Micro-econometric and Micro-Macro Linked Models: Sequential Macro-Micro Modelling with Behavioral Microsimulations

Jann Lay

1 Introduction

Analyzing the poverty and distributional impact of macro events requires understanding how shocks or policy changes on the macro level affect household income and consumption. It is clear that this poses a formidable task, which of course raises the question of the appropriate methodology to address such questions. This paper presents one possible approach: A sequential methodology that combines a macroeconomic model with a behavioral micro-simulation. We discuss the merits and shortcomings of this approach with a focus on developing country applications with a short to medium run time horizon.[1]

Most analyses of the poverty and distributional impact of macro shocks have turned to Computable General Equilibrium (CGE) models, which typically incorporate different representative household groups with a given within-group income distribution. Yet, recent empirical findings on distributional change indicate that changes within household groups distributions account for an important share of overall distributional change (Bourguignon et al. 2005a, b). At first sight, an obvious solution to this problem seems to increase the number of household groups, or even to incorporate all households from representative household surveys into

---

This chapter is a re-print of: Lay, J. (2010). Sequential macro-micro modelling with behavioural microsimulations. International Journal of Microsimulation, 3(1), 24–34.

[1]Davies (2009) reviews applications linking macro models to micro-simulation models in developing and transition country contexts. His focus is on the applicability of different types of such models to specific questions and contexts. A more technical survey including applications is provided by Colombo (2010) who concentrates on alternative methods to link macro and micro models.

J. Lay (✉)
GIGA German Institute of Global and Area Studies, Hamburg, Germany
e-mail: jann.lay@giga-hamburg.de

the CGE model. Similarly—yet without providing heterogeneous feedback into the CGE model—micro-accounting techniques on the basis of household survey data that apply changes in factor prices at the individual level using household survey-data could be used to increase household heterogeneity. In an assessment of Russia's accession to the World Trade Organization (WTO), Harrison et al. (2000) however find differences in poverty and distributional outcomes between a model with ten representative household groups and a model with 55,000 households to be negligible.

Such evidence does not imply that household heterogeneity would not matter for a true understanding of the poverty and distributional impacts of macroeconomic shocks. It merely shows that even full heterogeneity of households in terms of factor endowments and consumption patterns does not make a difference in a standard CGE model. Microeconomic evidence on the drivers of changes in income distributions however suggests that applied CGE models (including those combined with household-survey-based micro-accounting models) may fail for a different reason: The importance of individual heterogeneity and decisions taken at the individual level for distributional and poverty outcomes; in other words, the importance of "individual behavior." On the labor market, individual decisions include entry into the labor market, falling into unemployment or switching between sectors or occupations. Of course, CGE models can be extended to include for example unemployment and/or endogenous labor supply. Yet, in order to capture the income distribution implications, decisions would have to be taken by "real" individual household members. This implies to introduce individual "fixed effects" and eventually requires the estimation of structural labor market models (Bourguignon et al. 2005a, b) that would need to be integrated in a general equilibrium framework. The estimation of such structural labor market models is by no means a trivial exercise and embedding them into a general equilibrium framework an additional challenge.[2]

This paper presents a less ambitious and more pragmatic approach. The sequential macro-micro approach that links a macroeconomic model, for example an applied CGE model, to a behavioral micro-simulation model has two distinguishing features. First, it is sequential. A counterfactual scenario is generated in the macro (CGE) model. Then, specific poverty and distribution-relevant link variables, for example wages and employment, are passed to a micro-simulation model. Second, the micro-simulation has behavioral components. For the micro-simulation, individual and household decisions are modeled using microeconometric techniques on household and employment survey data. Through the micro-simulation, the combined model hence incorporates individual "fixed effects" into the analysis.

The paper is structured as follows. We first outline some important characteristics of macro models used as part of a sequential model and present a stylized specification of a labor market that produces the link variables for our illustrative macro-micro model. We then provide a simple representation of household income

---

[2]See Blundell and MaCurdy (1999) for a survey of structural labor supply models and Creedy and Duncan (2002) for a discussion of their application in micro-simulation models. See Cogneau (2001) and Cogneau and Robilliard (2001) for attempts to integrate such models into general equilibrium models.

generation that forms the core of the of our prototype behavioral micro-simulation. We describe the simulation mechanics of the micro model. The next section presents two applications of this approach before we assess its strengths, weaknesses, and challenges. The final section concludes.

2 A Stylized Macro-Micro Model with a Behavioral Micro-simulation

2.1 The Macro Model and the Link Variables

The sequential approach presented in this paper requires a macro model that produces changes in distribution and poverty-relevant (aggregate) variables that are passed to a micro-simulation model. These variables, which we label "link variables," are prices and quantities on factor and goods markets. Link variables from factor markets include real wages for different types of labor, returns to land and different types of capital. Factor quantities, for example the sectoral composition of labor, may also be passed from a macro model to a micro-simulation. Finally, goods prices and quantities may operate as link variables. The developing country applications presented in this paper use applied trade-focused CGE models.[3] Yet, other types of macro models with very different foci and features, including other forms of general equilibrium models (real business cycle models, and stochastic dynamic general equilibrium models) and macroeconometric models, may be more suitable in different contexts and for different questions. The illustrative framework presented in the following is general enough to allow the reader to imagine the application of a sequential macro-micro approach using very different models both at the macro and micro level, and different link variables.

