financial sector, subsidies on agricultural production prices, workfare programs, cash transfer programs and scaling up of other poverty alleviation policies, and other policies or shocks. There is no perfect model and different variants focus on different dimensions of the issues: behavioral responses, imperfection in the

4 A brief review of these various approaches is provided in Brown *et al.* (2007), citing the typology drawn by Baekgaard (1995).

markets, dynamic longer-term effects. The choice of the model will depend on the question asked by the researcher. For example, if the researcher is interested in the short-term impact of reduction in tariffs, then using a static top-down Walrasian CGE model linked to simple micro accounting may be the best methodological choice. Whereas if the question is about the interaction of trade liberalization with labor markets, the sectoral adjustment of the economy, and the growth and distribution effects of trade reform, then a dynamic model that perhaps accounts for the imperfections of the labor market should be a better suited model. This chapter presents some applications of these approaches, but its main objective is to describe the methods and their underpinning links with the theory. For more examples of these techniques, the interested reader could consider Bourguignon and Spadaro (2006) and Bourguignon *et al.* (2008) two volumes on the impact of macroeconomic policies on income distribution and poverty.

21.2 STATIC DISTRIBUTION ORIENTED MACRO–MICRO MODELS BASED ON WALRASIAN CGE

It will be convenient to start this review of static distribution-oriented micro–macro models with the simple *top-down* approach, based on a Walrasian CGE. This approach typically applies to identifying the first-round effects of a macro policy reform, say a trade or an indirect tax reform, on the welfare distribution in the population of households without taking into account macro feedback effect of these changes. The symmetric *bottom-up* approach for evaluating the macro effects of a micro-oriented policy is equally simple although it is less easy to justify why feedbacks are ignored in that case. Various ways of taking these feedbacks into account will then be discussed, before considering the possibility of building fully integrated macro–micro models, where both macro and micro effects of a policy reform are jointly simulated.

21.2.1 Top-down approach: justification and limitation

Figure 21.1 illustrates this approach. A change in the exogenous variables ΔX[5] in the CGE model at the top generates changes $\Delta \mathbf{p}$ in the price system, including wages and the returns to fixed assets, and a change in real aggregates $\Delta \mathbf{Y}$. The components of $\Delta \mathbf{p}$ are the 'linkage variables' between the macro and the micro part of the full model. When plugged into the micro model at the bottom, those changes in prices generate in turn for each household i in the household survey used as the database for the micro model a change in welfare, ΔW_i, a change $\Delta \mathbf{c}_i$ in the vector of net consumptions and a change $\Delta \mathbf{l}_i$ in the vector of labor supplies – assuming various types of labor and possibly several adults in the household.

The top-down approach stops at that level. Clearly there seems to be something wrong in doing so since the feedback effect from the micro to the macro model that may

[5] Throughout this chapter, bold notations refer to vector variables, and normal characters to scalar variables.

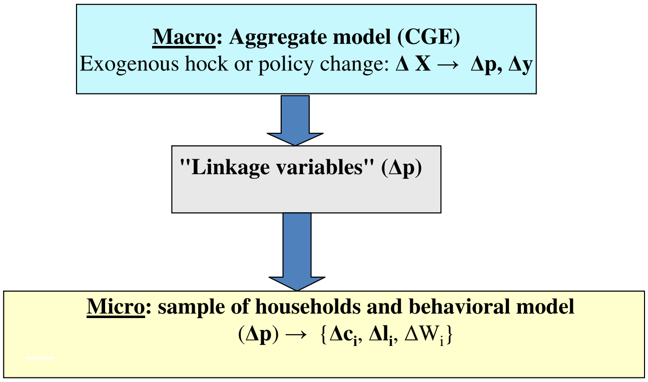

Figure 21.1 Distribution effects of macro policies: top-down macro–micro approach.

go through taking into account the aggregate change in labor supply, and possibly in total net consumption, resulting from summing the $\Delta\mathbf{c}_i$ ($\Delta\mathbf{C}$) and the $\Delta\mathbf{l}_i$ ($\Delta\mathbf{L}$) is not taken into account. In that sense, one may consider that this top-down approach essentially describes the first-round effects of the exogenous (policy) change, ΔX. Practically, however, not taking into account the second round effects, i.e. the macro effects of behavioral responses $\Delta\mathbf{c}_i$ and $\Delta\mathbf{l}_i$, may be quite misleading. It is clear, for instance, that the effect of that behavioral response on the budget of the government is simply ignored.

Of course, this is not completely true. If correctly specified, the CGE model at the top must include some representation of an 'aggregate' or representative household that responds to changes in the price system through changes in its vector of consumption and labor supply. Let $\Delta\mathbf{C}^*$ and $\Delta\mathbf{L}^*$ be those responses. This means that the outcome of the CGE model, Δp, used as 'linkage variables' in Figure 21.1 includes the behavioral response of this representative household. Under these conditions, the issue is not whether household behavioral responses in the micro model are 'ignored' but whether they are accurately represented by the behavior of the representative household as specified in the CGE model. In other words, the issue is whether:

$$\Delta\mathbf{C}^* \neq \text{or} = \Delta\mathbf{C} = \sum_i \Delta\mathbf{c}_i$$

$$\Delta\mathbf{L}^* \neq \text{or} = \Delta\mathbf{L} = \sum_i \Delta\mathbf{l}_i$$

The answer to that question is straightforward. It depends on the aggregation properties of the consumption and labor supply behavior of households in the micro model. If the behavior of all households is assumed to be identical (but their characteristics are different) and is specified as a linear expenditure model of consumption and leisure, then $\Delta\mathbf{C}$ and $\Delta\mathbf{L}$ will be identical to $\Delta\mathbf{C}^*$ and $\Delta\mathbf{L}^*$ and there is no need to take into account the micro–macro feedback for evaluating the effect of the policy on the distribution of welfare (ΔW_i). The top-down approach is sufficient. On the contrary, if aggregation is not perfect, either because households do not have the same behavior or because their behavior cannot be aggregated because of strong non-linearities, then the top-down approach gives only some approximation of the effect of the policy change ΔX.

The aggregation properties of micro behavior have been intensively studied in the literature. Note that the argument above refers to 'perfect aggregation' where aggregate demand depends on aggregate (full-time) income and not on the distribution of individual characteristics, including wage rates, in the population of households. The case of 'exact aggregation' where aggregate demand depends on both aggregate income and some simple general characteristics of the income distribution is not consistent with this top-down approach.[6]

A case that can be easily handled while maintaining the full validity of the top-down approach is where households can be divided into groups with identical consumption and labor supply behavior with perfect aggregation property. It is then sufficient to introduce in the CGE model as many representative households as there are groups of behaviorally identical households. Perfect aggregation ensures that the consumption and labor supply of this representative household behaves as the sum of individual households in the corresponding group. At the same time, the presence of various representative households in the CGE model permits to take into account some distributional features of the policy being analyzed at the top level of the micro—macro framework. CGE models with several representative households are quite common. For instance, Adelman and Robinson (1978), one of the first CGE models explicitly dealing with income distribution, was built along these lines. The difference with the present top-down macro—micro approach was that the income distribution within groups of behaviorally identical households was represented parametrically rather than generated endogenously on the basis of a sample of households.[7]

There are numerous examples of the top-down approach to the evaluation of distributional effects of macro policies in the literature. Note, moreover, that in many of them, the authors do not really model the response of households to price changes as this response or, more exactly, its difference with the way the aggregate response is modeled in the CGE model is ignored. For instance, Lokshin and Ravallion (2008) analyzing the effect of lowering the tariff on wheat imports in Morocco simply estimate the change in the welfare of households in a household survey through:

$$\Delta W_i \cong -\sum_k c_{ik} \Delta \mathrm{p}_k,$$

[6] For a recent survey of aggregation issues and an in-depth discussion of 'exact aggregation', see Blundell and Stoker (2005).

[7] For a review of the way CGE models include distributional features through representative households, see Lofgren *et al.* (2003).

where Δp is obtained by simulating the drop in wheat tariff in a conventional CGE model and c_{ik} is the initial (net) consumption of good k. Indeed, it is well-known from the envelope theorem that for an optimizing household the welfare variation due to a change in prices can be evaluated without taking into account behavioral responses. In welfare economics, the preceding expression corresponds to the 'equivalent variation' of income that makes consumer i indifferent between the initial and the new vector of prices. Then, an income equivalent of the change in social welfare may be approximated by adding individual equivalent variations weighted by the marginal social welfare of the income of consumer i in a standard social welfare function. 'Equivalent variations' may also be added to initial income. Social welfare as expressed by combining mean income and standard inequality measures may then be evaluated using the resulting individual incomes.[8]

In Chen and Ravallion (2004) the separability of the macro and micro parts in the top-down approach allows the authors to refer only marginally to the underlying CGE model, estimated and discussed by other authors in a separate paper, when evaluating the effects of the entrance of China in the World Trade Organization (WTO). However, two problems are to be mentioned in connection with this kind of approach: (i) there is an issue of consistency between aggregate preferences in the CGE model and micro individual consumption preferences, and (ii) it is not clear whether the envelope theorem that applies only to marginal changes without any consumption constraint is valid in this context.

21.2.2 Bottom-up approach

The approach symmetric to the preceding one consists of starting from the bottom applying policy changes to the micro model of household behavior. Individual behavioral changes are then aggregated and fed into the macro model to analyze overall changes in the economy. Of course, the linkage variable in that case must be taken as being exogenous in the CGE model at the top (in the same way as the linkage variables in the top down approach were exogenous with respect to the micro model).

This approach is represented in Figure 21.2. The policy change is now denoted ΔZ and it is assumed to refer to the features of the tax-benefit system. The behavioral micro model applied to each household i in the database of households generates changes in the behavior of households, including changes in a variable of macro interest, Δx_i. Summing over all households yields an aggregate change ΔX, which is now the linkage variable between the micro and the macro part of the whole model. Then this change is fed into the macro CGE model as an exogenous shock, generating then changes in prices and in the structure of the economy, $\Delta \mathbf{p}$ and $\Delta \mathbf{Y}$. Presumably, these changes should feed back onto the behavior of households in the micro database but the bottom-up procedure stops short of that step.

8 On the relationship between inequality measures and social welfare functions, see Cowell (2000).

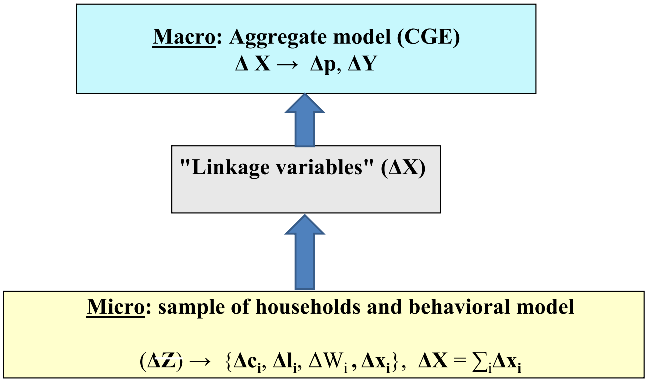

Figure 21.2 Micro-oriented policies: bottom-up micro—macro approach.

As a matter of fact, most studies of micro-oriented policies (ΔZ) stop once ΔX has been determined. Typically, the modeling of the micro effects of a reform in the tax-benefit system faced by households focuses on labor supply behavior. An econometric model is estimated that relates the labor supply of individuals or households to their budget constraint. Then the effect of a reform in the tax-benefit system on labor supply is obtained by simulating that model with the new budget constraint implied by the reform. The outcome of the whole analysis is the aggregate change in labor supply, ΔX in Figure 21.2. However, analysts seldom go beyond this and feed that change in the macro model, which should specified in such a way, of course, that aggregate labor supply is exogenous. This would be, but apparently without very much value added for policy makers, presumably because they can intuitively infer the impact of more or less labor supply on the level and structure of economic activity or unemployment. However, the overall budget cost, or benefit, of the tax-benefit reform and the ways to cover depend on these indirect effects the identification of which requires some kind of general equilibrium computation.[9]

An example of a study that fits the logical structure in Figure 21.2 is the evaluation of a diabetes prevention campaign in Australia by Brown *et al.* (2007). It involves three steps: (i) a micro model describing the effect of the campaign on diabetes prevalence within a sample of households, (ii) a conventional household labor supply model that identifies the influence of diabetes on employment and the likely impact of diabetes prevention on labor supply, and (iii) a CGE model that gives an estimate of the economy-wide effects of that change in labor supply, the cost of the prevention program and the saving in treatment costs in terms of economic aggregates such as GDP, employment, levels of consumption and investment.[10]

9 For a review of micro-based analyses of the effect of tax-benefit reforms on labor supply, see Blundell and MaCurdy (1999).

10 Actually the model is cast in a dynamic framework, but the logic remains the same as the one described here.

Maybe more than in the top-down approach, ignoring the feedback effects of the macro model on microeconomic behavior seems extremely restrictive here. With more labor supply because of the diabetes prevention campaign, it is to be expected that real wages will fall possibly inducing a reduction in the initial labor supply effect of the program. This feedback is in no way included in the micro model used in the first place, whereas it was present, although not in a fully consistent way in the first step of the top-down approach.

21.2.3 Recursive two-way linkage

Going a step further than in the top-down or in the bottom-up approach requires taking into accounts the feedbacks ignored in these approaches and evaluating their effects. This is equivalent to combining the two approaches and to go up and down iteratively in Figures 21.1 and 21.2 until no change is observed between two iterations.

This recursive approach requires some consistency between the macro and the micro parts of the whole model. In particular, it is important, as in the bottom-up approach that the linkage variable is exogenous in the macro model. For instance, extending the top-down approach illustrated in Figure 21.1 to make it recursive would require specifying the aggregate labor supply, and possibly the total household consumption, and in effect all the endogenous dimensions of household behavior in the micro model, as *exogenous* in the macro model. In other words, no representation of aggregate household behavior is needed in the CGE model, unlike with the top-down approach. More fundamentally, the aggregate CGE model should not include any endogenous representation of aggregate household behavior since this would essentially neutralize the micro part of the whole macro—micro framework.

This is illustrated in Figure 21.3. In the macro model, there are then two types of exogenous variables. Those variables that represent the macroeconomic context, including policy variables (ΔX) and those variables that result from the aggregation of

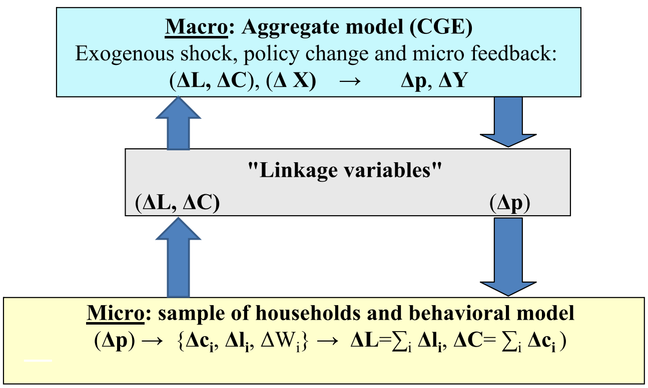

Figure 21.3 Distribution effects of macro policies: recursive two-way approach.

individual behavior in the micro model applied to the household database. In Figure 21.3 the latter essentially are the changes in aggregate labor supply and in aggregate consumption expenditures **ΔL** and **ΔC**. Of course, all types of endogenous outcomes from the micro model could be used, provided they are exogenous and play some role in the aggregate CGE model.

If this recursive procedure converges – convergence being defined here as no difference in endogenous variables between two complete top-down/bottom-up cycles – then this approach is clearly equivalent to a *fully integrated macro–micro model*. In such a CGE model, the household sector does not consist of a few aggregate representative households but actually of a full sample representative of the whole population. Convergence of this kind of fixed point algorithm is not granted, however. The Jacobian of the whole system must satisfy some properties for convergence to hold. Models may be sufficiently complex for these properties to be difficult to check.

This recursive approach has been used extensively. One of the first macro–micro models, DYNASIM, was solved using this kind of approach within a sequential dynamic framework. Typical recent applications include for instance Aaberge *et al.* (2004), Arntz *et al.* (2008), and Mussard and Savard (2010).

21.2.4 Simultaneous or fully integrated approach

Progress in computing power permits being more direct and avoiding the preceding problem of convergence. It is sufficient to include in a standard CGE with representative households framework as many 'representative' households as there are households in the database. Cockburn (2006) and Rutherford and Tarr (2008) have done it respectively with 3000 and 50 000 households apparently without major problems. Whether through the recursive method or directly through powerful model-solving software, handling the macro and the micro part of macro–micro models of income distribution within a fully Walrasian context is therefore feasible, even though this may often be a rather heavy operation.

