Discussion Paper No. 08-054

# Linking CGE and Microsimulation Models: A Comparison of Different Approaches

Giulia Colombo

ZEW

Zentrum für Europäische
Wirtschaftsforschung GmbH

Centre for European
Economic Research

Discussion Paper No. 08-054

# Linking CGE and Microsimulation Models: A Comparison of Different Approaches

Giulia Colombo

# Linking CGE and Microsimulation Models: A Comparison of Different Approaches

Giulia Colombo*

Juli 2008

## Abstract

In the paper we describe in detail how to build linked CGE-microsimulation models (using fictitious data) following three main approaches: one in accordance with the fully integrated approach and the other two according to the layered approach – the so-called Top-Down and Top-Down/Bottom-Up approaches. After this, we implement the same policy reform in each of the three models. Results show that all three approaches yield different results especially in terms of income distribution and poverty, although analysed within the same economy and under the same policy simulation. We then analyse in more detail the TD/BU approach as developed by Savard (2003) and, in order to avoid possible deviations due to data inconsistencies, we propose an alternative way of taking into account feedback effects from the micro level of analysis into the CGE model.

**JEL classification:** C68, C15, C35, D31

**Keywords:** CGE models, microsimulation, income distribution.

---

* Centre for European Economic Research (ZEW)
Department of Labour Markets, Human Resources and Social Policy
L7,1 Mannheim
Tel +49 621 1235 367
colombo@zew.de
I would like to thank in particular Stefan Boeters and Nicole Gürtzgen for their precious comments.
Thanks are also due to Marco Missaglia for his constant support.

# Zusammenfassung


Die wirtschaftswissenschaftliche Literatur zu Ungleichheit und Armut verknüpft immer häufiger makroökonomische berechenbare allgemeine Gleichgewichtsmodelle (CGE-Modelle) und Mikrosimulationsmodelle, die auf Individualdaten beruhen. Die Verknüpfung dieser beiden Modellwelten macht es möglich, aus politischen Reformen oder ökonomischen Schocks resultierende Veränderungen der Einkommensverteilung für heterogene Agenten unter Einbeziehung gesamtwirtschaftlicher Rückkopplungseffekte zu analysieren.

Dieser Aufsatz vergleicht die Güte dreier konkurrierender Ansätze zur Verknüpfung der Mikro- und Makroebene. Der erste Ansatz ist ein so genanntes integriertes Modell, bei dem die verfügbaren mikroökonomischen Daten unmittelbar in ein CGE-Modell eingespeist werden. Die beiden anderen Ansätze sind geschichtete Modelle, bei denen makro- und mikroökonomische Modellierung separat erfolgen und die Verknüpfung zwischen beiden Modellwelten durch Übergabe einiger ausgewählter Parameter erfolgt. Der Top-Down-Ansatz verknüpft ein verhaltensbasiertes Mikrosimulationsmodell mit dem CGE-Modell über ein spezifisches Gleichungssystem, das Variablen bzw. Parameter wie Preise und Beschäftigung von der Makro- zur Mikroebene übergibt. Das Top-Down/Bottom-Up-Modell geht weiter, indem zusätzlich Feedback-Effekte von der Mikro- an die Makroebene berücksichtigt werden.

Um die Leistungsfähigkeit dieser Ansätze zu analysieren, wird für jedes Modell dieselbe wirtschaftspolitische Reform anhand der Mikro- und Makrodaten für eine hypothetische Ökonomie simuliert. An diesem Beispiel zeigt sich, dass die drei Modellansätze zu markant unterschiedlichen Ergebnissen führen können. Angesichts der bestehenden Vielfalt der möglichen berechenbaren Mikro-Makro-Modelle sind bei Simulationen daher Robustheitschecks unbedingt erforderlich. Im Einzelnen zeigt sich: Das integrierte Modell liefert tendenziell ungenauere Ergebnisse für die Armuts- und Ungleichheitsmaße als die geschichteten Modelle. Die Resultate im Top-Down/Bottom-Up-Modell reagieren sensibel auf die Variablen, die für die Übergabe von der Mikro- zur Makroebene genutzt werden, und auf Inkonsistenzen der verfügbaren mikro- und makroökonomischen Daten. Wie das Papier zeigt, können die Verzerrungen durch Dateninkonsistenzen im Top-Down/Bottom-Up-Modell verringert werden, wenn die Variablen bzw. Parameter, die von der Mikro- an die Makroebene übergeben werden, in Veränderungen und nicht in Niveaus ausgedrückt werden.

# Non-technical summary

The economic literature on the topic of poverty and inequality has increasingly been linking macroeconomic computable general equilibrium models (CGE models) to microsimulation models based on individual data. Linking these two models allows for an analysis of heterogeneous agents which also takes into account the macroeconomic effects resulting from political reforms or economic shocks.

This paper rates three competing approaches to linking the micro with the macro level of analysis. The first approach is a so-called integrated model which feeds the available microeconomic data directly into a CGE model. The remaining two models are layered models, in which the macro- and microeconomic models are shaped separately and then linked by passing certain selected parameters from one level of analysis to the other. The Top-Down approach links a behavioural microsimulation model with the CGE model via a specific system of equations, which passes variables or parameters (such as price and occupation) from the macro- to the microsimulation model. The Top-Down/Bottom-Up model goes even further and takes into account the feedback effects from the micro- to the macro level of analysis.