If a macro model is built as part of a sequential macro-micro model, its labor market specification is the key component and will have to be compatible with the micro-simulation model that we present below. The following representation of a labor market should be thought of as being embedded, for example, in an applied multisectoral CGE model that distinguishes between formal and informal production sectors. The associated labor markets are assumed to exhibit structural imperfections with different clearing mechanisms for these sectors. For the simplicity of exposition, we abstract from other factors of production and assume that the formal and informal sector produce the same good. Let total employment be fixed and assume that factors are fully employed. Hence, total employment will be the sum of formal and informal employment, $L = L^f + L^{if}$. In a simple neoclassical world with full mobility of labor between formal and informal sectors, wages in the formal and informal sectors, which produce with different technologies ($f_f(L^f), f_{if}(L^{if})$), will be the same. Employment in formal and informal sector, respectively, and hence the

---

[3]See Robinson (1989) for a survey and van der Mensbrugghe (2003) and Lofgren et al. (2002) for standard applied model in the tradition of Dervis et al. (1982).

formal labor share in this economy will be determined by the equation of marginal labor products in formal and informal sectors, as expressed in Eq. (1).

$$\frac{w^f}{p} = \frac{w^{if}}{p} = f'_f(L^f) = f'_{if}(L^{if}) \tag{1}$$

Now assume that different wage setting mechanisms exist: In the formal sector, wages are rigid, for example due to the presence of bargaining by trade unions or efficiency wages. This rigidity can be represented by a "wage curve," as in Eq. (2) where the real formal sector wage becomes a function of the ratio of formal to informal employment.

$$\frac{w^f}{p} = g\left(\frac{L^f}{L - L^f}\right) \tag{2}$$

Without unemployment, the informal sector will now absorb the remaining workforce and the informal sector wage will adjust such that labor demand by the informal sector equals "residual" labor supply. This is depicted in Fig. 1 below where Ec illustrates the competitive equilibrium and wc/p0 the corresponding wage. With WC, the wage curve, the equilibrium wage and employment levels are represented by E_0. The formal sector wage w_0^f/p_0 will now be higher than the informal sector wage w_0^{if}/p_0. Accordingly, formal sector employment L_0^f will be lower than in the competitive case L_c^f.

Real wages and employment in formal and informal sector, respectively, constitute the link variables in our illustrative macro-micro model.

$$\frac{w^f}{p}, \frac{w^{if}}{p}, L^f, L^{if}$$

We now consider a policy experiment that shifts formal labor demand and leads to a new equilibrium in E1. The formal sector wage increases to w1f/p1 and formal employment to L_1^f. The informal sector wage will increase as well from w_0^{if}/p_0 to w_1^{if}/p_1. Hence, the counterfactual values for our link variables[4] that will be passed to the micro-simulation will be

$$\frac{w_1^f}{p_1}, \frac{w_1^{if}}{p_1}, L_1^f, L_1^{if}.$$

---

[4]With real data, the base values for wages and employment levels will typically not be empirically consistent between the macro model, i.e. the SAM, and the micro-simulation model. This inconsistency is "resolved" by passing relative changes from the macro to the micro model. In the simple representation here for example an x percent increase in formal employment and a y(z) percent increase in formal (informal) sector wages.
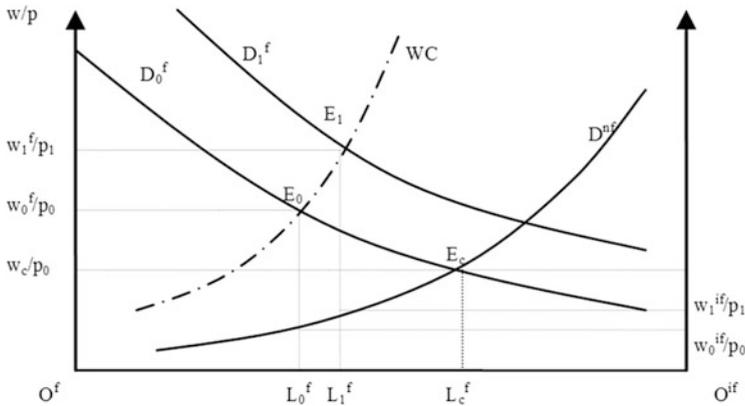
Fig. 1 Formal and informal labor markets. Source: Authors compilation

2.2 A Prototype Income Generation Model: The Micro-simulation

This section describes a prototype micro-simulation model that can be used in combination with the above CGE model to simulate the poverty and distributional impacts of shocks. The basis of the micro-simulation is a household income generation model that needs to be compatible with the above CGE model. For good reasons, we avoid the term consistency here and refer to compatibility instead, as the macro and micro models will not be strictly consistent, neither theoretically nor empirically. We will return to this very important issue in more detail later. The household income generation model is estimated from household survey data with individual-level employment information.

In the micro-simulation, we hence model the household income generation process.[5] This implies that individuals make occupational choices and earn wages or profits accordingly. These labor market incomes plus exogenous other incomes, such as transfers and imputed housing rents, comprise household income. The components of the income generation model are thus an occupational choice and an earnings model. In the choice model, individual agents can choose between wage-employment and self-employment.[6] We thus ignore labor market

---

[5] The following section borrows from Robilliard et al. (2002). A more detailed discussion of a similar labor market specification can be found in Alatas and Bourguignon (2005).

[6] We will use self-employment and informal sector employment interchangeably.

participation choice in this illustrative model. The occupational choice model is assumed to be slightly different for household heads other household members. Once occupational choices are made, earnings are generated accordingly either in the form of wages or as profits for the self-employed. Being self-employed means being part of what might be called a "household-enterprise," in which all self-employed members of a household pool their incomes. The wage-employment market is segmented: the wage setting mechanisms are assumed to differ for skilled and unskilled labor as well as for females and males, which implies that there are four wage labor market segments.