Calibration, in particular, must now be handled both at the macro level, as usually done in CGE modeling, and at the micro level. This is something easily done when dealing with a few representative household groups under the assumption of perfect aggregation within each group. It requires more care when dealing with the full diversity of households.

Consider for instance the case of consumption behavior when the consumption function is simply specified as the standard Stone–Geary model:

$$p_k c_{ik} = p_k C_k(\boldsymbol{X_i}) + \alpha_k(\boldsymbol{X_i})\left[y_i - \sum_j p_j C_j(\boldsymbol{X_i})\right], \tag{21.1}$$

where c_{ik} is the consumption of good k by household i, $\boldsymbol{X_i}$ its observed sociodemographic characteristics and y_i its income, which should itself be endogenous. α_k is the propensity to consume good k and C_k the minimum consumption of good k. Both are assumed to depend on household characteristics. These functions can be estimated econometrically on a cross-section of household budget data through straight linear ordinary least squares (OLS) on a cross-section of household expenditures (e_{ik}), assuming that all of them face the same set of prices, p_k, set to unity without loss of generality. The residuals $\hat{u}_{ik}$ of these regressions are given by:

$$e_{ik} = p_k \hat{C}_k(\boldsymbol{X_i}) + \hat{\alpha}_k(\boldsymbol{X_i}) \left[y_i - \sum_j p_j \hat{C}_j(\boldsymbol{X_i}) \right] + \hat{u}_{ik} \tag{21.2}$$

When simulating model (21.2) with changes in prices and in household incomes, it may matter whether the residual term is homoscedastic or heteroscedastic, especially if it depends on endogenous variables like y_i. In that case, a change in the distribution of income, due for instance to changes in the structure of wages, affects the aggregate demand not only through the observed heterogeneity with respect to the characteristics X but also through the unobserved heterogeneity in the $\hat{u}_{ik}$, itself dependent on the distribution of income. Things are even more serious when the coefficients of the Stone–Geary model are assumed to be heterogeneous, on top of observed characteristics X. There is indeed no way this unobserved heterogeneity can be explicitly dealt with.

As it is already the case with straight CGE modeling, practically all theoretical specifications can be used when extending the model to take into account the heterogeneity of households and the distribution of income among them. The resulting complexity of the micro–macro framework does not really raise conceptual or computational problems. The difficulty lies at the calibration stage as there is an infinity of behavioral parameters consistent with the data when taking into account the behavioral heterogeneity of households. Panel data would be needed to identify this heterogeneity which is somewhat contradictory with static micro–macro modeling.

21.3 STATIC MACRO–MICRO DISTRIBUTIONAL MODELS WITH REAL OR APPARENT LABOR MARKET IMPERFECTIONS

An important dimension in the analysis of income distribution is the allocation of individuals or households in specific groups, locations or economic sectors where remunerations or prices are different. Somehow this supposes that some market imperfections are present that prevent wages to equalize across sectors. These imperfections may be apparent or real. Production sectors may pay different levels of wages for observationally identical individuals because they employ individuals with heterogeneous productivities in perfectly competitive markets or because workers in some sectors have been able to secure some non-competitive advantage. Employment sectors thus are important determinants, on top of individual endowments, of the distribution of income.

In such a context, the linkage variables that link the macro and the micro parts of the macro–micro model should include not only the price system, but also some representation of the way individual workers are allocated across sectors. In some instances, this requires also employment levels to be passed on from the macro to the micro part of the model.

As an illustration, imagine the CGE model in the macro–micro framework allows for changes in sectoral levels of employment with fixed relative wage across sectors of production. Then the simulation of a reform in the trade policy at full employment leads to a change in the weight of the various sectors of the economy as well as, possibly, in the general level of wages, and therefore to a change in the distribution of earnings and income. It is this change the analyst is interested in. This case is analyzed below before considering other types of market imperfections and the effect of factor reallocation on the distribution of household income rather than individual earnings.

In an economy comprising two sectors A and B, the structure of relative aggregate wages across sectors is fixed, but the absolute level of real wages is flexible. Assume that a macro policy contributes to increasing the value-added price of sector A relatively to that of sector B. Labor is thus reallocated from B to A. Assuming that earnings are distributed in a different way in the two sectors, something that is observed in the micro database, how will the distribution of earnings will be modified due to that sectoral shift of employment?

There are various ways of addressing that issue. The simplest one consists of reweighing the two sectors in the original sample of earners in the population. However, more elaborate approaches can be used with different assumptions about the sectoral allocation of individual workers across sectors.

21.3.1 Random changes or reweighting

A simple way of implementing this employment shift at the micro level is to withdraw randomly earners employed in sector B and to add randomly an equal number of workers in sector A. At the same time, all wages must be scaled up or down according to the change in the general level of wages in the CGE model. The change in the overall distribution of relative earnings thus results from the change in the weights of the two different distributions in sectors A and B in the whole population of earners.

This method is equivalent to the reweighting procedure used when updating population samples to fit changes in the geographical structure of the population. If it is realized that location A is undersampled and location B oversampled, a simple adjustment in sampling weights will ensure that the sample is fully representative of the population. Presumably, the same can be done with sectors of employment.

In both cases, the real issue is whether those people who move from one sector to another, or who were missed in the initial sampling phase are randomly distributed in their sector or location of origin. This is quite doubtful and requires going beyond this simple micro adjustment procedure.

Before doing so, this might be the place to mention an issue that was ignored in the Section 21.2, that of the homogeneity of the price changes predicted by the CGE model. In Figures 21.1 and 21.3, it is assumed that all households face the same change in prices, including wages. However, it is conceivable that they actually face different changes, most notably on the labor market. In other words, the percentage change predicted in the aggregate wage level of a given type of labor in the CGE model does not necessarily trigger the same proportional change in all individual wages in that category of labor in the micro database. The CGE aggregate change may be met at the micro level by a whole variety of individual changes, with of course different effects on the distribution of earnings. Individual dynamics of this type were ignored in the basic macro–micro framework, in the same way as it is ignored in the simple reweighting approach to employment changes.[11]

21.3.2 Rational individual choice of employment sector: the Roy model

One way of rationalizing different levels of wages for labor of a given type in different sectors of the economy is to assume different tasks in the various sectors of the economy for a given type of labor and workers with heterogeneous productivities in the various tasks self-selecting rationally and competitively in the various sectors. This is sufficient to generate average wage differentials and distinct earnings distributions across sectors. This basic model named after Roy (1951) is very helpful in analyzing earnings distributional issues (see Neal and Rosen, 2000).

Formally, the Roy model can be specified as follows. Say that sector A employ unskilled workers for task a, whereas sector B employs unskilled workers for task b. Assume that there is some heterogeneity in the productivity of the workers in the two tasks so that a worker i is paid:

$$\ln w_{ai} = \ln \pi_a + \ln \mu_a + u_{ai}$$

$$\ln w_{bi} = \ln \pi_b + \ln \mu_b + u_{bi}, \tag{21.3}$$

in sectors A and B, respectively, where (π_a,π_b) is the price of a unit of effective labor in tasks a and b, respectively, (μ_a,μ_b) is the average productivities in the two tasks within the whole population, and (u_{ai},u_{bi}) stands for the heterogeneity of individual productivities.

[11] Reweighting procedures of the type just described are rather trivial. They are sometimes combined with other criteria to correct sample biases. For instance, the under-reporting of some type of income may be compensated by increasing the relative weight of demographic groups with more of that income. For general procedures doing so, see Creedy and Tuckwell (2004) or Robilliard and Robinson (2003). This kind of reweighting is of less relevance in the present framework.

These two variables are assumed to be distributed with zero expected value in the population.

The supply of workers in sectors A and B will depend on the price difference $\pi_a - \pi_b$, on the difference in individual productivities in the two tasks, $\mu_. + u_{.i}$, and possibly on individual or household characteristics. In the simple case where only wages matter in the employment sector decision, a worker will chose sector A iff:

$$u_{ai} - u_{bi} > \ln\mu_b\pi_b - \ln\mu_a\pi_a \tag{21.4}$$

This sector allocation rule replaces the random allocation process and the reweighting procedure discussed in the preceding section. The proportion of workers supplying labor in sector A is given by:

$$\Pr\{u_{ai} - u_{bi} > \ln\mu_b\pi_b - \ln\mu_a\pi_a\} = L_a(\mu_b\pi_b/\mu_a\pi_a),$$

where the function L_a depends on the distribution of the two individual productivity terms.[12] The average (log) productivity of these workers is given by:

$$E\{\ln\mu_a + u_{ai}|u_{ai} - u_{ai} > \ln\mu_b\pi_b - \ln\mu_a\pi_a\} = \ln P_a(\mu_a, \mu_b\pi_a)$$

The same definitions can be applied to task b and sector B. What is essential here is the fact that average wages, $W_a = P_a\pi_a$, $W_b = P_b\pi_b$, differ across sectors and also differ from what would be observed without self-selection ($\mu_a\pi_a$ and $\mu_b\pi_b$). The supply of effective labor for task a in sector A is thus given by:

$$S_a = N \cdot L_a\left(\frac{\mu_b\pi_b}{\mu_a\pi_a}\right) \cdot P_a(\mu_a, \mu_b\pi_b/\mu_a\pi_a) \tag{21.5}$$

where N is the size of the population.

To close the model, it is sufficient to introduce the demand for effective labor in task a in sectors A and for task b in sector B in the macro model. Given the supply of the two tasks given by (21.5), the macro model then returns the competitive price to be paid for each unit of task a and b (π_a and π_b) which can in turn be plugged into the micro-simulation and the self-selection rule (21.3). The solution is obtained recursively as in Figure 21.3, the main difference being that labor supply is now sector (or task) specific.

The calibration of the micro model is straightforward. Earners are observed to be in sector A or B, and therefore either their wage in A or their wage in B is observed. Model (21.3) can be estimated with some additional self-selection equation that relies on the two regimes in (21.3) and possibly on some additional variables reflecting the

12 In a more general specification, the utility of each job may depend on some observed (X_i) and unobserved (ν_i) sociodemographic characteristics of worker i, on top of earnings. The probability for worker i to chose sector A would then be given by: $\Pr\{u_{ai} - u_{bi} + X_i\beta + \nu_i > \pi_b - \pi_a\} = L_a(\pi_b - \pi_a + X_i\beta)$, where β is a set of parameters to be estimated.

socioeconomic determinants of individual productivities.[13] Sticking to the first case, the econometric model writes:

$$\ln w_i = \ln(\pi_a \mu_a) + u_{ai} \text{ if } \ln(\pi_a \mu_a) + u_{ai} \geq \ln(\pi_b \mu_b) + u_{bi}$$

$$\ln w_i = \ln(\pi_b \mu_b) + u_{bi} \text{ if } \ln(\pi_a \mu_a) + u_{ai} \geq \ln(\pi_b \mu_b) + u_{bi}, \qquad (21.6)$$

with some normality assumptions on the probability distribution of the individual productivities. This very standard regression switching econometric model yields estimates of $\pi_a \mu_a$ and $\pi_b \mu_b$ and the variances of the *u*s. The correlation between the *u*s must be taken as arbitrary in this simple model whereas μ_a and μ_b are obtained through normalizing the prices π_a and π_b to unity.

Residuals *u*s are observed for the regime an earner has actually chosen, whereas the residual in the other regime is generated as a pseudo-residual drawn randomly in the distribution of that individual productivity under the self selection condition (21.4). Extended models and richer data permit us to estimate the covariance between the two productivity terms and to have a better estimate of the unobserved residual.

Although this would seem to be a rather attractive way of introducing the distribution of earnings within a full general equilibrium framework relying on macro-level wage differentials, there are few applications of this model at this stage. Cogneau and Robilliard (2008) have pioneered this approach but, until now, had few followers.

There are various reasons why this approach to the modeling of earnings and income inequality within a micro—macro framework is burdensome. First, the straight Roy model has been shown not to work very well by Heckman and various coauthors within a two-sector context. In particular, Heckman and Sedlacek (1990) show that there is a lack of coherence between the Roy model and aggregate labor demand in US manufacturing and non-manufacturing sectors. Although they propose an extended model that works better, a second difficulty is that generalizing these models to more than two sectors quickly becomes prohibitively heavy from a computational point of view. Third, the microsimulation stage described above makes it difficult to use standard software to solve the model in an integrated way and it is not completely clear *a priori* why a recursive procedure of the type depicted in Figure 21.3 would converge. Fourth, it is rather restrictive to assume that individual productivities are independent across sectors as done above. Doing otherwise, however, requires a richer dataset than readily available cross-sections of earnings and sectors of employment.

Heckman and various coauthors have been able to overcome the previous difficulty by using data with some panel dimension or by assuming that individual productivities in various tasks were correlated with some observed individual attribute like cognitive and

13 See Heckman and Sedlacek (1985) and Heckman and Honore (1990) for the estimation of the straight and extended Roy model.

non-cognitive skills. Some of these models were then integrated into a general equilibrium framework, but of a rather different nature from the kind of static multisectoral model considered up to now. More will be said on this macro—micro link below.[14]

At this stage, it will only be concluded that the Roy model seems *a priori* a nice way to bridge macro and micro to deal with distributional issues while taking into account a major macro determinant of inequality, which is the difference in average wages for different types of labor across sectors of economic activity. It is to be hoped that more attempts will be made in that direction to get a better sense of the usefulness of that approach.

21.3.3 Modeling rationing and household income generation on the labor market

There are two important limitations in the preceding approach to integrating intersectoral wage differentials into a distribution oriented micro—macro framework. The first is that this is done within a perfect competition framework where all individuals chose the type of work and the sector they prefer. Instead, one could think that the labor market is imperfect, and supply and demand of the various types of labor in different tasks do not adjust through wage flexibility. If this is the case, the micro—macro link would require to model wage, and possibly other price rigidities at the macro level and the way individuals are rationed in the labor market on the micro side, working in the sector they do not prefer at ruling wages or being unemployed.

A second limitation of the Roy model is that it actually applies to individual earners and thus provides an interesting way of describing the determinants of the distribution of individual earnings. However, analysts may ultimately be more interested by the distribution of income among households and consumption units. At that level, important choices include not only the sector of employment — including unemployment — but also the labor force participation of the various household members.

An approach that partly remedies these limitations consists of introducing a partial equilibrium household income generation model on the micro side of the whole macro—micro model. By partial equilibrium it is meant that labor supply or sector employment choice of household members do not explicitly depend on the endogenous prices and wage generated by the macro part of the model. They depend essentially on family and individual characteristics, and on endogenous sectoral employment levels given by the CGE model.

The representative work of this approach to macro — micro modeling is the analysis by Bourguignon *et al.* (2004) and Robilliard *et al.* (2008) of the distributional impact of the Indonesian crisis and potential effect of alternative policies to cope with the

[14] See Heckman and Vytlacil (2008).

exogenous shocks. As both unemployment and the shift from formal to informal employment were important consequences of the crisis, it seemed necessary to model the labor market with some rigidities that would generate excess supply of workers for formal jobs and possibly unemployment. At the micro level, this required introducing some rationing scheme among labor force participants and possibly allowing some of them to become inactive. A second difficulty was to take into account the correlation of employment status among the members of the same household.

A simplified representation of the way this was done is as follows. Consider first, to simplify, the case of individual labor force participants who apply for a job in the formal sector of the economy, but assume that because of some wage rigidity there is excess supply of workers. Also assume that the employment decision is discrete, only full time jobs are available. Let L_i $(=0,\ 1)$ be the notional labor supply of individual i in the formal sector and a_i $(=0,\ 1)$ the rationing scheme that applies to him/her in case he/she applies for a job. His/her employment status e_i $(=0,\ 1)$ is then given by:

$$\begin{aligned} &e_i = 1 \text{ (employed in the formal sector) if } L_{\mathrm{i}} = 1 \text{ and } a_i = 1 \\ &e_i = 0 \text{ (other employment or inactive) if } L_{\mathrm{i}} = 0 \text{ or } a_i = 0. \end{aligned} \tag{21.7}$$

At the macro level, assume then that the CGE model with some type of wage rigidity provides a value for total employment, E. In the base run there is full adequacy between the number of individuals observed in the micro database as being employed in the formal sector and the simulation of the CGE model. Consider, however, a shock leading to a simulated change ΔE in total formal employment. The top-down macro–micro structure in Figure 21.1 can be used but now aggregate sectoral employment levels are part of the linkage variables. The problem then is that of adjusting the number of people with a formal job in the micro database. Whom should get a job if ΔE is positive, whom should lose his/her job if ΔE is negative?

One possible answer is the reweighting procedure discussed above, which is equivalent to considering that movers are taken randomly from the population they come from and allocated randomly to the employment group they go to. This is clearly unsatisfactory. Whether through notional labor supply or rationing, such a change in aggregate employment is necessarily selective. People at the 'border' of the implicit labor supply or employment criterion will be the movers and they are not randomly distributed in the population.

A rigorous treatment of that issue would require building a model of the labor market with matching of workers and employers. This is clearly a major step forward that is hard to take, in particular because of heavy data requirements.[15] A more pragmatic solution

[15] Note, however, that work along the lines of Abowd and Kramarz (1999) goes in that direction. What is missing is the full link to a macro model.

consists of simply modeling the notional labor supply and the employment decision in (21.7) on the basis of observed variables.

Namely, let the probability of individual i being employed in the formal sector be:

$$\begin{aligned}\Pr\{e_i = 1\} &= \Pr\{L_i = 1 \text{ and } a_i = 1\} \\ &= \Pr\{\boldsymbol{A_i}\cdot\beta + w_i\gamma + \boldsymbol{Z_i}\delta + \eta_i \geq 0 \text{ and } \boldsymbol{z_i}\delta' + \epsilon_i \geq 0\}.\end{aligned} \tag{21.8}$$

In this specification, the variables A_i stand for the income of alternative occupations for individual i (informal work, farming, etc.), w_i for his/her wage in the formal sector, $\boldsymbol{Z_i}$ for individual *and* family characteristics, including the income of other family members, and $\boldsymbol{z_i}$ for pure individual characteristics like education. η_i and ϵ_i stand for unobserved characteristics, and are assumed to be distributed as independent normal variables with zero mean.[16] The first criterion in the probability expression in (21.8) corresponds to standard labor supply behavior, whereas the second describes the rationing behavior that depends solely on individual characteristics.

The parameters of the model may be estimated from a standard micro dataset. However, this requires some precaution since some of the variables entering the model are not necessarily observed. This is the case in particular of the formal wage rate of those individuals not employed in the formal sector or alternative occupations' income for those in that sector. On the other hand, the family dimension of the model requires taking into account some possible correlation of unobserved determinants η_i and ϵ_i across family members.

Assuming this model can be estimated satisfactorily, then adjusting overall employment to the simulated change ΔE can be done in a simple heuristic way. Say that total employment must be reduced. Then the most likely movers are those who are observed in formal employment but with the lowest *a priori* probability (21.8) of being employed in that sector. This is equivalent to assume that there is a waiting queue to enter the formal sector and candidates are ranked according to some criterion. A contraction of employment in that sector then would mean dropping off those employed in the formal sector with the lowest criterion value, whereas an expansion would mean taking on board those outside the formal sector with the highest criterion value.

Formally, the score used to define the queue of people waiting to get employed in the formal sector may be based on the second condition in the expression of probability (21.8). Let it be:

$$S_i = \boldsymbol{z_i}\cdot\hat{\delta}' + \hat{\epsilon}_i,$$

[16] A model of this type is used by Abowd and Farber (1982). It was first introduced in the econometric literature by Poirier (1980). Of course, different specifications than the one used here can be used to describe the functioning of the labor market in presence of rigidities and rationing.

where $\hat{\epsilon}_i$ is a pseudo residual drawn in the distribution assumed for estimating the parameters in (21.8) consistently with the observed employment status. Then reallocating individuals in or out of the formal sector so as to accommodate the simulated aggregate change in employment in the macro model ΔE can be done simply by incrementing all the residuals in the same way until the number of people passing the double criterion in (21.8) be the desired one.

It can be seen that this procedure of reallocating individuals across sectors is in effect of the same nature as the reweighting procedure described above. There is a big difference, though. It is that now this reweighting is *selective*, or, in some sense, conditional on observed individual and household characteristics. This is bound to have strong implications in terms of income distribution.

The simple model above can be generalized in various directions. (i) The choice of activity may involve more than two alternatives as, for instance, with formal, informal and inactivity or more elaborate choices. This can be dealt with discrete choice models among many activities using multilogit rather than probit specifications. (ii) It is important to take into account the interdependency between the members of the same household. It is unlikely, for instance, that a household member will chose inactivity if the head of the household is him or herself inactive or unemployed. Likewise, the sectoral choices may not be as strict in such circumstances. Estimating this kind of occupational choice models simultaneously for various members of a household would clearly be cumbersome. The most common practice is to estimate that behavior first for the household head, conditionally on the characteristics of all family members and then for other household members conditionally on the decision, and the corresponding income of the head. (iii) A further point to take into account is that any occupational choice model relies on unobserved returns to alternative activities for which it is necessary to get estimates. This means that the estimation of model of type (8) must take place jointly with the estimation of models that describe how earnings in specific activities depend on individual characteristics while taking into account of selection biases.

Putting all these components together leads to a true household income generation model that describes the way in which the distribution of income depends on basic aggregate variables: the prices as in standard full equilibrium models, but also employment levels in case of market rigidities. Examples of this kind of model are now common. A first generation of such models is described in a thorough way in Bourguignon *et al.* (2003), which also gives various examples of application in Latin America and Asia.

Top-down links with imperfectly competitive CGE models were first introduced by Bourguignon *et al.* (2004) and Robilliard *et al.* (2008) in analyzing the effects of the Indonesian crisis on the distribution of income. This was certainly a case where the assumption of market imperfections and rigidities was justified. To illustrate the power of the macro—micro method just described, Figure 21.4 compares its results to the

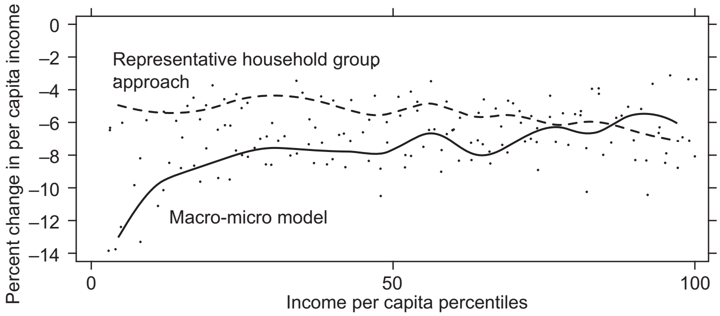

Figure 21.4 Simulated distributional impact of the Asian crises in Indonesia (Robilliard *et al.*, 2008).

simulated incidence of the crisis when using the standard representative households approach in the CGE modeling and assuming constant distribution within each corresponding group of households in the micro database. The difference is staggering. Of course, the change produced by the macro—micro model is much closer to what was actually observed.[17]

Various authors have applied this approach since the mid 2000s. They include, for instance, Boccanfuso and Savard (2007) on Mali, Ahmed and O'Donoghue (2010) on Pakistan, Bussolo and Lay (2006, 2008) for Brazil and Colombia, and Go *et al.* (2010) on South Africa. There are fewer applications in developed countries, where in effect economy-wide CGE models are found to be less adequate because of lesser structural changes and possibly more rigidity in those economies. However, an application of the present macro—micro framework to the redistribution system in Germany is provided by Peichl and Schaefer (2009), and in Switzerland by Muller (2004). It is also worth mentioning that the top-down approach can be extended so as to fully integrate the macro and micro parts of the model. This was done with a recursive algorithm as in figure 3 in Bourguignon and Savard (2008).

As mentioned above, the way employment allocation is done at the micro level in this recent literature is somewhat *ad hoc.* An attempt at justifying (21.8) on theoretical grounds was made in Bourguignon *et al.* (2004), but was not fully convincing. If there is agreement on the fact that economies are frequently out of equilibrium and workers are not necessarily in the occupation they would prefer, then the appropriate labor market models should be used, namely those models that combine search and sorting, possibly with different types of labor and tasks, as in be calibrated using panel data on employment and job mobility. This goes much beyond the scope of the models reviewed here.

[17] For a more rigorous comparison of the simulation of the distributional impact of a crisis and the actual impact, see Ferreira *et al.* (2008) using a similar approach in the case of Brazil.

21.3.4 Final reflections on static macro—micro modeling

Three remarks are in order in concluding this review of static macro—micro models of income distribution. First, the difference with the representative household approach that consists of introducing various representative households in the CGE model and assuming constant distribution within these household groups in the micro database must be emphasized. As long as there are enough dimensions in the bridge between macro and micro (the linkage variables), microsimulation shows much more variation across individuals or households in the micro impact of macro changes — and of course of micro-oriented policies — than the representative household approach.

It would be a mistake, however, to consider that this kind of microsimulation gives information on the way specific households or individuals would actually be affected by the policy being simulated. Microsimulation, whether arithmetic, behavioral, linked to a macro model or not, gives information on the way a cross-section of households is affected by a policy reform, without really being possible to say that simulated changes apply to specific households in the sample used for the computation. In other words, earners and households in the sample are in effect themselves 'representative' earners or households and the simulation says something about the change in the behavior or the welfare of those 'representative' agents. Things would be different if the model were dynamic and relied on panel data.

The last remark has to do with the time dimension of the macro—micro modeling of policy reform. The approach discussed here is essentially an *ex ante* micro evaluation of the reform being simulated. If micro datasets are available at regular time intervals, then it becomes possible to use before/after comparisons and to partly validate the whole approach. Discrepancies between what is observed *ex post* and what is simulated *ex ante* may be due to various causes: (i) the CGE model or the modeling in the micro part are unsatisfactory or their implicit time horizon differs from the time gap between simulation and *ex post* observation, (ii) other observed important changes took place that must be included in the initial simulation for the comparison to be meaningful and (iii) changes independent from the reform are difficult to observe or cannot be modeled. Some of these discrepancies can be evaluated with a thorough treatment of *ex ante* and *ex post* data. Not enough of this reality check is generally done in applied work.[18]

21.4 DYNAMIC MACRO—MICRO MODELING

As the latter remarks suggested, merging macro and micro approaches is especially difficult when the focus of the analysis is economic growth. The main problematic assumptions are that, on the one hand, macro dynamics are somehow independent from

[18] Such an approach is followed in Ferreira *et al.* (2008).

distribution or heterogeneity at the micro level and, on the other hand, inter-temporal decisions of individual agents are seldom 'aggregated' at the macro level. In other words, when it comes to growth macro—micro feedbacks become extremely difficult to model.

For example, the recent microeconomic literature on inequality of opportunities (see Roemer, 1998; Roemer *et al.*, 2003; Bourguignon *et al.*, 2007) has shown that unequal initial distribution or unequal access to education, health and other human development factors leads to inefficient inequality traps. In such a situation, investment opportunities are missed and the final result is a lower rate of growth. However, this lower growth would not be predicted by a standard macro model that ignores these micro distributional issues.

Another example can be found in the macroeconomic literature on endogenous growth. This literature has highlighted the importance of accumulation of knowledge, via private or public spending on education and/or R&D. However, most empirical models of endogenous growth do not really consider how accumulation of knowledge really happens at the micro level. Or which are the incentives that work and push agents to invest more in knowledge. Once again the macro—micro feedbacks are absent.

In the reminder of this chapter, we focus on some simpler preliminary attempts to link macro and micro approaches to study the dynamics of economic systems. This section briefly reviews the vast literature on dynamic microsimulation.[19] This literature originated from an endeavor to combine macro and micro, but most of the recent contributions have a microeconomic focus. An illustrative example of how these micro techniques can be combined with a macro (CGE) dynamic model will be presented in Section 21.5.

21.4.1 Introduction to dynamic microsimulation: definition and most common uses

As mentioned in the introduction, the seminal work on dynamic microsimulation was by Orcutt and dates back to the end of the 1950s. Orcutt's own words clearly describe the motivation and value added of dynamic microsimulations:

> *Existing models of our socio-economic system have proved to be of rather limited predictive usefulness. This is particularly true with respect to predictions about the* effects of alternative governmental actions *and with respect to any predictions of a* long-range *character. [… Besides] current models of our socio-economic system only predict aggregates and fail to predict* distributions *of individuals, households, or firms in single or multi-variate classifications.*[20]

19 A few reviews of dynamic microsimulation are already available (see, e.g. Harding, 1996; Mitton *et al.*, 2000; Zaidi and Rake, 2001; Bourguignon and Spadaro, 2006; Harding and Gupta, 2007; O'Donoghue, 2008). The objective of this section is not to provide an additional exhaustive review of the relevant literature, but actually to present a brief guided tour of the most relevant aspects of this literature. A more specific goal is to highlight the challenge and usefulness of linking microsimulation and general equilibrium methods.

20 From Orcutt (1957, p. 116), emphasis added.

Since his original work on DYNASIM, significant progress, both in terms of modeling and data availability, has been achieved. However, it is still true that dynamic microsimulation models are the main tool to address long-term distributional issues related with changes of government policies or other shocks. It is also still true that the unique contribution of a dynamic microsimulation model is that it projects samples of the population forward in time — simulating life paths of individuals.

A simplified microsimulation framework can be formalized as follows. Household *per capita* income (or its version accounting for economies of scale), y, is normally used as the indicator of household welfare; any economic policy is thus assessed in terms of its impact on this indicator. The distribution D of income y at time t can be expressed as the product of the joint distribution of all relevant household or individual characteristics $\boldsymbol{X}$ and the distribution of income conditional on these characteristics:

$$f_t(y) = D_t = \int \cdots \int_{c(x)} g_t(y|\boldsymbol{X}) v_t(\boldsymbol{X}) \mathrm{d}x, \tag{21.9}$$

where $f_t(y)$ is the density function of the distribution of income and the summation is over the domain $C(\boldsymbol{X})$ on which $\boldsymbol{X}$ is defined. Define an income generation model describing household *per capita* income (y) at time t as a function of household members' characteristics or endowments ($\boldsymbol{X}$), the market rewards for those characteristics ($\boldsymbol{\beta}$) and unobservable components (ϵ):

$$y_t = F(\boldsymbol{X_t}, \boldsymbol{\beta_t}, \epsilon_t) \tag{21.10}$$

Knowing the distribution of ϵ_t, it is then possible to get the conditional distribution $g_t(y|\boldsymbol{X})$ which depends itself on $\boldsymbol{\beta_t}$. Then the distribution of income is determined through (21.9).

This simplified framework is useful for illustrative purposes, but a few extensions are necessary to use it in actual empirical analyses. First of all, equation (21.10) is a *reduced form* of a generic microsimulation model. In most applications, this equation is replaced by a set of structural equations. A common setup includes at least two equations — one for the labor market participation choice and one for the determination of earnings. In such a case, the β of equation (21.10) would then represent the whole array of different parameters influencing both participation choices and earnings, including the tax-benefit system for instance. Another important extension is that households' welfare is clearly also influenced by the prices of consumption goods (or inputs in production activities run by self-employed workers, e.g. small farmers). To take into account the consumption side, equation (21.10) should thus include these prices as done in the static framework (see Figure 21.3).

A crucial element in dynamic microsimulations is to determine how the distribution D changes over time. In this simplified version it is possible to distinguish two sources of dynamic change. The first one is given by the changes in the parameter $\boldsymbol{\beta}$, i.e. the market rewards to the characteristics $\boldsymbol{X}$, other parameters such as those affecting labor participation decision and prices. An example may clarify this. Workers with higher education levels may start earning higher wages (β_{skill} goes up) because of a technological change that affects the skill premium or a change in direct taxation affects the decision to participate in the labor market for some member of the household. In both cases D will be affected.

The second source of dynamic shifts in the income distribution D is represented by changes of the distribution of individual and household characteristics ($\boldsymbol{X}$). In other words, alterations of the structure of the population in terms of age, education, households' size and composition shape the distribution of income of that population. 'Life events' such as marriage and household formation, fertility, mortality, migration and investment in education are the ultimate factors involved in the change of the structure of the population.

Clearly these two sources of dynamic change are not independent one from the other and, in the real world, they are simultaneously determined. For example, decisions to participate in the labor market may be affected by fertility choices or, *vice versa*, investments in education may depend on expected higher future wages for skilled jobs. The problem of estimating and running a fully simultaneous microsimulation framework is overcome by making some simplifying assumptions (more on this below).