The following set of equations describes the household income generation model. Household income Y_{hh} is earned by k_{hh} members, who are (and remain) active on the labor market [Eq. (3) below]. They are active either in the formal (with $DF_i = 1$, a dummy variable for formal sector employment) or informal sector $((DF_i - 1)(-1) = 1$ if $DF_i = 0)$ and earn the corresponding wages w_i^f, w_i^{if}. In addition, the household receives an exogenous nominal income $\bar{y}_{hh}$, for example transfers or remittances. All these components are real values, i.e. deflated with prices p. In practice, p will be assumed to be one in the initial situation. Per capital income y_{hh} is obtained by dividing household income by household size $Y_{hh}/hsize$ so that $(y_1, y_2, \ldots, y_n)$ denotes the distribution of income when each observation is weighted with household size.

$$Y_{hh} = \frac{1}{p} \sum_{i=1}^{k_{hh}} \left(w_i^f DF_i + w_i^{if}(DF_i - 1)(-1) + \bar{y}_{hh} \right) \qquad (3)$$

Individual occupational choices—between informal and formal activities—can be described by the following functions, which are assumed to be different for household heads (h) and other household members (o). We suppress the individual index here. Equation (4) shows that the household head's probability of being employed in the formal sector is a function of a linear expression with a constant term c^h and personal and household characteristics X^h, which can include for example education, age, and households composition variables.

$$P(DF^h = 1|X^h) = g^h(c^h + X^h\alpha^h) \qquad (4)$$

The choices of other household members are assumed to depend not only on their own individual characteristics X^o, but also on the household head's occupational choice.

$$P(DF^o = 1|X^o, DF^h) = g(c^o + X^o\alpha^o + \gamma^o DF^h) \qquad (5)$$

Equations (6) and (7) express wages w in the formal (f) and informal (if) sectors, respectively, in log-linear form with X, a vector of personal characteristics, and u, a random error.

$$\ln\frac{w^f}{p} = c^f + X\beta^f + u^f \tag{6}$$

$$\ln\frac{w^{if}}{p} = c^{if} + X\beta^{if} + u^{if} \tag{7}$$

The model just described gives the household income as a non-linear function of observed and unobserved individual and household characteristics. This function depends on two sets of parameters, which include the parameters of the wage equations for informal and formal activities and the parameters in the utility associated with different occupational choices for household heads and other family members. The occupational choice equations as well as the corresponding wage equations can be estimated from standard household survey data. Estimating Eqs. (4) and (5) using discrete choice models (with dichotomous choices hence logit or probit models) and (6) and (7) using Ordinary Least Squares (OLS) (or other adequate estimation techniques[7]) yields the following parameter vector:

$$\left(\widehat{c}^h, \widehat{\alpha}^h, \widehat{c}^o, \widehat{\alpha}^o, \gamma^o, \widehat{c}^f, \widehat{\beta}^f, \widehat{c}^{if}, \widehat{\beta}^{if}\right).$$

In addition, we obtain $\widehat{u}^f$ and $\widehat{u}^{if}$ as observed residuals from the wage equations. However, we only observe formal wages for individuals employed in the formal sector. As the micro-simulation will allow individuals to switch between formal and informal activities, we simulate a residual for the non-observed wage, here by a random draw from a normal distribution with the respective (formal or informal) observed variance.[8] We face a similar problem in the latent utility models necessary to estimate Eqs. (4) and (5). In these models, residuals cannot be observed and are hence generated from the distribution underlying the respective model, here either a normal (probit) or logistic (logit) distribution. Residuals have to be drawn consistent with the observed occupational choice, i.e. the utility an observed formal wage earner relates to formal employment has to be higher than the utility associated with informal employment. Statistically, this implies to draw these residuals conditional on the observed choice. These simulated residuals are denoted $u1_i$ and $u0_i$. With

---

[7]Selection bias is a problem in estimating earnings equations in different sectors/occupations (corresponding to different labor market choices) that is difficult to resolve. We return to this point later.

[8]This number does not have to be a random number. It may be reasonable to assume that the observed residual has important informational content with regard to unobserved characteristics, such as ability. A possible alternative to a random draw is then to scale the observed residual in accordance with the observed variances of formal and informal wages, respectively.

ind, an indicator function that assumes a value of 1 (0) if the condition in brackets is (not) fulfilled, we thus have.

$$DF_i^h = ind\big(\hat{c}^h + X_i^h\hat{\alpha}^h + u1_i > u0_i\big) \tag{8}$$

$$DF_i^o = ind\big(\hat{c}^o + X_i^o\hat{\alpha}^o + \gamma^o DF^h + u1_i > u0_i\big) \tag{9}$$

Here DF_i^h and DF_i^o will hence assume their observed values. This implies that the sum of these two dummies—defined either for household heads or other household members—over all individuals will give the total number formal sector employees L_0^f, consistent with the initial value of this link variable from the macro model. This is illustrated in Eq. (10).

$$L_0^f = \sum_{hh} \sum_i^{k_{hh}} \big(DF_i^h \oplus DF_i^o\big) \tag{10}$$

Accordingly, average wages in the formal sector can be expressed as follows.

$$\frac{w_0^f}{p} = \frac{\sum_{hh} \sum_i^{k_{hh}} \left[\big(DF_i^h \oplus DF_i^o\big) \cdot \exp\big(\hat{c}^f + X_i\hat{\beta}^f + \hat{u}_i^f\big)\right]}{L_0^f} \tag{11}$$

Similar expressions can be written down for informal sector employment and the corresponding wage, such that we can replicate all link variables in the initial equilibrium $w_0^f, w_0^{if}, L_0^f, L_0^{if}$. Remember that this replication is based on the observed characteristics of the individuals (all X), unobserved and partially simulated characteristics (all u), and the estimated parameters.