In more formal terms, the dynamic change of the $\boldsymbol{X}$ can be represented as follows:

$$\boldsymbol{X_{t+1}} = \boldsymbol{\phi}\left(\gamma_t, \boldsymbol{X_t}, \hat{\beta}_{t+k}, \hat{\epsilon}_{t+k,}\right), \tag{21.11}$$

where ϕ is a vector and the hat notation (^) represents expectations in future periods. Equation (21.11) is a compact and useful formalization because, with some small variation, it can encompass the wide range of dynamic microsimulation models found in the literature.

As the few examples in the section below will illustrate in more detail, function ϕ can be thought of as: (i) a trivial arithmetic addition of one year if the age of individuals is the only variable considered; (ii) a transition probability matrix, with transitions from one status to another such as being alive, employed, married, having a child, etc.; and (iii) a survival function measuring the duration of time before a life event (or change of status) happens. In general, however, $\boldsymbol{X}$ is a vector and $\boldsymbol{\phi}$ may be much more complex.

Note also that equation (21.11) explicitly states that the *next* or *future* value of $\boldsymbol{X}$ depends exclusively on the current value of $\boldsymbol{X}$ and future expectations for the other arguments. Given that the future value does not depend on past values except for the present one, the process represented by equation (21.11) is Markovian. However, this may not be a satisfactory assumption for some dynamic micro processes and more complex representations that include 'memory' of past states may need to be introduced. This would be the case, for instance, for pension rights that depend on the whole sequence of past earnings. Finally, although the notation in equation (21.11) is for a discrete time model, it is easy to transform it to denote continuous time; in fact, most survival functions operate in continuous time.

To draw a parallel with the static microsimulation models of the previous sections, it is possible to classify dynamic microsimulations into two groups: arithmetic (or probabilistic) and behavioral. The classification will depend on the type of the variables $\boldsymbol{X}$ under consideration in equation (21.11) and on the choices of the modeler. Some of the $\boldsymbol{X}$ are by their nature 'rigid' over time (e.g. age), while others are 'flexible' and may imply behavioral decisions at some time period t or, still, the dynamics of some other variable may be due to a mix of behavioral choices and random elements.

Including heterogeneity at the individual level, modeling labor supply and consumption behavior at the micro level, as well as modeling long-term changes of the socioeconomic structure of a population by following individuals throughout their life paths are all important advantages of dynamic microsimulation models. Owing to these advantages, these models are the only modeling tool that can be used to assess changes in income distribution related with: (i) different demographic scenarios and, in particular, aging; (ii) changes of long-term government programs such as pensions, health and long-term care, social insurance, financing of education, etc.; and (iii) inter-temporal processes and behavioral issues such as wealth accumulation, nuptiality and fertility, labor market mobility and intergenerational transfers. Given their considerable resource requirements, most large dynamic microsimulation models are built and maintained in developed countries,[21] but some examples are also available for developing countries.[22]

21.4.2 Methodological issues of dynamic microsimulations

Comprehensive dynamic microsimulation models aiming at projecting the characteristics of a population have to simulate many decisions at the micro level and therefore present numerous challenges. The most frequent methodological challenges faced by researchers include, amongst other, choices of behavioral versus stochastic specification of micro processes of type (21.11), discrete versus continuous time, cohort versus population model. This section considers some of these methodological challenges, whereas Section 21.4.3 will focus more on data quality and estimation issues.

The first and most commonly faced methodological challenge in dynamic microsimulation consists of *aging* the population. The term aging indicates not only the age of the individuals, but also other characteristics, such as their education, marital status,

21 The literature reviews mentioned in footnote 19 report on many examples of large-scale dynamic microsimulation models for Organization for Economic Cooperation and Development (OECD) countries.

22 See Cogneau and Grimm (2008) who study the dynamic effects of the AIDS epidemic on income distribution and poverty in Cote d'Ivoire or Ferreira and Leite (2003) who use microsimulation methods to investigate Millennium Development Goals achievement in Brazil.

location, etc. As discussed before, in the so-called *static* aging approach, the micro data at a point of time is reweighted so that the relevant characteristics of the distribution (the Xs in equations 21.9–21.11 above) are changed to match some external data source. For example, a demographic projection provides the age distribution of a certain population for a given future period; the initial micro dataset is reweighted in such a way as to reproduce this given age distribution.[23] Other commonly targeted characteristics include education achievement, occupation, marital status, affiliation to certain government programs (pensions, conditional cash transfers, etc.).

There are advantages and disadvantages to static ageing. This approach can exactly target demographic (or other) projections when they are available, and thus avoid the risk of generating unrealistic projections associated with more sophisticated but potentially mis-specified aging approaches. However, static aging may quickly become difficult to apply when the number of characteristics to project increases. It may be the case, for example, that one needs to age a population taking jointly into account age, education, area of residence and fertility. If the age groups are very disaggregated (less than 5 years per group) it may happen that the original sample has too few observations for certain combinations of the mentioned characteristics. If required by the external data source used to calibrate these projections, a large weight may be assigned to a group with very few observations in the initial sample and this is likely to produce unstable predictions.

Static aging cannot trace specific individuals across different moments in time. This is a major limitation since it means that analyses of mobility requiring a ('projected') panel structure cannot be carried out.

Dynamic ageing overcomes these limitations because, as shown in equation (21.11), it consists of simulating the attributes (or characteristics X) of each individual at time $t+1$ using the attributes at time t. In this case, a set of equations links the characteristics of a specific individual from one period to the next effectively generating a *simulated* panel dataset. The central issue in dynamic aging is the specification of micro *processes* of type (21.11). A first best option would be to have behavioral equations based on

[23] As mentioned in the text, there are some connections between static ageing in a microsimulation setting and the projection methods used by demographers. A commonly used demographic projection method – the cohort component method – starts with a base-year population classified into different age and gender groups. Then rates of mortality, fertility and migration are assumed for each group, and used to project their evolution through time. This procedure provides, for every year of the projection, a distribution of the population by age and sex. However, this method does not reveal anything about the identity of the individuals within any of the groups. In other words, demographers are not interested in following individuals along their life paths. This is the key interest of dynamic microsimulation. 'Anonymous' demographic projections are useful because they provide a 'macro' benchmark for the mircrosimulations. Specific individuals followed by the microsimulations end up in certain demographic groups; the size of these groups should correspond to the size given by the 'macro' demographic model. An application of this type of demographic macro and microsimulation micro projection of a population by age, gender and education level is described in Section 21.5.2.1.

microeconomic theory. In other words, agents' behavior – in terms of choices concerning marital status, area of residence or type of job throughout their life paths – would be determined by their preferences and other constraints such as, for example, the economic policy environment. In this context, agents' behavior could be represented by a structural model. Most of the times, this is too demanding given the available micro datasets, economists' knowledge of microeconomic behavior and econometric estimation techniques. Therefore, second best options are used. They consist of using models in reduced form that combine deterministic and stochastic components. A stochastic component incorporates the effect of incomplete model specification or reflects the genuine random nature of a process.

Three examples (marriage, fertility and education) are helpful to clarify some of these methodological issues.

The first example considers the marriage decision. This is normally modeled in two steps: (i) individuals are selected to be eligible for marriage, and (ii) men and women are matched to one another. The main objective of the so-called mate-matching techniques, i.e. the techniques used to generate synthetic marriages, is to 'replicate the central tendency as well as the dispersion of the assortative mating patterns observed in the real world'.[24] Once females and males are selected into the 'marriage market', normally because of some of their demographic characteristics (age, single) and a Monte Carlo lottery, they are matched using a mix of deterministic and stochastic methods. The likelihood of a potential couple to form a household is estimated using data on the gender differences in age, education and current or potential earnings and some other relevant characteristics, such as area of residence, ethnicity, etc. The estimated probability is then used to rank all potential couples and, in the 'stable marriage' approach (see Gale and Shapely, 1962) an optimization method is used to generate a sequence of marriages.[25]

It has been shown that this approach has several drawbacks and that it produces a distribution of marriages that is not very realistic. Alternative matching methods introduce stochastic components in the preceding approach. In these approaches, a 'Monte Carlo' lottery generates a random number from a uniform distribution over the interval (0,1). A marriage occurs if that number is smaller than the probability estimated from the matching equation and the newly married are taken out of the list of potential marriages.

Accurately simulating the marriage decision can be fairly important. As Perese (2002) points out, age and income differences in married couples influence other important household decisions including, amongst other, fertility, participation in the labor market,

24 Perese (2002, p. 2).

25 Once the ordered (according to the likelihood of a match or compatibility amongst the man and the woman) list of all possible marriages is compiled, the optimization process works as follows. The best first pair, i.e. the most likely, is married. Then all other potential marriages that included either the man or the woman from the first marriage are dropped from the list and all remaining ones are reranked. Then the new best pair is married and the process is repeated.

savings and migration. Furthermore, social security benefits and pensions, both in terms of eligibility and amounts, are affected by the composition of the household.

The second example is about the fertility decision. This can be modeled by combining an estimated equation and some external projection of aggregate population growth. The estimated equation is derived from a probit function that represents the probability that a woman give birth in a given time period (see Cogneau and Grimm, 2008). Following population economics, this probability increases with marriage, rural residence and age (up to a point and at a decreasing rate), and it diminishes with number of children already in the family, education, the size of the household (in particular the number of other women) and urban residence. The estimated probabilities can then be used to rank women from the most likely to the least likely to have a child and a number of births can be adjusted so as to match external fertility projections. A random component can also be introduced to determine the gender of the child.

The final example describes the microsimulation of education decisions. As in previous cases, different estimation techniques are available. Using household survey data, a first set of methods estimates the probability of going to school and that of dropping out. These probabilities are functions of individual characteristics, previous educational attainments and the level of education of the parents.[26] Other methods rely on models with more structure. For instance, Bourguignon *et al.* (2003) use a structural model with the following three set of equations: (i) an occupational choice equation that, for young individuals, determines the probabilities of either not attending school, attending school and not working, or attending school and simultaneously working; (ii) an equation that determines the distribution of educational attainments, i.e. an individual's choice of final educational attainment (in terms of years of schooling); and (iii) a wage equation determining the wage levels. These equations interact since the predicted wage level enters the occupational choice equation. This structure embeds a more realistic behavior of agents and it allows more interesting analysis of policy reforms.

As briefly illustrated by these examples, modeling key demographic and economic behaviors at the micro level can be fairly complex. Depending on the objectives of the microsimulation, additional life events or behavior (e.g. mortality, savings and wealth accumulation) may need to be included. Then, one of the following two approaches can be used. The first approach consists of building a *cohort model*. This type of model ages a specific sample composed of unrelated individuals of the same age. The models' data and computing requirements are lower than those of the second approach, labeled *population models*, which include a sample of individuals of different ages, and require tracking both the descendants and the ascendants of individuals.

[26] Note that to estimate these transitions in and out of education, cross sectional surveys have to report on the current enrollment and enrollment at a previous (usually 12-month) period.

A final methodological issue concerns the *treatment of time* and the related issue of transition matrices versus survival/hazard functions. Many microsimulation models are organized around discrete time intervals, usually of a year. Every year, agents in the model are 'aged' and specific events happen to them. Cumulatively many events occur that, at the end, represent the whole life path of agents. These discrete time models usually rely on yearly transition matrices that represent the probabilities of change in status. They have two main drawbacks. First, it is not possible to take into account intrayear phenomena. Transitions in and out of employment that last less than a year cannot be captured by a model with a yearly cycle, but this may not be major problem. A second and more important shortcoming is that results are not always robust to the sequencing of events. In the real world, many decisions (and related life events) are interdependent. A woman's decision to participate in the labor market is affected by the number of children she has and her marital status. These, in turn, are simultaneously affected by her participation decision. However, in a discrete time model, a fixed sequencing of events determines the direction of causality. In such a model, an agent is faced first with the choice of family formation and then with the choice of labor force participation; the latter is affected by the former, but not *vice versa*.[27]

Survival (or hazard) functions, which represent the risk of facing a *range* of events, are an alternative approach that overcomes these limitations. Depending on the characteristics of the agent, these functions model the duration of a certain status and operate with continuous rather than discrete time. Even if this approach is theoretically appealing, its practical implementation remains difficult. Most data are still collected on a discrete annual timeframe and estimation of joint probabilities in continuous time poses serious econometric difficulties. Some of these are discussed Section 21.4.3.

21.4.3 Data, estimation and validation issues

The data and estimation requirements for dynamic aging are formidable even when the analysis is limited to a relative small set of endogenous behaviors. Besides, even when data are available and estimation feasible, the projected macroeconomic trends obtained by aggregating micro decisions may not match results from macro models (see below) or demographic projections. Therefore, most microsimulation models combine static and dynamic (or behavioral) aging techniques or adopt some form of *alignment/calibration* of the model results with external projections. In a few cases, validation of the model results is also carried out. A brief description of these issues follows here.

Starting with the data, comprehensive datasets are needed for (i) setting the starting values of the simulations, (ii) estimating behavioral and other functions, and (iii) validating

27 In the relevant literature, this path dependency is often defined as the notion of competing risks. Different outcomes (e.g. having a child, deciding to participate) are *competing* to manifest themselves and the sequencing in their realization may not be fixed as requested in a discrete-time model. See Galler (1995).

the model. A long comprehensive panel (or longitudinal) dataset can be used to serve these three purposes. However, panels are available for few countries, and may suffer from problems of attrition and measurement errors; they may not represent the whole population and may not include all the socioeconomic variables of interest. Moreover, they are dated and represent the past. Practically, most applications end up combining data from different sources ranging from censuses to cross-section surveys and administrative data. Harding (2007) provides a convincing description of data challenges when discussing the Australian case. In this case, the availability of a longitudinal dataset that samples around 7000 households and has accumulated five waves from 2001 to 2002 represents a vast improvement in the availability of longitudinal data within Australia. However, even with a good-quality and almost 10-year long panel dataset, problems remain. In particular, the relatively small sample size, apart from introducing sampling errors, limits the possibility of estimating behavior with models that reasonably include many explanatory variables.

Beyond the issue of data quality and availability, there are estimation problems. Klevmarken (2008) gives a clear and lucid description of them. Many economic variables have distributions that are heavily skewed and with high kurtosis. In that respect, data may suffer from censoring. On the other hand, extreme outliers are common for variables measuring incomes, assets and wealth. Another difficulty is that many decisions and life events that need to be captured in a dynamic microsimulation model are essentially discrete. OLS estimation may not be appropriate in these cases and non-linear estimation becomes cumbersome when considering more than two discrete variables simultaneously. The model structures have then to be recursive or block-recursive despite the lack of strong theoretical reasons to think that the recursive structure of the model represents reality. As mentioned above in these recursive models, the sequencing of the estimation and simulation is crucial since variables determined in the first stages can affect variables downstream but not the other way around.

Perhaps the most serious problem in these microsimulation models is the assumption of stability of the parameters (for reduced form models) and even of the structure of the model (for structural models) over the time horizon under consideration. This assumption makes these models open to the Lucas' critique. If the model parameters (or its structure) are not autonomous with respect to policy changes, then it becomes impossible to use the model to study the effects of such changes. This problem is particularly serious for reduced form models where complex behavior is 'summarized' in a few parameters. Structural models that explicitly account for changes of policy regimes and allow for behavioral responses are less sensitive to this critique. In the end, model stability is an empirical question. However, it may not be possible to verify the stability of microsimulation models used for projections out of the sample and into the future, where policy variables assume values in wide intervals.

In some cases, researchers partially remedy these problems by validating the microsimulation model through comparing some runs with actual data. This validation can be a fairly formal process where the stochastic properties of a validation measure, e.g. the distance between the model results and actual observations, are derived. A common method consists of systematically changing the values of the model parameters by drawing new values within a multivariate normal distribution with estimated means and covariance matrix of the parameters. At every new draw the simulation is rerun and the variability of the results due to the uncertainty about parameter values can be assessed. This straightforward approach can become time-consuming with large models. Besides, variability of the model results can also depend on uncertainty about the model structure. Sensitivity analysis can be used to identify which parts of the model, if changed, have the largest impact on results. However, the ultimate question of whether a model produces an accurate representation of a society can only partially be answered. Only progress in data quality, economic theory and econometric inference methods can help produce more convincing answers to this ultimate question.

21.4.4 Linkages with macro models

Most dynamic microsimulation models simulate the behavior of consumers and workers, but have no market-clearing mechanism (they do not include the suppliers of goods and demanders of labor, i.e. the firms). This creates the need for having a source of information on macro variables such as the level of wages, the change in relative goods prices, unemployment rate and others.