Based on this micro replication of the initial situation, we can now micro-simulate the distributional and poverty implications of the changes in the link variables given by the macro model. In the simulation, the link variables $\frac{w_1^f}{p_1}, \frac{w_1^{if}}{p_1}, L_1^f, L_1^{if}$ will hence be used as target values. This implies that individual earnings and occupational choices have to change such that they reproduce these targets on the aggregate level. There are a number of ways how this can be achieved. Obviously, the required individual changes in occupational choices can be obtained by varying the coefficients or the observed or unobserved individual characteristics. A typical choice in applied micro-simulation models is to vary the constant(s). Hence, the chosen parameters are adjusted and occupational choices change accordingly, until the results of the micro-simulation are consistent, at an aggregate level, with the given aggregates. Formally, the following constraint describes the consistency requirement where c_1^h is the constant in the heads' occupational choice equation that is consistent with L_1^f.

$$L_1^f = \sum_{hh} \sum_i^{k_{hh}} \left[\left(ind \left(c_1^h + X_i^h \widehat{\alpha}^h + u1_i > u0_i \right) \right) \oplus \left(ind \left(\widehat{c}^o + X_i^o \widehat{\alpha}^o + \gamma^o DF_1^h + u1_i > u0_i \right) \right) \right]$$

(12)

Varying only the constant $\widehat{c}^h$ (to c_1^h) implies that we assume that the macro shock only induces household heads to switch occupation. Other household members' occupational choices are only affected through the possible change in the head's occupational choice, i.e. in the case of $DF_1^h \neq DF^h$. As this kind of behavior may not be realistic, we can alternatively assume that the constants of both the heads and other household members vary. However, without an additional restriction, changes in the two constants cannot be uniquely determined. A possible solution is to add a variable Δ to the constant term. In practice—when such equations are solved for real households from a household survey—we will typically be able to find a unique solution for Δ in Eq. (13).

$$L_1^f = \sum_{hh} \sum_i^{k_{hh}} \left[\begin{array}{l} \left(ind \left(\left(\widehat{c}^h + \Delta \right) + X_i^h \widehat{\alpha}^h + u1_i > u0_i \right) \right) \\ \oplus \left(ind \left(\left(\widehat{c}^o + \Delta \right) + X_i^o \widehat{\alpha}^o + \gamma^o DF_1^h + u1_i > u0_i \right) \right) \end{array} \right]$$

(13)

Using either approach to adjust the constant (or both constants) in the occupational choices, will thus enable us to replicate the changes in formal as well as formal employment given by the CGE model. Our very simple income generation model allows us to proceed step-wise. We first solve for changes in occupational choices, and simulate wages in the next step. The reason is that wages do not enter the occupational choices of individuals, as they might in a more complex—or structural—income generation model. However, changes in occupational choices enter the equation for aggregate wages, as the (observed and unobserved) characteristics of the individuals in the respective sectors change. As in the case of occupational choices, we can vary the constants in the respective sectors to equate wages given by the CGE model and those in the micro-simulation.[9] For the formal sector, this requires Eq. (14) to hold with c_1^f, the new formal sector wage equation constant.

$$\frac{w_1^f}{p_1} = \frac{\sum_{hh} \sum_i^{k_{hh}} \left[\left(DF_{1i}^h \otimes DF_{1i}^o \right) \cdot \exp \left(c_1^f + X_i \widehat{\beta}^f + \widehat{u}_i^f \right) \right]}{L_1^f}$$

(14)

The equation for the average informal sector wage can be derived accordingly. The solutions for the constants in the choice Eq. (13) and the wage Eq. (14) can be obtained using numerical solution algorithms, for example Gauss-Newton techniques. With counterfactual occupational choices and corresponding wages DF_1, w_1,

---

[9]Alternatively, we may choose to vary the coefficient for education implying that we expect the macro shock to affect wages through its impact on returns to education.

we can now compute the counterfactual household income Y_{1hh}, as illustrated in Eq. (14). We assume that exogenous transfers $\bar{y}_{hh}$ are constant in nominal terms.

$$Y_{1hh} = \frac{1}{p_1} \sum_{i=1}^{k_{hh}} \left(w_{1i}^f DF_{1i} + w_{1i}^{if}(DF_{1i} - 1)(-1) + \bar{y}_{hh} \right) \ with \ DF_1$$

$$= DF_{1i}^{\,h} \oplus DF_{1i}^{\,o} \tag{15}$$

With constant household size these counterfactual household incomes now yield a counterfactual income distribution that can be described by $(y_{11}, y_{12}, \ldots, y_{1n})$.

3 Applications

The above prototype macro-micro model is intended to provide an introduction into the basic mechanics of a macro-micro model with a behavioral micro-simulation. Which macro model to choose and which transmission channels to highlight eventually depends on the research or policy question and the context, in which it is placed. The two applications that we present in the following are based on recursive-dynamic, trade-focused national CGE models.[10] The first application examines the possible poverty impacts of a Doha round scenario of further multi-lateral trade negotiations for the case of Brazil. The second assesses the poverty and distributional implications of the Bolivian gas shock.

As in the above model and most developing country applications, the focus is on the labor market, as reflected by the link variables that include average wages in different labor market segments, employment levels and the occupational composition of employment. The respective specification of the labor market represents the transmission channels considered to be of particular relevance for the policy and shock under consideration. The Brazilian model focuses on movements between agricultural and non-agricultural sectors, while the Bolivian model concentrates on formal-informal segmentation in the urban labor market. In both applications, the labor market is further segmented along skill levels.

The micro-simulation models used in the subsequent applications share the reduced-form character of the above prototype model. Employment volumes in the respective labor market segments, for example unskilled agricultural employment, and wages are adjusted according to the results from the macro model. These adjustments are not triggered by individual responses to prices, for example relative wages—as they would in a (more) structural labor market model. As above, adjustments are obtained by changing the parameters of the estimated household income generation model.

---

[10]See van der Mensbrugghe (2003) for a technical description of the basic characteristics of the CGE model used in both applications.

3.1 The Poverty Impacts of Trade Liberalization in Brazil

Using this type of sequential model, Bussolo et al. (2006) ex-ante assess the poverty and distributional impacts of different uni- and multilateral trade liberalization scenarios for Brazil.[11] The labor market specification of the CGE model distinguishes between skilled and unskilled labor. While skilled workers are fully mobile across sectors, the labor market for the unskilled is segmented between agriculture and non-agriculture. This dual labor market for unskilled workers is modeled following a simple Harris-Todaro specification where the decision to migrate is a function of expected income in the non-agricultural sectors relative to the expected income in the agricultural sectors.

The micro model is linked to the macro model through changes in the following set of variables: First, changes in agricultural and non-agricultural labor income of unskilled labor; second, changes in labor income of skilled labor; third, changes in the sectoral (agriculture vs. non-agriculture) composition of the unskilled workforce. In addition, the micro-simulation takes into account that unskilled and skilled labor supplies grow at different rates. These rates—also assumed to be exogenous in the CGE model—are derived from past trends of labor supply growth in the respective categories.

In accordance with the structure of the CGE model, the micro model thus simulates the decision to move from agriculture into non-agricultural sectors (or vice versa) only for unskilled workers. This simulation is based on a sectoral mover-stayer model that is estimated for heads and non-heads separately—as in the above prototype model. For this estimation, Bussolo et al. (2006) make use of a distinguishing feature of the PNAD.[12] In contrast to many other household surveys, the PNAD provides information on employment histories, which allows the authors to identify movers between sectors and their characteristics at the time of moving. These characteristics include the type of land right the movers held or whether they were self-employed before they moved out of agriculture. These characteristics enter as explanatory variables into the mover-stayer model. As in the prototype model, the household income generation model is completed by Mincer-type wage equations for unskilled labor in agriculture and non-agriculture as well as for skilled labor. Individual labor incomes are aggregated as described above.

The mover-stayer model can be used to illustrate the behavioral content of the micro-simulation model. For example, Bussolo et al. (2006) find a strong negative influence of own landholdings on the propensity to move. In contrast, higher educational achievements are making individuals more likely to move into non-agricultural employment. In addition, occupational choices of members of the same household are strongly correlated. In the simulation, individuals with no

---

[11]Changes in global prices and trade flows following multilateral liberalization scenarios are derived from global models. For details see Bussolo et al. (2006).

[12]The Pesquisa Nacional por Amostra de Domicílios (PNAD) is a regularly conducted representative household survey

landholdings, better education and—in case of non-household heads—and moving household heads will hence be the first ones to move from agriculture into non-agricultural employment. Which individuals (landless, but better educated) move first, can make a difference in distributional outcomes. The movers' characteristics will determine the composition of those who remain in agriculture (more with own landholdings, but less educated) as well the earning prospects in the non-agricultural sector (ceteris paribus better with better education).

With these components, the micro-simulation involves two steps: First, unskilled labor moves out of agriculture until the new share of unskilled labor in agriculture given by the CGE is reproduced. Second, wages/profits are adjusted according to the CGE results taking into account the sectoral movements of unskilled labor from agriculture into non-agricultural sectors. Adjustments are achieved through the same procedures as in the above prototype model, i.e. the computation of new constants in the choice and wage equations, respectively, using numerical solution algorithms.

The analysis suggests that the economic effects of multilateral liberalization are rather limited for Brazil. Accordingly, poverty would remain largely unaffected by such reforms. In contrast, a full liberalization scenario implies quite substantial welfare gains that are concentrated among some of the poorest groups of the country, in particular those in agriculture. This scenario is also most interesting from a methodological viewpoint, as it highlights the benefits of a behavioral micro-simulation. Under full liberalization, the rural poor benefit more than proportionately, a result driven—on the macro level—by an export boom in agriculture and agricultural processing industries, growing labor demand and associated higher wages. However, following full liberalization, a larger number of workers remain in agriculture compared to the baseline scenario. Given that moving out of agriculture may substantially improve the income situation of a household, one may expect full liberalization to weaken poverty reduction, an expectation supported by the observation that moving households are on average poorer than those remaining in agriculture (for example because they are landless). However, this is not the case, as the gain in agricultural incomes more than compensates the reduced benefits from lower migration flows (for example because they are better educated than those who stay in agriculture).

3.2   The Poverty Impacts of the Bolivian Gas Boom

Lay et al. (2008) examine the poverty effects of the gas boom Bolivia experienced in the late 1990s and early 2000s. Their analysis attempts to disentangle the effects of the resource-boom/bust from other shocks that the Bolivian economy experienced at the same time. The market for unskilled labor is segmented between rural and urban areas. The two segments are linked through rural-urban migration, modeled as in the Brazilian case as a function of the corresponding wage differential. In contrast, skilled labor is assumed to be fully mobile across all production

sectors. Within the urban economy, unskilled workers are mobile between formal and informal sectors, but wage differentials observed in the base period are assumed to persist. These differentials point to systematically lower labor productivity in informal sectors.

Almost as in the prototype model, the macro model is linked to the micro-simulation through the following set of variables: (1) the share of unskilled workers in the formal sector, (2) the share of skilled workers in the formal sector, (3) mean wages for skilled workers, (4) mean wages for unskilled workers, and (5) mean informal profits.[13] Informal profits are understood as mixed income received by self-employed workers. Accordingly, they are calculated as the sum of skilled and unskilled labor income as well as informal capital income.

The two basic components of the income generation model in the Bolivian application are again a model of occupational choices that represents the choice between formal and informal employment as well as earnings functions that correspond to the respective sector of employment. Employment is assumed to be informal if the individual is self-employed/non-remunerated household member and/or works in an enterprise with less than five employees. If individuals happen to be in (or switch to) the formal sector they are assumed to earn a wage, whereas individuals in the informal sector are assumed to be (or become) part of a household enterprise and contribute to the profits earned by this enterprise. The choice between informal and formal activities is modeled separately for household heads, spouses, and other household members. In contrast to the above specifications, the equations of the choice model are interrelated through the head's wage (and choice) that enters the occupational choice model of spouses and other household members. Again, occupational choices are hence assumed to be sequential with the household head deciding first. In line with the CGE model, the micro-simulation distinguishes between unskilled and skilled labor. Separate wage equations for skilled and unskilled labor, respectively, hence describe earnings for individuals employed in the formal sector.[14] The micro-simulation again adjusts the constants to produce counterfactual occupational choices, earnings, and, eventually, household incomes and the corresponding distribution of income.