One solution is to project these macro variables according to the trends of the recent past; however, there are two issues: (i) these macro assumptions are not necessarily consistent with each other and (ii) they are not necessarily consistent with the aggregated micro behavior from the microsimulation.

As for the cases of static microsimulation described in the first part of the chapter, a more ambitious solution is to link the microsimulation to a dynamic CGE model and this is the topic of Section 21.5.

21.5 GIDD MODEL AS AN EXAMPLE OF A GLOBAL DYNAMIC MACRO–MICRO MODEL[28]

Previous sections in this chapter described in detail the various ways of linking CGE models to microsimulation frameworks. Advantages as well as limitations and issues for further research have been highlighted. However, most of what has been described applies to single-country *static* applications. This section instead describes a modeling tool – GIDD – capable of generating 'reasonable' predictions of how *global* inequality and poverty might evolve under different future scenarios. Thus, GIDD provides a useful example of a *global dynamic* macro–micro modeling framework.

[28] This section draws heavily from work that appeared in Bussolo *et al.* (2010, 2011, 2012).

21.5.1 Motivation and originality

While building on past efforts, GIDD introduces important new features. It is the first macro–micro global simulation model: its extensive coverage allows it to address questions that would not be tractable with other methods. For example, GIDD can assess growth and distribution effects of global policies such as multilateral trade liberalization or mitigation of climate change damages, among others. The global nature of the modeling framework permits decomposing inequality dynamics into a component due to changes in average income between countries and a component due to widening disparities within countries. A second important novelty is that GIDD can explicitly consider long-term time horizons during which changes in the demographic structure may become crucial components of both growth and distribution dynamics. The explicit long-term features of GIDD can capture the impacts of aging and other demographic changes, such as the skill composition of a population, which may become crucial components of both growth and distribution dynamics.

GIDD is based on data from household surveys for 132 countries in the world covering about 91% of the total world population for the base year of 2005, see Table 21.1. Over 1 million households are sampled in 63 developing countries, while household information for 21 developed countries comes from the Luxemburg Income Study dataset. Detailed survey data for these 84 countries is combined with more aggregate information (usually vintiles) for the remainder of the world.

From the base-year household-based data, an initial global income distribution is estimated. A series of counterfactual distributions can then be simulated by linking a global CGE model with the micro data in a top-down approach (as described in Section 21.2). Counterfactual distributions are obtained by applying three main changes to the initial distribution:

(i) Demographic changes, including aging and shifts in the skill composition of the population (shifts in education).

Table 21.1 GIDD population coverage by region

Region	Countries surveyed	Covered population	Actual population	Percent covered
World	132	5513123	6076509	91
East Asia and Pacific	10	1749255	1817232	96
Eastern Europe and Central Asia	14	474468	474468	100
High-income countries	30	767291	974612	79
Latin America	24	503418	515069	98
Middle East and North Africa	8	192128	276447	70
South Asia	6	1336922	1358294	98
Sub-Saharan Africa	40	489642	663305	74

(ii) Shifts in the sectoral (agriculture/non-agriculture) composition of employment.
(iii) Economic growth, including changes in relative wages across skills and sectors.

The remainder of this section will describe the details of how these changes are transmitted from the global CGE to the micro data, in other words the solution algorithm for this global summary of interesting applications and counterfactual analyses carried out with GIDD will conclude this section.

21.5.2 GIDD solution algorithm

Recalling the definitions and notation used in Section 21.4.1, the objective of the GIDD methodology is to define *counterfactual* values for the income distribution D_t. This can be done by simulating counterfactual values for endowments (or characteristics $\boldsymbol{X}$) and market rewards for these endowments (represented by $\boldsymbol{\beta}$ in equations 21.9–21.11). This is not a minor task and becomes even more challenging when done for many countries. To keep the process manageable, the functional form of equation (21.10) has to be defined in a simple fashion using only those independent variables available for all countries in the sample. GIDD's right-hand side variables (the $\boldsymbol{X}$s) include age, education endowments and sector of employment of the household head (country subscripts excluded for simplicity):

$$y_{i,t} = \alpha_t + \beta_{1,t}\left(D^{\mathrm{A}}_{i,t}D^{\mathrm{A}}_{i,t}\right) + \beta_{2,t}\left(D^{\mathrm{NA}}_{i,t}D^{\mathrm{s}}_{i,t}\right) + \beta_{3,t}\left(D^{\mathrm{NA}}_{i,t}D^{\mathrm{us}}_{i,t}\right) + \sum\nolimits_{k=1}^{k}\gamma_{k,t}F_{k,t} + \epsilon_{i,t}, \tag{21.12}$$

where D^{A} and D^{NA} are dummy variables taking the value of 1 if the household head is employed in the agricultural sector or in the non-agricultural sector, respectively; D^{s} and D^{us} are dummy variables identifying skilled and unskilled household heads, respectively. F_k captures the proportion of household members in each of the K age cohorts. The βs are rewards (prices) to education endowments conditional on the sector of employment and γs are income determinants associated with household composition (earnings of additional workers, family size, etc.). Finally, ϵ summarizes all other income determinants. The counterfactual expression to (12) is:

$$\begin{aligned}\hat{Y}_{i,t+1} = {}& \hat{\alpha}_{t+1} + \hat{\beta}_{1,t+1}\left(\hat{D}^{\mathrm{A}}_{i,t+1}\hat{D}^{\mathrm{s}}_{i,t+1}\left(\hat{F}_{k,t+1}\right)\right) + \hat{\beta}_{2,t+1}\left(\hat{D}^{\mathrm{NA}}_{i,t+1}\hat{D}^{\mathrm{s}}_{i,t+1}\left(\hat{F}_{k,t+1}\right)\right)\\ &+ \hat{\beta}_{3,t+1}\left(\hat{D}^{\mathrm{NA}}_{i,t+1}\hat{D}^{\mathrm{us}}_{i,t+1}\left(\hat{F}_{k,t+1}\right)\right) + \sum\nolimits_{k=1}^{k}\hat{\gamma}_{k,t+1}\hat{F}_{k,t+1} + \epsilon_{i,t}\end{aligned} \tag{21.13}$$

where the demographic characteristics, endowments and returns to these endowments have been modified in accordance with the counterfactual scenario. Note in particular that changes in employment sector – and to a lesser degree in skill – explicitly depend

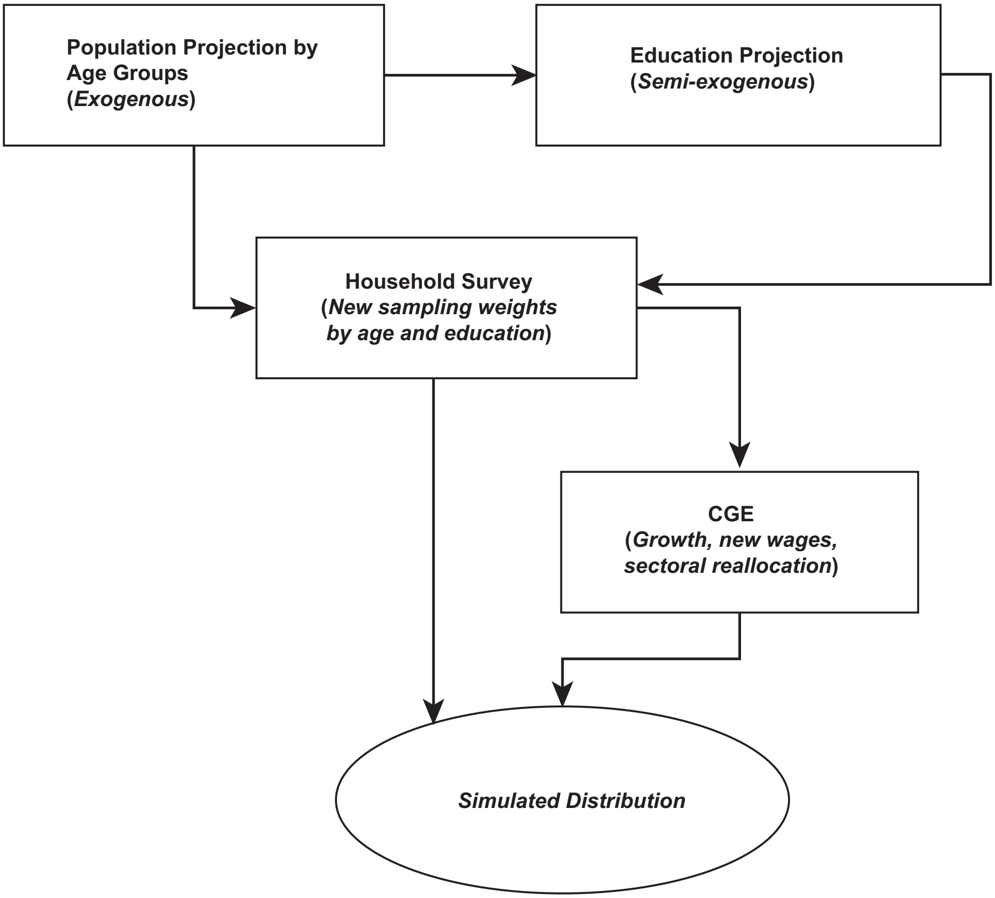

Figure 21.5 GIDD methodological framework.

on the composition of the household/ The intercept captures the *per capita* economy-wide growth rate $(\hat{\alpha}_{t+1})$.[29]

In reality, the model parameters change simultaneously; however, for simplicity and tractability, GIDD modifies each of these sequentially, as shown in Figure 21.5.

The first step consists of accounting for changes in the size of groups formed by age and education characteristics (top boxes of Figure 21.5). The impact of these changes on the labor supply is used as an input into the CGE model (link between the middle and bottom rectangles). In the second step, the CGE model produces a counterfactual scenario for a set of linkage aggregate variables (LAVs), which include overall economic

29 The intercept in fact captures the residual average rate of growth, after all other changes – demographic, endowments, and prices – have already been taken into account.

growth, growth in relative incomes by skill (education) and sector, sectoral reallocation of labor, and a new vector of consumer prices. Finally, the changes in the LAVs are passed on to the counterfactual income distribution that has already been adjusted for changes in the age and education structure (bottom link in Figure 21.5). These steps are described in more detail below.

21.5.2.1 Demographic and educational changes

The first step consists of changing the structure of the population in terms of age and education. This entails two separate substeps: (i) an exogenous aggregate (or macro) projection for these two characteristics is constructed, and (ii) all individuals become older and specific individuals, in each of the base-year surveys, actually acquire a certain level of education. These micro-level changes will have to be *aligned* with the aggregate changes of the first substep.

Macro projections for joint age/education-level population groups are normally not available. However, micro data shows that for most countries in the world, younger generations tend to be better educated than older ones. With aging the population becomes better educated; in other words, the joint distribution of age and education endowments is changing in a definite way (Lutz and Goujon, 2001). In GIDD, these macro projections were run as follows. For each country-specific survey, the population is partitioned into *m* groups according to age and educational attainments. In particular, age cohorts are set to correspond to the age cohorts used by the UN in its demographic projections[30] and three education categories are considered: primary, secondary and tertiary. The fact that the average educational attainment increases as younger better educated cohorts replace older less educated cohorts can be labeled the 'pipeline' effect – hence the process is referred to as 'Semi-exogenous' in Figure 21.5.

A stylized version of the projection methodology is shown in Table 21.2. In this hypothetical country at the starting time, 60% of young people are skilled, i.e. have a high level of education, and 40% are unskilled. The old cohort has lower levels of education: 30% of old individuals are skilled. With the passing of time the younger cohort replaces the old one; by 2030, both young and old cohorts have the same distribution of education. In the right panel of the table, the same situation is depicted in symbols: there are four initial (year 2000) population age/education subgroups (P_m) and four final (year 2030) age/education subgroups ($\hat{P}_m$).

Note that this approach assumes that the new (i.e. in 2030) young cohort has not improved its education *vis-à-vis* the young cohort of the initial year. This strong assumption is equivalent to keeping the within-cohort school entrance and graduation rates constant over time or, in other words, it is equivalent to maintain the same

[30] For more details on the UN demographic projections, see: http://esa.un.org/unpp/.

Table 21.2 Demographic projections: age and education

	Skilled	**Unskilled**	**Skilled**	**Unskilled**
2000				
Young	60	40	P_1	P_2
Old	30	70	P_3	P_4
2030				
Young	60	40	$\hat{P}_1$	$\hat{P}_2$
Old	60	40	$\hat{P}_3$	$\hat{P}_4$

efficiency and policies in the education sector. More expansionary education policies may accelerate the increase in the average educational attainment beyond the pure pipeline effect and, if available, can be easily accommodated in the macro estimation process.

Once the macro target growth rates for the m (age/education) population groups are obtained – using the UN age-cohorts projections combined with the pipeline effect (or some other assumption on education) – the subgroups $\hat{P}_m$ in Table 21.2 are determined and the first substep is completed. The second substep is to determine which exact individual, in each of the base-year surveys, will actually acquire a certain level of education while becoming older. Ideally a full micro model that accounts for the joint fertility (including marriage) and education (and possibly health) choices would be required to identify these specific individuals. However, estimating such a model is a difficult task even for a country with excellent household data and it quickly becomes unfeasible – at least with the currently available data – at the global level. Therefore GIDD methodological choice is to use a simpler non-parametric (i.e. non-behavioral, or static aging) approach to represent changes in the demographic structure. A new set of m age/education groups is produced by reweighting the original surveys, such that each additional member of the population within each partition is an exact replica of the average member of each partition before the reweighting.

More formally, the problem is that of calculating the new weights. The original weights can be represented by a matrix $\mathbf{W} = [w_{m,n}]$, where m stands for the age/education groups used in GIDD and n for individuals within the population. Of course, $w_{m,n} = 0$ *if individual n does not belong to the age/education m.* The sum of all weights in $\mathbf{W}$ generates the total population and the sum over n for a given m produces the total weight of the age/education population subgroups P_m:

$$
\begin{aligned}
P &= \sum_{m=1}^{M} \sum_{n=1}^{N} w_{m,n} = \mathbf{i}'_m \mathbf{W} \mathbf{i}_n \\
P_m &= \sum_{n=1}^{N} w_{m,n} = \mathbf{W}\mathbf{i}_n \ \forall m = 1, \ldots, M,
\end{aligned}
\tag{21.14}
$$

where $\mathbf{i}_n$ and $\mathbf{i'}_m$ are identity column and row vectors, respectively. The target new age and education population groups (i.e. the macro projections discussed above) is obtained through the reweighting procedure:

$$\hat{P}_m = \sum_{n=1}^{N} a_{m,n} w_{m,n} = (\mathbf{A} \circ \mathbf{W})\mathbf{i}_n \ \forall m = 1, \ldots, M, \tag{21.15}$$

where $\mathbf{A} = [a_{m,n}]$ is a matrix of multipliers that ensure that the m constraints on the targeted structure of the population $\hat{P}$ are satisfied and $(\mathbf{A} \circ \mathbf{W})$ is the Hadamard product. This system has $(m \times n) - 1$ variables, but only m constraints, and is therefore underdetermined.[31]

An example using a situation similar to the one depicted in Table 21.2 may clarify the reweighting procedure. Imagine a classification of four age/education groups ($m = 4$) and a sample of eight individuals ($n = 8$).

Equations (21.15) above can be rewritten as follows:

$$\begin{aligned} \hat{P}_1 &= a_{1,1} w_{1,1} + a_{1,2} w_{1,2} \\ \hat{P}_2 &= a_{2,3} w_{2,3} + a_{2,4} w_{2,4} \\ \hat{P}_3 &= a_{3,5} w_{3,5} + a_{3,6} w_{3,6} \\ \hat{P}_4 &= a_{4,7} w_{4,7} + a_{4,8} w_{4,8} \end{aligned} \tag{21.16}$$

This basically assumes that individuals 1 and 2 belong to group 1 (young and skilled in the terminology of Table 21.2), individuals 3 and 4 to group 2 and so on. From the *macro* projections, we know the $\hat{P}_m$ and we have the initial individual (household) weights $w_{m,n}$s. The system of equations (21.16) is used to determine the $a_{m,n}$s.

In the system above we have eight unknowns and four equations, so this system is not enough to determine the multipliers.[32] We need additional constraints and GIDD has two possible solutions. These solutions are to add equations to make the system exactly identified, or to solve an optimization problem that minimizes the distance between the original matrix $\mathbf{W}$ and the final matrix $(\mathbf{A} \circ \mathbf{W})$. Both solutions are available in GIDD.