As in the Brazilian case, the micro-simulation reveals the importance of individual characteristics that determine the sign and the strength of distributional change. Lay et al. (2008) find that—for both unskilled and skilled labor—the very poor are affected most by increasing informality. These results can be rationalized by looking at, first, who moves into informality and, second, the size of the income loss for movers relative to both their initial income and the income losses incurred by other individuals. The size of the income loss depends on individual

---

[13]The authors note on formal profits (Lay et al. 2008): "Although formal profits account for an important share in value added, they are not passed to the micro-simulation for two reasons. First, most formal profits are retained and invested. Second, capital income is likely to be measured very poorly in household surveys. As formal profits increase considerably during the gas boom, we may systematically ignore an inequality-increasing factor."

[14]Rural incomes are taken into account through a simple micro-accounting exercise.

characteristics (as the returns to these characteristics differ between formal and informal activities) and on whether an individual joins an already existing household enterprise or establishes a new one. The estimation, which underlies the micro-simulation, shows that less educated younger (and hence poorer) individuals tend to move into informality first. With regard to the size of the income losses, the estimation results for wages and profit functions indicate that the income loss of moving into informality is higher for more educated individuals, at least in absolute terms, when they move into an existing household enterprise. However, it may also happen that establishing an informal enterprise increases earnings for a skilled individual—conditional of course on other individual characteristics. For an unskilled individual, by contrast, moving into informality will always imply an income loss.

Overall, Lay et al. (2008) find that the gas boom has both unequalising and equalising distributional impacts that tend to offset each other. As net distributional change is limited, growth generated by the boom also reduces poverty and the boom hence does not completely bypass the poorer parts of the Bolivian population. Poverty reduction with little distributional change can be observed despite increasing informality. Additional stylized micro-simulations by Lay et al. (2008) illustrate that lower formal employment can lead to a significant rise in urban poverty and that the very poor are affected most by increasing informality. Yet, considerable overall increases in informal profits compensate this possible negative impact.

3.3   Strengths and Weaknesses

The macro-micro approach presented above and illustrated by the two case studies brings together two strands of literature, macro models, here applied CGE models, on the one hand, and microeconometric poverty and distributional analyses, on the other, which were largely separated from each other. While CGE analyses tend to suffer from being too stylized and not being well informed by micro data, poverty and distributional analyses are often merely descriptive and lack an assessment of the causes of distributional change and the related transmission channels. The sequential approach that combines a CGE model and a behavioral micro-simulation attempts to get the best out of these two "modeling worlds."[15]

A general advantage of a sequential over more complex models is its tractability: While it remains tractable both at the macro and the micro level, it still allows for sufficiently detailed and disaggregated analyses. This is of course more so when the micro model has behavioral components. The case studies above have illustrated the value added of introducing behavior or "individual fixed effects" into the micro-simulation model. In such a micro-simulation, the poverty and distributional impact of policies, as in reality, depend on the characteristics of the households or even individuals.

---

[15]This section borrows heavily from Lay (2007).

However, getting the best of two fairly different modeling worlds comes at the cost of a lack of both theoretical and empirical consistency. Sequentially combining a macro and micro model typically implies the imposition of a number of ad-hoc assumptions that are not satisfying from a theoretical perspective. While the "degree of consistency" between the macro and the micro model however differs between applications, the combined model will lack the theoretical consistency of a general equilibrium model and it is difficult—if not impossible—to resolve all the data discrepancies between national accounts, on the one hand, and household survey data, on the other. Individual responses from estimated relationships may not be conforming to theoretical expectations and the combined model may have leakages—in contrast to the consistent system of flows of an applied CGE model. Theoretically, changes in the behavior of economic agents are driven by relative price changes, whereas the micro-simulation typically only features a reduced-form representation of labor market behavior where prices do not appear as explanatory variables. Empirically, problems arise from the large differences in national accounts and household data, in particular with regard to labor value added, although some authors, e.g. Robilliard et al. (2002), manipulate survey weights to reach "empirical consistency."

Furthermore, quite a few economists may argue that the combination of an applied CGE model with a micro-simulation based on a reduced-form labor market representation may not be a good idea after all. Both types of models suffer from serious shortcomings and combining the two may compound these problems by adding new problems and distracting the researcher from the shortcomings of the "single" models. This is a critique that should be taken seriously. From our own experience in building sequential models, we have become increasingly aware that the additional problems that arise from combining the models, for example in terms of empirical consistency, leave less time for the scrutiny needed to estimate a household income generation model from household survey data or less time to do the sensitivity analyses so often called for in applied CGE analyses (Harrison et al. 1993). We therefore dedicate the following paragraphs to the weaknesses of the single components of a combined macro-micro model without, however, forgetting about their strengths.

The shortcomings of the income generation models are very specific to the respective application and they are discussed at length elsewhere, for example in Bourguignon et al. (2005b). We just want to highlight two typical problems: Selectivity and parameter validity. Estimating earnings equations that correspond to different sectoral or occupational choices entail selection problems. In the presence of unobserved heterogeneity, for example in terms of entrepreneurial ability, it is fairly likely the same unobserved characteristics that make you choose a specific sector also determine the earnings in the respective sector. This selection on unobservables biases the coefficients of an Ordinary Least Squares estimation of the respective equations and would have to be accounted for. It is not trivial to correct for selectivity bias, although the so-called Heckman correction or one of its variants is very common in applied work. To be empirically valid, however, an instrument is needed that explains the sectoral choice, but not earnings in the respective sectors. Such a variable is typically very difficult to find.

There may also be reasons for challenging the validity of the estimated parameters in the household income generation model. Typically, the behavioral equations, e.g. those governing occupational choices, are estimated from cross-sectional data. It is hence assumed that the observed variation in behavior between individuals is used to simulate behavioral change of (other) individuals in time, for example in the Bolivian case study.[16] The Brazilian model relies on employment histories and therefore avoids this problem, but the type of information used reflects to a certain extent short-term behavior. Even if panel data was available, constant parameters would have to be assumed for the simulation period, which apparently becomes an increasingly problematic assumption the longer time horizon of the analysis.

Despite these problems, micro-simulation models based on household income generation models provide a powerful tool to assess the final distributional impact of changes in "distributional drivers," as they reflect the welfare implications of discrete changes in individual behavior, such as labor market entry or sectoral movements. The impact of individual transitions out of agriculture in the Brazil study demonstrates the possible magnitude of these discrete individual changes on household welfare. Finally, it should be stressed that the household income generation models of the type presented in this paper have been shown to do fairly well in reproducing historical patterns of poverty and distributional change (Lay 2007).

The applications from above both use CGE models to trace the transmission channels and quantify the magnitude of the effects of the respective shock. Although widely applied, these models have been criticized for a number of reasons. Analytically, most CGE models rely on the neoclassical framework, although a number of structural characteristics and rigidities are incorporated in most developing country applications. Whether and how structural characteristics and rigidities are taken into account differs between country applications and the research question at hand, as illustrated by the case studies above. Two areas where applied CGE models do not capture the economic realities very well, are the rural and the urban informal sector. It is well known that neoclassical price setting and supply responses in agriculture, is at best a very rough approximation of the reality in most developing countries. In addition, disaggregated input-output data for agriculture are typically not available and agricultural surveys suffer from a lot of problems related to measurement, seasonality, and temporary shocks. Furthermore, the insights from agricultural household models regarding non-separability of production and consumption in rural households (Singh et al. 1986) have not yet entered standard models.[17] More research effort also needs to be dedicated to modeling the informal urban sector. Its heterogeneity in terms of technology,

---

[16]Although this assumption seems to be very restrictive, it can be plausibly made e.g. in the context of occupational choices, which are explained mainly by individual educational attainment, age, and household composition variables.

[17]See Lofgren and Robinson (1999), who integrate a rural household model into a standard CGE model, for an exception.

import penetration, export orientation, and linkages to the formal sector are not reflected in applied CGE models.

However, even with all these improvements, eventually the results of a CGE model will be driven by the assumptions made.[18] Econometricians challenge the empirical relevance of applied CGE models on grounds of the calibration technique based on very restricted functional forms, typically (nested) CES functions. McKitrick (1998) shows the choice of the functional form to make a considerable difference in the results. Yet, in the developing country context, data to estimate these functions is typically not available and the calibration approach overcomes these data restrictions. Furthermore, it is well known that model results are very sensitive to the assumed trade and production elasticities. Harrison et al. (1993) therefore suggest to perform systematic sensitivity analyses and to provide confidence intervals for the results. Such sensitivity analyses, however, are not common in applied work.

Finally, an assessment of the validity of CGE model results also depends on the purpose of the model. If the analysis is expected to provide a precise numerical estimate of the effects of a specific policy change, the above criticisms have to be taken very seriously. In contrast, if CGE models are seen as a rather stylized, yet empirically underpinned, analytical tool to better understand the transmission channels of a shock through counterfactual analysis and approximate their relative importance, the critique is less relevant. This is not to say that the numbers resulting from CGE models are without meaning. They should be taken as the results of a model, given a specific set of assumptions.[19]

4 Conclusions

We have presented and discussed a sequential methodology that combines a macroeconomic CGE model with a behavioral micro-simulation. More specifically, we have shown how micro-simulations based on household income generation models allow the researcher to incorporate individual fixed effects into macro-micro analysis. This is achieved by linking aggregate drivers of poverty and distributional change, such as wages and sectoral employment, to a micro-simulation that is being "forced" to reproduce the changes given by the macro

---

[18]See De Maio et al. (1999) and the reply by Sahn et al. (1999) for an exemplary discussion on specific aspects of CGE models applied to developing countries. These aspects include the macroeconomic and labor market closures as well as the assumption on price setting mechanisms. De Maio et al. (1999) challenge the results of a study by Sahn et al. (1997) on the poverty impacts of structural adjustment in Sub-Saharan Africa as reflecting only the assumptions made in the CGE models, and not reality.

[19]In some CGE applications, including some of those presented in this paper, there is a tendency to treat CGE results as forecasts.

model. We also explain common empirical operationalizations of this link and the micro-simulation procedures commonly used in the literature.

The presented sequential macro-micro approach has been illustrated using two case studies that examine the poverty and distributional impact of macroeconomic shocks, the typical research and policy research to which this kind of model is and should be applied. Examples from these applications have demonstrated the importance of individual heterogeneity in the analysis of these shocks and have underpinned the value added of such methods with behavioral components.

Beyond its ability to capture individual heterogeneity, one of the merits of the approach is its flexibility. However, this flexibility—embodied in a number of fairly ad-hoc assumptions—comes at the cost of theoretical inconsistency. While the macro models rely on consistent theoretical frameworks, the reduced-form models underlying the micro-simulation do not fulfill the requirements, for example in terms of functional forms. Furthermore, empirical inconsistency between national accounts and household survey data that becomes apparent in macro-micro applications is known to be notorious (Round 2003; Robilliard and Robinson 2003).

Finally, we have argued that combining an applied CGE model and a micro-simulation model does not resolve the problems associated to either of those techniques. These problems include a number of typical microeconometric problems that arise from the estimation of income generation models, the basis of the micro-simulation model. Similarly, CGE models suffer from well-known, often-discussed, but less frequently addressed shortcomings. Despite these problems and challenges, the alternative to the proposed models can only be a general equilibrium model that incorporates heterogeneous individuals. As argued in the introduction, researchers are still far from building an applied model based on a micro-based general equilibrium theory. On the route to building such a model, it may be helpful to improve existing macro-micro models through more and better validation exercises. In addition, micro-simulations may also be linked to more different types of general equilibrium models with a more explicit focus on the operation of labor markets.