21.5.2.2 Macroeconomic changes

The new population groups are used as inputs in a global CGE. A general equilibrium is needed to capture in a theory consistent fashion the multiple complex changes that affect

[31] Note also that there is an additional implicit constraint: the targets $\hat{P}_m$ of the subgroups add up to the total population $\hat{P}$ (either originally or following normalization by the user); this makes one of the equations linearly dependent of the others and can be dropped.

[32] Note also that there is an additional implicit assumption: the subgroup targets $\hat{P}_m$ add up to the total population $\hat{P}$ (either originally or following normalization by the user); this makes one of the equations linearly dependent of the others and allows us to drop it.

growth and income distribution. For example, population aging is generally correlated with declining saving rates and changing demand patterns, while, other things being equal, rising average skill endowments reduces the observed skill wage premia. In an increasingly globalizing world, the direction and magnitude of these changes will also be affected by the changing patterns of international flows of goods, services and capital.

The global CGE model used in GIDD is the World Bank LINKAGE model, but other CGE models could be used. A detailed description of LINKAGE is available elsewhere (van der Mensbrugghe, 2006) so just three particularly relevant aspects of LINKAGE (for the purposes of GIDD) are briefly described here: (i) the multifactor nested production function that includes labor (skilled and unskilled), capital, land and natural resources; (ii) the demand structure that is modeled through an extended linear expenditure system (ELES), with cross-price and income elasticities; and (iii) the sector-specific productivity growth trends that are 'calibrated' to be consistent with historical evidence.

The nested constant elasticity of substitution (CES) production functions embed various substitution possibilities across inputs and determine demands for the various types of factors. Depending on the assumptions on the supply side (which is set to be consistent with GIDD population subgroups) and on intersectoral mobility, returns to factors of production can be type- and sector-specific. In standard GIDD applications, capital and well as skilled labor are perfectly mobile across sectors within a country, while the market for unskilled labor is segmented into farm- and non-farm categories. Within each segment, labor is perfectly mobile across activities, but mobility across segments is limited by a migration function which responds to changes in the farm/non-farm wage premia.[33]

The allocation of household budget (for a single representative household in each country) across saving and a vector of consumption commodities is determined simultaneously through maximization of an ELES. The system captures various substitution possibilities across commodities as well as a gradual shift in demand towards commodities with higher income elasticities (e.g., manufacturing and services) over time.

Total factor productivity trends are also sector- and factor-specific, and are calibrated to be consistent with historical evidence as well as World Bank's near- and medium-term GDP growth forecasts.

21.5.2.3 Microeconomic changes

Once *shocked* with demographic changes, sector- and factor-specific productivity growth, and any other exogenous change (in policy variables or other shock), LINKAGE generates a set of aggregate changes in linkage variables. These include changes in the sectoral allocation of workers, in the factor returns and in the price of goods. Let us consider these three changes in more detail.

[33] The LINKAGE model also allows for international mobility of labor and capital as well as changes in the unemployment rate, but none of these possibilities are currently modeled within GIDD.

At the macro level, the CGE model will predict the number of workers moving out of the traditional agricultural sector into the relatively modern industrial and service sectors. At the micro level, the macro constraint of moving a number A of workers out of agriculture and into manufacturing and service activities can be obtained through a large number of combinations.

GIDD employs the methodology described above in Section 21.3.3. It estimates a set of parameters that describe the conditional probability function of being a worker in the non-agricultural sector given some personal characteristics. It then ranks the workers in the agricultural sector according to their probability score and assigns migrant status to the A workers with the highest score. Currently, this procedure is implemented at the household level — where the head of household makes the migration decision and takes the rest of the household members with her — although the methodology can also be applied at the individual level.

A new income has to be assigned to these migrants. This is done using a Mincer equation that is estimated for workers in agriculture (A) and non-agriculture (NA) sectors such as:

$$\ln(Y)_{j,s} = X_j\beta_s + \epsilon_{j,s}, s = (\mathrm{A}, \mathrm{NA})$$

Movers retain their endowments X_j and the residual term $\epsilon_{j,\mathrm{A}}$, but they will receive the rewards β_{NA} that are receiving the incumbents in the non-agricultural sector. In addition, their initial residual will be rescaled to account for differences in the distribution of unobservables between the agricultural and non-agricultural sectors.

This process of choosing movers and assigning new remunerations completes the first step of microeconomic changes. However, this alone does not guarantee that average incomes by sector and by education at the micro data level match the aggregate LAVs. Therefore, two additional changes are carried out: (i) a new vector of earnings $\beta'_{s,b,t}$ is imposed on each worker, conditional on that worker being in sector s and having an educational attainment e, and (ii) distributionally neutral adjustment is imposed on every worker so that the change in the average *per capita* income from the CGE is matched by the change in the average *per capita* income in the household surveys.

There are two potential difficulties in translating the price changes of the CGE model into the micro data. (i) Following the implementation of the reweighting and migration routines certain changes have already taken place both in the average survey income and its distribution. Therefore, the macro constraints on changing returns to sector and skills $[\gamma_{s,l}]$ as well as the average income $\overline{y}$ are imposed *net* of the changes that have already taken place up to this stage. (ii) Achieving full consistency between macro and micro data is often difficult if not impossible.[34] Since there is no

[34] See the discussion in Bourguignon *et al.* (2008) for a more detailed statement of this consistency problem and some examples.

Table 21.3 Global income inequality

Index	2000	2030	Dispersion only	Convergence only
Gini	0.672	0.626	0.673	0.625
Theil	0.905	0.749	0.904	0.749
Mean log deviation	0.884	0.764	0.893	0.759

Source: Authors' own calculations using data from GIDD.

guarantee that the first period wages in the CGE model match the labor earnings in the micro data, directly passing the changes in factor returns from the former to the latter may result in inconsistent evolution of wage premiums in the two models. In extreme cases, wage gaps may even be reversed in one model but not in the other. In order to hedge against these potential complications while ensuring maximum consistency between the macro and micro outcomes, GIDD adjusts the ratios between wage premiums rather than wages themselves.

21.5.3 Business as usual scenario and the global distribution of income in 2030

The procedures just explained are initially used to establish a 'business-as-usual' scenario where past trends are minimally adjusted and no policy change or other shocks are applied. This allows us to highlight how the global income distribution may evolve in the next two decades and to identify what are the main drivers of its changes. Convergence in average growth rates across countries is a major driver in the reduction of global inequality. More specifically, since reductions in average incomes differentials are weighted by population, rapid growth of populous poor countries like China and India have a huge impact on global inequality. Overall, the Gini coefficient of the global income distribution falls from 0.672 in 2000 to 0.626 in 2030. In Table 21.3, the overall change in inequality is decomposed into two parts. In the 'Dispersion'-only column, only the within country inequality is changed, but average income disparities among countries are left unchanged. In the 'Convergence'-only column, within country inequality is unaffected and only average country incomes are modified. Table 21.3 shows that the reduction in global income inequality between 2005 and 2030 is the outcome of two opposing forces: the inequality-reducing convergence effect and the inequality-enhancing dispersion effect.

Three main findings can be stressed. (i) Even with significant changes of within-country inequality levels, all the potential reduction of global inequality can be accounted for by the projected convergence in growth rates of average incomes across countries. (ii) The aggregate impact of the changes of the within-countries component of inequality appears to be minor even though specific countries, and specific households' types within countries, may experience large distributional shifts.[35] (iii) A main cause of within-country inequality changes is the adjustments of factor rewards.

To highlight the importance of the second and third findings consider Figure 21.6 Within each country, income distribution is affected by two sets of factors: shifts in the *demographic structure* of the population, in terms of aging and education attainment, and changes in *rewards for individuals' characteristics*, such as their education level, experience, sector of employment, etc. In the real world these demographic and economic shocks occur simultaneously and jointly determine inequality changes, but Figure 21.6, based on GIDD model, decomposes the total change into these two demographic and economic components.

Controlling for other factors, both the level and dispersion (inequality) of household income tend to increase with the age and education of the household head.[36] Therefore as the population shares of groups with more income inequality rise, one may expect to see higher inequality.[37] However, as shown by tick marks in Figure 21.6, there is no clear pattern in changes in inequality driven by demographic forces and these changes are contained within a band that is much narrower than total inequality changes, without being negligible, far from it.

On the other hand, widening gaps in factor rewards, and particularly in the premium paid for higher skills, tend to produce larger changes in inequality and generally determine the overall direction of the effect. This is shown in Figure 21.6, where for large changes in inequality, the distance between the blue and red markers (i.e. the change in inequality attributable to changes in economic factors) increases, a sign that economic factors are the most important determinant for the final level of inequality.[38]

35 The relatively weak impact of within country inequality changes on global inequality has been stressed by various authors on historical data (see, e.g. Bourguignon and Morrisson, 2002).

36 The relationship between household income and age of the household head is positive in approximately 70% of the sample countries, while the age-income profile is positive in 60% of the countries.

37 The literature on the evolution of income inequality identifies three channels that determine the effects of demographic change: (i) given an upward-sloping age-earnings (incomes) profile, aging will increase inequality between old and young groups (Deaton and Paxson, 1997), and (ii) different age groups are characterized by different within-group inequality and inequality tends to be higher among older-age cohorts (see Deaton and Paxson, 1997; Jenkins, 1988; Mookherjee and Shorroks, 1982). With everything else remaining constant, when older cohorts become more populous, as is the case with lower population growth rates, aggregate inequality increases. These two channels affect aggregate inequality without any change in the age premium, i.e. with a fixed age-earnings profile; however, the channel (iii) considers changes in inequality due to changes of the life-cycle income profile. As the population ages, older high-wage and more experienced workers tend to become less scarce and the wage premium they initially receive will be reduced (Higgins and Williamson, 1999). This third channel works through the labor market and contributes to attenuating the inequality increases brought about by the first two channels. This channel is explored in more detail as part of the discussion on price-wage adjustments.

38 Some of the changes in inequality shown in Figure 21.6 may seem implausible when compared with some *ex post* evidence; however, the aim of Figure 21.2 is not to present forecasts of income inequality, but rather to show what may happen, *cet. par.*, to inequality in a specific scenario for the evolution of the global economy. In particular, redistribution policies can do much to attenuate those changes.

In sum, changes in income inequality over the next 25 years are likely to be driven mainly by changes in the rewards for individual characteristics. Considering jointly the change in these rewards, the projected growth rate and the initial level of inequality, it is possible to describe a general pattern for the evolution of within-country inequality as shown in Figure 21.6. Countries with low initial inequality and fast growth, and thus upward pressure on skilled wages, are likely to experience a worsening distribution of income. Countries with slower growth rates and greater initial inequality in income are likely to see inequality fall. The results therefore illustrate a 'convergence' of income distributions across countries, which can be interpreted as a manifestation of the Kuznets hypothesis or as a consequence of the globalization-induced equalization of factor prices.

To translate these results into a more practical and policy relevant perspective, GIDD has been used to study what happens to a specific income group during the 2005–2030 time period. The group under consideration is labeled 'global middle class' (GMC) and comprises people whose income levels are between the average incomes of Brazil and Italy, in purchasing power parity terms.[39,40] The combination of the convergence and divergence components described earlier drive a dramatic increase in the size of the GMC and its profound compositional change in favor of developing country nationals. A key conclusion asserts that developing country members of the GMC are likely to become an increasingly important group within their own countries, will increase their political influence and possibly provide continued momentum for policies favoring global integration.

21.5.4 Policy simulations

The macro (CGE)–micro (GIDD) modeling framework described here is easy to criticize. Measuring global income distribution, accounting for the general equilibrium effects of growth patterns and global policy changes, and predicting the future are all very difficult things to do in economics, and they all require assumptions and some shortcuts that can then be easily attacked. However, if one accepts the premise that the ability to 'predict' – obviously subject to great uncertainty – the plausible worldwide distributional implications of large shocks and policy changes in the future, then it is not easy to propose a clearly superior alternative to this.

39 In 1993 PPP prices, the lower threshold is \$303 per person per month, while the upper threshold is \$611 per person per month. This means that *per capita* earnings of members of the GMC are 10–20 times above the international poverty line of \$1 a day. These income thresholds are due to the GMC definition proposed by Milanovic and Yitzhaki (2002).

40 A few recent studies focus on the evolution of the middle class. See, e.g. World Bank (2007), Bussolo *et al*. (2009), Ravallion (2009), Birsdall (2010), Banerjee and Duflo (2008), and Kharas (2010).

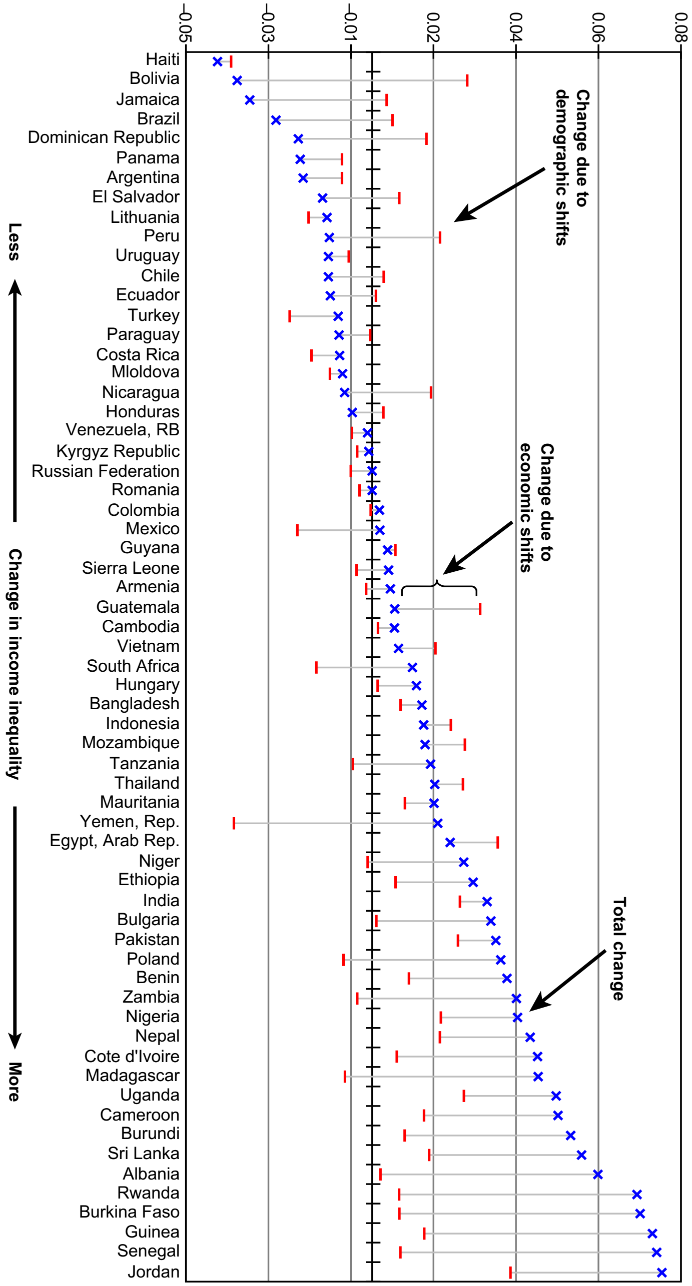

Figure 21.6 Within countries changes in inequality due to demographic and economic shifts. Final change in Gini indices for individual countries (×) and change explained by the demographic shift (–). *Note*: For each country, two points are depicted in this graph: (i) the total change in the Gini index (the blue crosses, representing the Gini level calculated on the final distribution of incomes in the year

Table 21.4 Poverty is higher among agricultural households even if their incomes are less unequal

	Gini (%)	Population shares (%)	Average monthly income (2000, US$ PPP)	$1 poverty incidence (%)	Poverty share (%)
Agriculture	44.9	44.8	65.4	31.7	75.9
Non-agriculture	62.8	55.2	328.9	8.1	24.0
World	67.0	100	210.8	18.7	100

Source: GIDD database.

In this subsection two recent applications of GIDD framework are briefly described. Both applications consider very important global policies/issues: multilateral trade liberalization and climate change. It is also useful to describe these two case to illustrate the flexibility of GIDD approach. In fact, the first uses GIDD in a simpler, not dynamic, mode, i.e. the trade liberalization is implemented in a comparative statics fashion; the second considers the long run and the implication of climate change for poverty in one to two generations.

21.5.4.1 Free trade in agriculture and global poverty

Almost 45% of the population in the world lives in households where agricultural activities represent the main occupation of the head and a large share of this agriculture-dependent group, close to 32%, is poor. Agriculture households contribute disproportionally to global poverty: three out of every four poor people belong to this group (see Table 21.4). Thus, changing economic opportunities in agriculture – in particular, liberalizing agricultural trade – can significantly affect global poverty and inequality. Direct effects of this liberalization will entail changes in the international prices of agricultural products and in the returns of factors used intensively in agriculture, with these changes determining winners and losers.

In addition to the initial trade distortions, three facts about the agricultural sector determine the overall welfare effects of a global-scale removal of agricultural distortions: (i) the proportion of the world population whose real incomes depend on the agricultural sector, (ii) the initial position of the agricultural population in the global income

2030 minus the Gini level of the income distribution of the year 2000, the initial Gini) and (ii) the change attributable to the demographic change (the red dashes, representing the Gini level for the income distribution for 2030 obtained after changes in the demographic structure, minus the initial Gini level). Countries in the graph have been ordered according to the sign and magnitude of the *total* change in their Gini index, from the largest reduction to the largest increase in inequality. *(Source: Authors' calculation based on survey data and microsimulation)*

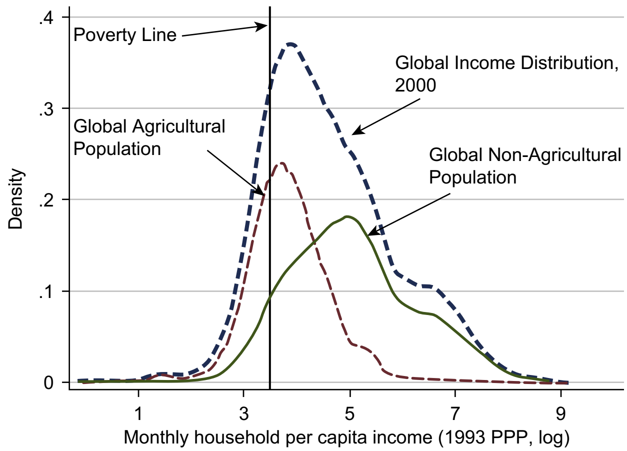

Figure 21.7 Identifying the agricultural population in the global income distribution.

distribution and (iii) the dispersion of incomes among the agricultural population.[41] Using the GIDD dataset, Figure 21.7 shows a kernel density for the global income distribution of household *per capita* income/consumption and kernel 'densities' for incomes/consumption of the population in and out of the agricultural sector, respectively.[42]

The area below the kernel 'density' for the agricultural sector is equal to 0.45, showing that 45% of the world population relies on agriculture for its livelihood. The distribution of the agricultural population is located to the left of the non-agricultural distribution implying that households in the agricultural sector earn, on average, just 20% of the (average) income of their counterparts in other sectors. The differences in shape between the two distributions corroborates what Kuznets hypothesized more than 50 years ago at the country rather than the global level, i.e. incomes in the traditional sector are less dispersed than in the modern industries. A more egalitarian traditional sector is depicted in the form of a taller and thinner distribution for agricultural population in Figure 21.7.

41 The estimates of the welfare effects of a global agricultural liberalization will also depend on the pattern of initial distortions (tariffs and subsidies) and, at least in the short term where no adjustment is possible, on the number of net consumers and net producers. Notice that, as explained in Section 21.3, our methodology allows for adjustments in the patterns of production (employment by sector changes) and consumption and thus consider the longer term.

42 The distributions for the agricultural and non-agricultural populations are not, strictly speaking, density functions since the area below the curve do not add to 1. The densities of the agricultural and non-agricultural population had been rescaled so that the area under the curve represents the proportion of the world population within these two groups.

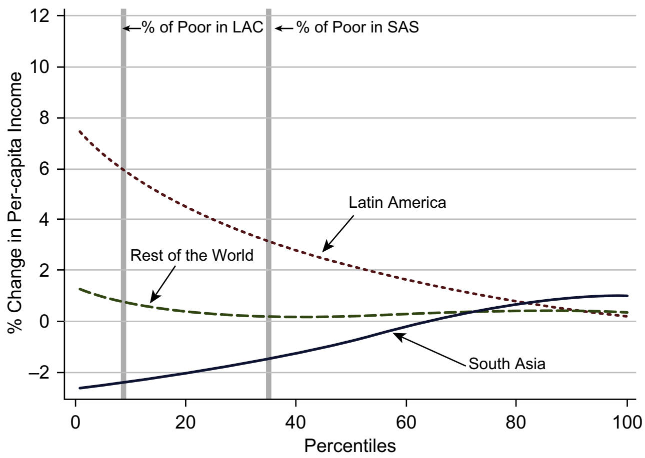

Figure 21.8 Regional growth incidence curves.

The removal of agricultural distortions has little impact on inequality at the global level. According to GIDD, the world's average monthly household income increases 0.3% after the reform. This should reduce global inequality. But it turns out that the increase in income dispersion within the agricultural sector almost completely offsets the effect of the reduction in the non-agricultural income premium on global inequality.

The increased dispersion in agricultural incomes is clearly illustrated by Figure 21.8 which plots the growth incidence curve (GIC) for Latin America, South Asia and the rest of the world. Farmers in many Latin America (LAC) countries are big winners from trade reform with impressive increases in their household income. By contrast, incomes of farmers in South Asia (SAS) shrink more than 3% after agricultural distortions are dismantled. The GIC for Latin America shows that the agriculture-based growth in the region is highly pro-poor; on the contrary, South Asia's reduction in agricultural incomes is highly regressive, with the poorest households losing from the reform. East Asia and, to a lesser extent, Sub-Saharan Africa benefit from the global reform, while the effects of the reform are progressive, albeit close to zero, for the rest of the world.

These changes in average incomes and dispersion have small effects on global poverty. Agricultural trade liberalization could increase the incidence of extreme poverty [less than $1 purchasing power parity (PPP) 1993 a day] by 0.2%, while moderate poverty (less than $2 PPP 1993 a day) is likely to fall by 0.3%. These small aggregate changes are produced by a combination of offsetting trends at the regional and country levels. With the elimination of all agriculture trade distortion, extreme poverty is reduced in all regions but in the Middle East and North Africa, where it is almost stable, and in South Asia, where it increases considerably. Since about 50% of all poor people live in South Asia, the worsening of poverty in this region counterbalances all the gains in the other parts of the world and an additional 9 million people fall into poverty. At the moderate poverty line, 14 million people escape poverty and most regions benefit from lower poverty incidence with the exclusion of Eastern Europe and Central Asia and Middle East and North Africa. Many non-agriculture households in South Asia are clustered below the \$2 a day poverty line and trade reform-related improvement in their incomes, versus the agricultural incomes' decline, explain the difference in global poverty results when the \$1 or \$2 a day lines are used.

21.5.4.2 Distributional impacts of climate change

In the second application GIDD is used to study the income distribution and poverty consequences of damages from global warming (see Medvedev and van der Mensbrugghe, 2010). The general equilibrium model with an integrated climate module and links from emissions to global temperature is solved through 2050, and climate change damages to agricultural productivity are calibrated using estimates in Cline (2007). In order to assess the magnitude and incidence of climate change damages, the baseline scenario (which incorporates climate feedbacks to agricultural productivity) is contrasted with an alternative scenario where the damage coefficient is set to zero (i.e., costless mitigation). The results show that a temperature increase of approximately 1°C above today's levels could raise the 2050 global moderate poverty headcount (\$2 per day poverty line) from 2.85% in a scenario with no damages to 3.01% when damages are taken into account. The limited global impact conceals a wider variation across regions, with increases in poverty ranging from 289 000 people in Latin America and the Caribbean to 2.7 million in South Asia and 6.2 million in Sub-Saharan Africa.

The adverse effects of global warming also vary by the main source of household earnings. Although climate change damages are concentrated in agriculture, the agricultural households are not necessarily the most affected. Due to a reduction in global output of agriculture of 1.5% (and nearly 12% in developing countries), prices for agricultural products rise and help close the wage gap between earnings in the farm and non-farm sectors. At the same time, however, the cost of the food basket rises for all consumers, including agricultural households. As a result, households in the farm sector are still likely to experience a reduction in their welfare due to higher consumption costs and the slower rate of growth in global GDP, but this reduction is likely to be less pronounced than the welfare losses for non-farm households. At the global level, these trends translate into a 0.2 percentage point increase in the non-farm poverty headcount while the headcount in agriculture rises by just 0.1 percentage points.

As the adverse impacts of global warming are more pronounced in the poor countries located closer to the equator, including climate change damages in the analysis results in an increase in the global Gini coefficient from 57.2 to 57.6 in 2050. The widening of inequality between countries is somewhat offset by the falling within component due to faster growth in the earnings of agricultural households, which tend to be concentrated in the left tail of the national distributions. These dynamics give rise to the global growth incidence curve in Figure 21.9. For each percentile, Figure 21.9 shows the difference in incomes (measured at the end of the period, i.e. in 2050) between a situation where environmental damages could be eliminated at no costs and a situation where environmental damages reduce agriculture productivity (as in the baseline scenario). The gains from eliminating, or basically not having, environmental damages are largest between the second and sixth deciles of the global income distribution. This means that households in this part of the distribution are likely to suffer the most from climate change.

21.6 CONCLUDING REMARKS

This chapter described in details the rewards and challenges of integrating state-of-the-art income distribution and general equilibrium analyses.

The largest reward of an integrated approach is the possibility of identifying winners and losers not only of changes of microeconomic policies, such as conditional cash transfers or some other targeted intervention, but also of macro policies, such as fiscal and monetary policies, trade liberalization, financial sector reforms, and other macro structural reforms. For these cases, formal evaluation techniques based on micro-simulation or possibly randomized experimentation techniques cannot be used. All individuals are affected by a macro policy or a macro shock in a way that varies across individuals and cannot be directly observed. A macro counterfactual, i.e. a situation without the macro policy or macro shock being analyzed, is needed to complete the analysis. Likewise, the macro—micro framework described in this chapter allows us to fully take into account the macro consequences of micro-oriented policies like changes in tax-benefit systems or labor market regulations.

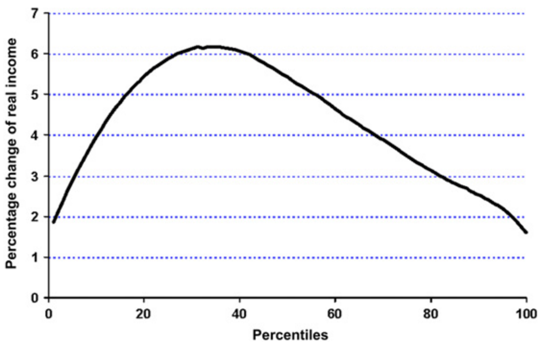

Figure 21.9 Global incidence of climate change damages: percent change in real income or consumption in 2050 relative to baseline with climate change damages *(Source: Simulations with GIDD model.)*

Although progress has been made, as shown in this paper, several big challenges remain. A practical challenge is the availability and quality of micro data. Although household-level surveys are now available for most countries, their quality in terms of thematic coverage, sample size and frequency can certainly be improved. Panel data are available in only a few countries and rarely on periods long enough so as to be able to estimate fundamental dimensions of individual behavior and, in particular, the heterogeneity of a specific behavior, be it economic like saving or demographic like fertility, across individuals. Owing to this, the way in which microeconomic behavior is modeled in integrated micro—macro models is most often rough, limited to reduced form models when not simply to transition probabilities estimated on the whole population and separately for individual changes that are clearly dependent on each other. Such an approach does not really allow representing satisfactorily individual behavior and the heterogeneity of the effects of policy reforms. Also, it is not immune to the Lucas critique.

Another important issue is that of the discrepancy between the income recorded in typical household surveys and the definition used in national accounts and implicit in the CGE part of the macro—micro models. As a consequence, the link between the two parts of the model essentially consists of a few macro variables that allow scaling up or down some components of household income in the micro database, typically wages for different labor skills, property income or dividends. By doing so, this approach ignores the fact that changes in some income components in the macro model may modify the distribution of the corresponding income component at the micro level. For instance, some reform of the pension system policies may lead households to acquire financial assets that are absent of their observed wealth in the micro database.

It must be realized that the main limitation of the macro—micro approach to distributional issues, as described in this chapter, is due as much to concepts as to data weaknesses. Empirical models cannot do better than theory, and a full dynamic theory that would permit us to include a full representation of lifetime individual behavior and its heterogeneity in the population within a dynamic and stochastic general equilibrium framework is simply not available at this stage. Some efforts are being made in that direction in the DSGE literature, e.g. in the line of Castaneda *et al.* (2003), but the representation of micro individual behavior is still much too simple to take into account the basic events that shape individual lives and welfare achievements.

Does this mean that the approaches described in this chapter are unsatisfactory to analyze the likely effects of macro policies on individual incomes as well as the macro effects of micro-oriented policies? Certainly not. The examples given in this chapter show that, even with rough assumptions about some dimensions of individual behavior, it is possible to shed light on major static and dynamic determinants of the distribution of income and overall welfare inequality in a society. This was seen for unemployment and occupational changes in a static framework, and with simply the aging of the population and the implied change in the educational structure of the population in the GIDD model. Being able to give evidence of the potential size of these effects is already a major progress, which has been made possible jointly by a better theoretical understanding of economic phenomena, better data, a better use of these data and finally by progress in computing power. There is little doubt that this is only the beginning of a journey that will lead us progressively to a more satisfactory understanding of the fundamental relationship between macro policies and the distribution of individual welfare within the population.

REFERENCES

Aaberge, R., Colombino, U., Holmøy, E., Strøm, B., Wennemo, T., 2004. Population ageing and fiscal sustainability: An integrated micro–macro analysis of required tax changes. Discussion Paper 367. Statistics Norway, Oslo.

Abowd, J., Farber, H., 1982. Job queues and the union status of workers. Ind. Labor Relat. Rev. 35, 354–367.

Abowd, J.M., Kramarz, F., 1999. Econometric analyses of linked employer–employee data. Labour Econ. 6, 53–74.

Adelman, I., Robinson, S., 1978. Income distribution in developing countries: A case study of Korea. Stanford University Press, Stanford, CA.

Ahmed, V., O'Donoghue, C., 2009. Global economic crisis and poverty in Pakistan. Int. J. Microsimulation 3, 127–129.

Arntz, M., Boeters, S., Gürtzgen, N., Schubert, S., 2008. Analysing welfare reform in a microsimulation–AGE model: The value of disaggregation. Econ. Model. 25, 422–439.

Bacon, B., Pennec, S., 2007a. Model validation: Microsimulation and macrosimulation, presentation to APPSIM Technical Advisory Group, National Centre for Social and Economic Modelling. University of Canberra, Canberra.

Bacon, B., Pennec, S., 2007b. APPSIM – family formation and dissolution. Working Paper 6. National Centre for Social and Economic Modelling. University of Canberra, Canberra. Available at www.natsem.canberra.edu.au.

Baekgaard, H., 1995. Integrating micro and macro models: Mutual benefits (Economics and Transportation). In: International Congress on Modelling and Simulation Proceedings, Volume 4. University of Newcastle, Newcastle, pp. 1–6.

Banerjee, Abhijit, V., Duflo, E., 2008. What is middle class about the middle classes around the world? J. Econ. Perspectives 22, 3–28.

Birsdall, N., 2010. The (indispensable) middle class in developing countries; or, the rich and the rest, not the poor and the rest, center for global development. Working Paper 207. Center for Global Development, Washington, DC.

Blundell, R., Macurdy, T., 1999. Labor supply: A review of alternative approaches. In: Ashenfelter, O., Card, D. (Eds.), Handbook of Labor Economics, first ed., vol. 3. Elsevier, pp. 1559–1695 (Chapter 27).

Blundell, R., Stoker, T.M., 2005. Heterogeneity and aggregation. J. Econ. Lit. 43, 347–391.

Boccanfuso, D., Savard, L., 2007. Poverty and inequality impact analysis regarding cotton subsidies: A Mali-based CGE micro-accounting approach. J. Afr. Econ. 16, 629–659.

Bourguignon, F., Morrisson, C., 2002. Inequality among world citizens; 1820–1992. Am. Econ. Rev. 92, 727–744.

Bourguignon, F., Pereira da Silva, L. (Eds.), 2003. The Impact of Economic Policies on Poverty and Income Distribution: Evaluation Techniques and Tools. World Bank and Oxford University Press, Washington, DC.

Bourguignon, F., Savard, L., 2008. Distributional effects of trade reform: An integrated macro–micro model applied to the Philippines, in Brazil. In: Bourguignon, F., Pereira, L., Bussolo, M. (Eds.), The

Impact of Macroeconomic Policies on Poverty and Income Distribution. Palgrave and the World Bank, Washington, DC, pp. 177–212.