References

Alatas, V., and F. Bourguignon. 2005. The evolution of income distribution during Indonesia's fast growth, 1980–1996. In *The microeconomics of income distribution dynamics*, ed. F. Bourguignon, F.H.G. Ferreira, and N. Lustig. Washington, DC: The World Bank and Oxford University Press.

Blundell, R., and T. MaCurdy. 1999. Labor supply: A review of alternative approaches. In *Handbook of labor economics*, ed. O. Ashenfelter and D. Card, vol. 3A, 1560–1695. Philadelphia, PA: Elsevier.

Bourguignon, F., A.S. Robilliard, and S. Robinson. 2005a. Representative vs. real households in the macroeconomic modeling of inequality. In *Frontiers in applied general equilibrium modeling*, ed. T.J. Kehoe, T.N. Srinivasan, and J. Whalley. Cambridge: Cambridge University Press.

Bourguignon, F., F.H.G. Ferreira, and N. Lustig. 2005b. A synthesis of the results. In *The microeconomics of income distribution dynamics*, ed. F. Bourguignon, F.H.G. Ferreira, and N. Lustig. Washington, DC: The World Bank and Oxford University Press.

Bussolo, M., J. Lay, and D. van der Mensbrugghe. 2006. *Structural change and poverty reduction in Brazil: The impact of the Doha Round*. World Bank Policy Research Paper 3833. Washington, DC: World Bank.

Cogneau, D. 2001. *Formation du Revenue, Segmentation et Discrimination sur le Marché du Travail d'une Ville en Développement: Antananarivo Fin de Siècle*. DIAL DT/2001/18.

Cogneau, D., and A. S. Robilliard. 2001. G*rowth, distribution and poverty in Madagascar: Learning from a microsimulation model in a general equilibrium framework*. TMD Discussion Paper 61. Washington, DC: International Food Policy Research Institute.

Colombo, G. 2010. Linking CGE and microsimulation models: A comparison of different approaches. *International Journal of Microsimulation* 3 (1): 72–91.

Creedy, J., and A. Duncan. 2002. Behavioural microsimulation with labour supply responses. *Journal of Economic Surveys* 16 (1): 1–38.

Davies, J.B. 2009. Combining microsimulation with CGE and macro modelling for distributional analysis in developing and transition countries. *International Journal of Microsimulation* 2 (1): 49–65.

De Maio, L., F. Stewart, and R. van der Hoeven. 1999. Computable general equilibrium models, adjustment, and the poor in Africa. *World Development* 27 (3): 453–470.

Dervis, K., J. de Melo, and S. Robinson. 1982. *General equilibrium models for development policy*. Cambridge: Cambridge University Press.

Harrison, G.W., R. Jones, L.J. Kimbell, and R. Wigle. 1993. How robust is applied general equilibrium analysis? *Journal of Policy Modeling* 15 (1): 99–115.

Harrison, G., T. Rutherford, and D. Tarr. 2000. Trade liberalization, poverty and efficient equity. *Journal of Development Economics* 71: 97–128.

Lay, J. 2007. Poverty and distributional impact of economic policies and external shocks. Three case studies from Latin America combining macro and micro approaches. In *Goettingen studies in development economics*, ed. H. Sautter and S. Klasen, vol. 18. Frankfurt: Peter Lang.

Lay, J., R. Thiele, and M. Wiebelt. 2008. Resource booms, inequality, and poverty: The case of gas in Bolivia. *Review of Income and Wealth* 54 (3): 407–437.

Lofgren, H., R. Lee Harris, and S. Robinson. 2002. *A standard computable general equilibrium (CGE) model in GAMS*. Microcomputers in Policy Research No. 5. Washington, DC: International Food Policy Research Institute.

Lofgren, H., and S. Robinson. 1999. *To trade or not to trade: Non-separable farm household models in partial and general equilibrium*. TMD Discussion Paper 37. Washington, DC: International Food Policy Research Institute.

McKitrick, R.R. 1998. The econometric critique of computable general equilibrium modeling: The role of functional forms. *Economic Modelling* 15 (4): 543–573.

Robilliard, A.S., F. Bourguignon, and S. Robinson. 2002. *Crisis and income distribution: A micro-macro model for Indonesia*. Washington, DC: World Bank.

Robilliard, A.S., and S. Robinson. 2003. Reconciling household surveys and national accounts data using a cross-entropy estimation method. *Review of Income and Wealth* 49 (3): 395–406.

Robinson, S. 1989. Multisectoral models. In *Handbook of development economics*, ed. H. Chenery and T.N. Srinivasan, 885–947. Philadelphia, PA: Elsevier.

Round, J.I. 2003. Constructing SAMs for development policy analysis: Lessons learned and challenges ahead. *Economic Systems Research* 15 (2): 161–183.

Sahn, D.E., P. Dorosh, and S.D. Younger. 1997. *Structural adjustment reconsidered: Economic policy and poverty in Africa*. Cambridge: Cambridge University Press.

———. 1999. A reply to De Maio, Stewart, and van der Hoeven. *World Development* 27 (3): 471–475.

Singh, I., L. Squire, and J. Strauss, eds. 1986. *Agricultural household models. Extensions, applications, and policy*. Baltimore: The Johns Hopkins University Press.

Van der Mensbrugghe, D. 2003. *Prototype model for real computable general equilibrium model for the state development planning commission*. PR China: Mimeo.

Jann Lay is an economist and heads the research programme on development and globalization at the GIGA German Institute of Global and Area Studies in Hamburg (Germany). As professor in development studies, he also teaches at the Georg-August-University in Goettingen (Germany). He has published articles in various internationally renowned journals and has worked as consultant for a number of development agencies. His current research interests include informal entrepreneurship in developing countries and the poverty and distributional implications of structural change in developing countries, specifically the impact of large-scale land acquisitions.