Bourguignon, F., Spadaro, A., 2006. Microsimulation as a tool for evaluating redistribution policies. J. Econ. Inequality 4, 77–106.

Bourguignon, F., Ferreira, F., Leite, P., 2002. Beyond Oaxaca–Blinder: Accounting for differences in household income distributions across countries. World Bank Policy Research Working Paper 2828. World Bank, Washington, DC.

Bourguignon, F., Ferreira, F., Leite, P., 2003. Conditional cash transfers, schooling, and child labor: Micro-simulating Brazil's bolsa escola program. World Bank Econ. Rev. 17 (2), 229–254.

Bourguignon, F., Ferreira, F., Lustig, N. (Eds.), 2003. The Microeconomics of Income Distribution Dynamics. Oxford University Press and the World Bank, Washington, DC.

Bourguignon, F., Robilliard, A.-S., Robinson, S., 2004. Representative versus real households in the macro-economic modeling of inequality. In: Kehoe, T., Srinivasan, T.N., Whalley, J. (Eds.), Frontiers in Applied General Equilibrium Modeling: Essays in Honor of Herbert Scarf. Cambridge University Press, Cambridge, pp. 219–251.

Bourguignon, F., Ferreira, F.H.G., Menéndez, M., 2007. Inequality of Opportunity in Brazil. Review of Income and Wealth. International Association for Research in Income and Wealth 53 (4), 585–618.

Bourguignon, F., Bussolo, M., Pereira da Silva, L. (Eds.), 2008. The Impact of Macroeconomic Policies on Poverty and Income Distribution – Macro–Micro Evaluation Techniques and Tools. World Bank and Palgrave Macmillan, Washington DC.

Brown, L., Harris, A., Picton, M., Thurecht, L., Yap, M., Harding, A., Dixon, P., Richardson, J., 2007. Linking Microsimulation and Macro-Economic Models to Estimate the Economic Impact of Chronic Disease Prevention. NATSEM, University of Canberra.

Bussolo, M., De Hoyos, R., Medvedev, D., 2009. The future of global income inequality. In: Estache, A., Leipziger, D. (Eds.), Stuck in the Middle. Brookings Institution, Washington, DC, pp. 54–74.

Bussolo, M., De Hoyos, R., Medvedev, D., 2011. Free trade in agriculture and global poverty. World Econ. 34 (12), 2019–2043.

Bussolo, M., De Hoyos, R., Medvedev, D., van der Mensbrugghe, D., 2008. Global Climate Change and its Distributional Impacts. World Bank, Washington, DC.

Bussolo, M., de Hoyos, R.E., Medvedev, D., van der Mensbrugghe, D., 2012. Global growth and distribution: China, India, and the emergence of a global middle class. J. Global. Develop. 2 (2), 3. (Berkeley Electronic Press).

Caldwell, S., 1996. Health, wealth, pensions and lifepaths: The CORSIM dynamic microsimulation model. In: Harding, A. (Ed.), Microsimulation and Public Policy. North-Holland, Amsterdam, pp. 505–522.

Caldwell, S., 2006. Dynamic microsimulation: Directions for improvement, paper presented to the 8th Nordic Seminar on Microsimulation Models Oslo.

Castaneda, A., Diaz-Gimenez, J., Rios-Rull, J.V., 2003. Accounting for the US earnings and wealth inequality. J. Polit. Econ. 111, 818–857.

Chen, S., Ravallion, M., 2004. Welfare impacts of China's accession to the World Trade Organization. World Bank Econ. Rev. 18, 29–57.

Cline, W.R., 2007. Global Warming and Agriculture: Impact Estimates by Country. Peterson Institute Press. Peterson Institute for International Economics, number 4037, January.

Cockburn, J., 2006. Trade liberalisation and poverty in Nepal: A computable general equilibrium micro simulation analysis. In: Bussolo, M., Round, J. (Eds.), Globalization and Poverty: Channels and Policies. Routledge, London, pp. 175–195.

Cogneau, D., Grimm, M., 2008. The impact of AIDS mortality on the distribution of income in Cote d'Ivoire. J. African Econ. 17 (5), 688–728.

Cogneau, D., Robilliard, A.-S., 2008. Simulating targeted policies with macro impacts: Poverty alleviation policies in Madagascar. In: Bourguignon, F., Bussolo, M., Pereira da Silva, L. (Eds.), The Impact of Macroeconomic Policies on Poverty and Income Distribution – Macro–Micro Evaluation Techniques and Tools. World Bank and Palgrave Macmillan, Washington, DC, pp. 213–246.

Cowell, F., 2000. Measurement of Inequality. In: Atkinson, A.B., Bourguignon, F. (Eds.), Handbook of Income Distribution. Elsevier, Amsterdam, pp. 87—166.

Cowell, F.A., Jenkins, S.P., 1995. How much inequality can we explain? A methodology and an application to the United States. Econ. J. 105, 421—430.

Creedy, J., Tuckwell, I., 2004. Reweighting household surveys for tax microsimulation modelling: An application to the New Zealand Household Economic Survey. Aust. J. Labour Econ. 7, 71—88.

Davies, J., 2009. Combining microsimulation with CGE and macro modelling for distributional analysis in developing and transition countries. Int. J. Microsimulation 2, 49—65.

Deaton, A., Muellbauer, J., 1980. Economics and Consumer Behavior, Cambridge Books. Cambridge University Press, number 9780521296762.

Deaton, A.S., Paxson, C.H., 1994. Intertemporal choice and inequality. J. Polit. Econ. 102, 437—467.

Deaton, A.S., Paxson, C.H., 1997. The effects of economic and population growth on national savings and inequality. Demography 34, 97—114.

Francisco, F.H.G., Leite, P.G., 2003. Meeting the Millennium Development Goals in Brazil: Can Microeconomic Simulations Help? Economia, Journal of LACEA - LATIN American and Caribbean Economic Association.

Ferreira, F., Pereira, L., Pichetti, P., 2008. Can the distributional impacts of macroeconomic shocks be predicted? A comparison of macro—micro models with historical data for Brazil. In: Bourguignon, F., Bussolo, M., Pereira da Silva, L. (Eds.), The Impact of Macroeconomic Policies on Poverty and Income Distribution — Macro—Micro Evaluation Techniques and Tools. World Bank and Palgrave Macmillan, Washington DC, pp. 119—174.

Ferreira, F.H.G., Ravallion, M., 2003. Global poverty and inequality: A review of the evidence. World Bank Policy Research Working Paper 4623. World Bank, Washington, DC.

Gale, D., Shapely, L.S., 1962. College admissions and the stability of marriage. Am. Math. Mon. 69, 9—15.

Go, D., Kearney, M., Korman, V., Robinson, S., Thierfelder, K., 2010. Wage subsidy and labour market flexibility in South Africa. J. Dev. Stud. 46, 1481—1502.

Gupta, A., Kapur, V., 2000. Microsimulation in Government Policy and Forecasting. Contributions to Economic Analysis 247. Elsevier, Amsterdam.

Harding, A., 1996. Microsimulation and Public Policy. North-Holland, Amsterdam.

Harding, A., 2007. Challenges and opportunities of dynamic microsimulation modeling, paper presented at the First General Conference of the International Microsimulation Association Vienna.

Harding, A., Gupta, A., 2007. Modelling our Future: Population Ageing, Social Security and Taxation. North-Holland, Amsterdam.

Harding, A., Percival, R., Schofield, D., Walker, A., 2000. The lifetime distributional impact of government outlays. Discussion Paper 47. NATSEM, University of Canberra.

Harris, A., O'Harra, J., Page, B., Perese, K., Sabelhaus, J., Simpson, M., Smith, J., 2002. Overview of the Congressional Budget Office Long-Term (CBOLT) policy simulation model. CBO Paper. Congressional Budget Office, Washington, DC.

Heckman, J., Honore, B., 1990. The empirical content of the Roy model. Econometrica 58, 1121—1149.

Heckman, J., Sedlacek, G.L., 1990. Self-selection and the Distribution of Hourly Wages. J. Labor Econ. 8 (1), S329—S363.

Heckman, J., Vytlacil, E., 2008. Econometric evaluation of social programs, part I. Handbook Econom. 8, 4779—5144.

Heckman, J., Sedlacek, G., 1985. Heterogeneity, aggregation, and market wage functions: An empirical model of self-selection in the labor market. J. Polit. Econ. 93, 1077—1125.

Higgins, M., Williamson, J.G., 1999. Explaining Inequality the World Round: Cohort Size, Kuznets Curves, and Openness. Staff Reports 79. Federal Reserve Bank of New York.

Jenkins, S., 1988. The Measurement of Economic Inequality. Papers 170. Australian National University - Department of Economics.

Jenkins, S.P., 1995. Accounting for inequality trends: Decomposition analysis for the UK, 1971—86. Economica 62, 29—63.

Keegan, M.K., 2007. Modelling the workers of tomorrow: the APPSIM dynamic microsimulation model, paper presented at HILDA Survey Research Conference. University of Melbourne. Available at http://

www.melbourneinstitute.com/hilda/conf/conf2007/HILDA%20Conf%20Papers%202007/All%20Papers/Keegan,%20Marcia_final%20paper.pdf.

Kelly, S., King, A., 2001. Australians over the coming 50 years: Providing useful projections. Braz. Electronic J. Econ. 4, 1–23.

Kharas, H., 2010. The emerging global middle class in developing countries. OECD Development Centre Working Paper 285. OECD, Paris.

Klevmarken, N.A., 1994. The impact of demographic changes on the income distribution: Experiments in microsimulation, paper presented at the ESPE Meeting Tilburg.

Klevmarken, N.A., 1996. Behavioural modelling in micro simulation models, paper presented to EUROMOD Meeting Cambridge.

Klevmarken, A., 2008. Dynamic microsimulation for policy analysis: problems and solution. In: Simulating an Ageing Population: A Microsimulation Approach Applied to Sweden. In: Baltagi, B.H., Sadka, E. (Eds.), Contributions to Economic Analysis Series, vol. 285. pp. 31–53. (Chapter 2).

Kuznets, S., 1955. Economic growth and income inequality. Am. Econ. Rev. 45, 1–28.

Lewis, W.A., 1954. Economic development with unlimited supplies of labor. Manchester Sch. Econ. Soc. Stud. 22, 139–181.

Lofgren, H., Robinson, S., El-Said, M., 2003. Poverty and inequality analysis in a general equilibrium framework: The representative household approach. In: Bourguignon, F., Pereira da Silva, L. (Eds.), The Impact of Economic Policies on Poverty and Income Distribution: Evaluation Techniques and Tools. Oxford University Press for the World Bank, New York, pp. 325–337.

Lutz, W., Goujon, A., 2001. The world's changing human capital stock: multi-state population projections by educational attainment. Population and Development Review 27 (2), 323–339.

Medvedev, D., van der Mensbrugghe, D., 2010. Climate change in Latin America: impacts and mitigation policy options. In: de Miguel, C., Lima, J.D., Giordano, P., Guzmán, J., Schuschny, A., Watanuki, M. (Eds.), Modeling Public Policies in Latin America and the Caribbean. Santiago, Chile: United Nations.

Milanovic, B., Yitzhaki, S., 2002. Decomposing world income distribution: Does the world have a middle class? Rev. Income Wealth 48, 155–178.

Mitton, L., Sutherland, H., Weeks, M., 2000. Microsimulation Modelling for Policy Analysis. Cambridge University Press, Cambridge.

Mookherjee, D., Shorrocks, A.F., 1982. A decomposition analysis of the trend in UK income inequality. Economic Journal, Royal Economic Society 92 (368), 886–902, December.

Muller, T., 2004. Evaluating the Economic Effects of Income Security Reforms in Switzerland: An Integrated Microsimulation Computable General Equilibrium Approach. University of Geneva, Geneva.

Mussard, S., Savard, L., 2010. Macro/micro modeling and Gini multidecomposition: An application to the Philippines. J. Income Distribution 19, 51–78.

Neal, D., Rosen, S., 2000. Theories of the Distribution of Labor Earnings. In: Atkinson, A.B., Bourguignon, F. (Eds.), Handbook of Income Distribution. North-Holland, Amsterdam, pp. 379–427.

O'Donoghue, C., 2008. Dynamic Microsimulation: A Methodological Survey, Department of Applied Economics. University of Cambridge, Cambridge.

Orcutt, G.H., 1957. A new type of socio-economic system. Rev. Econ. Stat. 58, 773–797.

Orcutt, G.H., Greenberger, M., Korbel, J., Rivlin, A., 1961. Microanalysis of Socioeconomic Systems: A Simulation Study. Harper & Row, New York.

Orcutt, G.H., Caldwell, S., Wertheimer, R., 1976. Policy Explorations Through Microanalytic Simulation. The Urban Institute, Washington, DC.

Orcutt, G.H., Merz, J., Quinke, H. (Eds.), 1986. Micronalystic Simulation Models to Support Social and Financial Policy. Elsevier, Amsterdam.

Peichl, A., Schaefer, T., 2009. FiFoSiM – an integrated tax benefit microsimulation and CGE model for Germany. Int. J. Microsimulation 2, 1–15.

Perese, K., 2002. Mate matching for microsimulation models. Technical Paper Series 2002-3. Congressional Budget Office, Washington, DC.

Poirier, D.J., 1980. Partial observability in bivariate probit models. J. Econom. 12, 210–217.

Ravallion, M., 2009. The developing world's bulging (but vulnerable) middle class. World Bank Policy Research Working Paper 4816. World Bank, Washington, DC.

Ravallion, M., Lokshin, M., 2008. Winners and losers from trade reform in Morocco. In: Bourguignon, F., Bussolo, M., Periera da Silva, L. (Eds.), The Impact of Macroeconomic Policies on Poverty and Income Distribution. World Bank and Palgrave, Washington, DC, pp. 27–60.

Robilliard, A.S., Robinson, S., 2003. Reconciling household surveys and national accounts data using a cross entropy estimation method. Rev. Income Wealth 49, 395–406.

Robilliard, A.-S., Bourguignon, F., Robinson, S., 2008. Examining the social impact of the Indonesian financial crisis using a macro–micro model, in Brazil. In: Bourguignon, F., Pereira, L., Bussolo, M. (Eds.), The Impact of Macroeconomic Policies on Poverty and Income Distribution. Palgrave and the World Bank, Washington, DC, pp. 93–118.

Roemer, J.E., 1998. Equality of Opportunity, Cambridge, MA: Harvard University Press.

Roemer, J.E., Aaberge, R., Colombino, U., Fritzell, J., Jenkins, S.P., Lefranc, A., M, Ive, P., Marianne, Pommer, E., Ruiz-Castillo, J., 2003. To what extent do fiscal regimes equalize opportunities for income acquisition among citizens? J. Public Econ., 87 (3–4), 539–565.

Rosenstein-Rodan, P., 1943. Problems of industrialization in Southern and Eastern Europe. Econ. J. 53, 202–211.

Roy, A.D., 1951. Some thoughts on the distribution of earnings. Papers, New Series. Oxford Economic 3, 135–146.

Rutherford, T., Tarr, D., 2008. Poverty effects of Russia's WTO accession: Modeling 'real households' with endogenous productivity effects. J. Int. Econ. 75, 131–150.

Scott, A., Evandrou, M., Falkingham, J., Johnson, P., Zaidi, A., 2003. Beyond 2020: lessons from the development of the SAGE dynamic microsimulation model for Britain. ESRCSAGE Discussion Paper 14. London School of Economics, London. available at. www.lse.ac.uk/collections/SAGE.

van der Gaag, N., de Beer, J., Willekens, F., 2005. Combining micro and macro approaches in demographic forecasting, paper prepared for Joint Eurostat–ECE Work Session on Demographic Projections Vienna.

van der Mensbrugghe, D., 2006. The LINKAGE Model Technical Documentation. World Bank, Washington, DC.

Wong, S., Kulmer, V., 2010. Poverty impacts of trade integration with the European Union: Lessons for Ecuador. Project Paper. ECLAC, Santiago.

World Bank, 2005. World Development Report 2006: Equity and Development. Oxford University Press for the World Bank, Oxford.

World Bank, 2007. Global Economic Prospects: Managing the Next Wave of Globalization. World Bank, Washington DC.

Zaidi, A., Rake, K., 2001. Dynamic microsimulation models: A review and some lessons for SAGE. SAGE Discussion Papers 2. London School of Economics, London.