In order to analyse the efficiency of these approaches, we simulate an identical economic shock with each model by using the micro- and macro-data for a hypothetical economy. This example shows that the three approaches can lead to distinctly different results. In the light of the existing diversity of possible computable micro-macro-models, simulations are essential for robustness checks. A closer look shows us that the integrated model tends to provide less accurate results for the poverty and inequality measures in comparison to the layered models. The results of the Top-Down/Bottom-Up model are sensitive to the variables used to communicate the feedback effects from the micro- to the macro level of analysis. Moreover, results are also affected by inconsistencies in the available micro- and macroeconomic data. As this paper shows, using variable or parameter changes instead of variable or parameter levels when passing from the micro- to the macro level can reduce deviations caused by data inconsistencies.

# 1. INTRODUCTION

Since the pioneering work by Adelman and Robinson (1978) for South Korea and Lysy and Taylor (1980) for Brazil, many Computable General Equilibrium (CGE) models for developing countries combine a highly disaggregated representation of the economy within a consistent macroeconomic framework and a description of the distribution of income through a small number of representative households (RHs).

However, in order to account for heterogeneity among the main sources of the changes in household income, several "representative households" are necessary. Despite this need for variety, the number of RHs is generally small in these models (usually less than 10).

Usually, the level of disaggregation depends critically on the questions that the model is expected to answer: the household account is broken down into a number of relatively homogeneous household groups to reflect the socioeconomic characteristics of the country or region under consideration. The degree of homogeneity is essential in the design of classifications, and especially in the classification of household groups, where one would like to identify groups that are relatively homogeneous in terms of income sources and levels, and expenditure patterns, and that may be able to reproduce the socioeconomic and structural stratification observed within the society and the economy under study. It is noteworthy anyway that a household classification based on income or expenditure brackets does not satisfy any of these requirements – except perhaps the last one. Indeed, consider for instance the poorest segment of the society (say the bottom decile of the income pyramid): it may include very different household heads, such as a landless agricultural worker and a urban informal sector worker, and policies aimed at improving conditions in the two cases are likely to be very different.

The CGE/RH framework sometimes also explicitly considers that households within a RH group are heterogeneous in a "constant" way. That is, in order to capture within-group inequality, it is assumed that the distribution of relative income within each RH follows an exogenous statistical law[1]. But the assumption that relative incomes are constant within household groups is not reflected in reality. Indeed, empirical analyses conducted on household surveys show that the within-group component of observed changes in income distribution is

[1] For early applications of this type of models, see Adelman and Robinson (1978), and Dervis et al. (1982), who specified lognormal within-group distributions with exogenous variances. More recent examples of this kind of models can be found in Decaluwé et al. (1999a), Colatei and Round (2000) and Agénor et al. (2001).

generally at least as important as the between-group component of these changes[2]. Thus, the RH approach based on this assumption may be misleading in several circumstances, and this is especially true when studying poverty. This argument may be better understood by presenting an example: consider a shock which increases the world price of a specific commodity, say maize, and reduces the world demand for this good. Under the small country assumption (that is, the country is price-taking on the world market), and assuming a demand elasticity with respect to price that is less than one in modulus, a country exporting this good will see a decrease in its exports and a domestic contraction of this sector. After the simulation of the shock with a CGE/RH model, suppose that we find a little change in the mean income of a RH group, say workers in the agricultural sector. In this case, poverty might be increasing by much more than suggested by this drop in income: indeed, in some households there may be individuals that lose their job after the shock, or that encounter more difficulties to diversify their activity or their consumption than others. For these individuals or families, the relative fall in income is necessarily larger than for the whole group, and this fall in their income is not represented by the slight fall in the mean income of the whole group. Suppose moreover that the initial income of these individual was low. Then poverty may be increasing by much more than what predicted by a simple RH model, which is based on the assumption of distribution neutral shocks. So, the RH approach does not capture the effects that a shock or a policy change may have on single individuals or households.

As it is well emphasized in Savard (2003), another significant drawback in linking the intra-group distribution change to a statistical law that is completely exogenous is that no economic behaviour is considered behind this change in within-group distribution[3].

---

[2] After Mookherjee and Shorrocks' (1982) study of UK, there are now other examples of "within/between" decomposition analysis of changes in inequality that indicate that changes in overall inequality are usually due at least as much to changes in within-group inequality as to changes in the between-group component. Among the applications to developing countries, see Ahuja et al. (1997), who applied this decomposition analysis to the case of Thailand, and Ferreira and Litchfield (2001) for Brazil.

[3] The intra-group distribution change is usually linked to a theoretical statistical relationship between average and variance of the lognormal distribution. Savard (2003) also underlines the fact that the average behaviour of a specific group is biased towards the richest in the group. Standard CGE models, indeed, use household groupings that take into account the total income and expenditure of each group and the behavioural parameters which are generally calibrated at the base year. In most of the models these parameters reflect the aggregate and not necessarily the average behaviour. Thus, as the richest of a group are endowed with most of the factors, their behaviour will be dominant in the group. Moreover, keeping in mind that when doing poverty analysis is very impor-

In order to overcome these problems, the recent literature has tried to develop new modelling tools which should be able at the same time to account for heterogeneity and for the possible general equilibrium effects of the policy reform (or the exogenous shock) under study. In view of the fact that most of the available economic models have either a microeconomic or a macroeconomic focus, and they do not address the question adequately, the recent literature has focused on the possibility of combining two different types of models. Most of the economic policies (structural adjustment programs or trade liberalizations, for example) and of the exogenous shocks commonly analyzed for developing countries (such as fluctuations in the world price of raw materials and agricultural exports) are often macroeconomic phenomena (or may have, at least, some structural effects on the economy), while poverty and inequality are mainly microeconomic issues. Thus, an approach that takes into account these important micro-macro linkages, seems to be the right answer to the problem. In particular, some authors have tried to link microsimulation models to CGE models[4], in order to account simultaneously for structural changes, for general equilibrium effects of the economic policies, and for their impacts on households' welfare, income distribution and poverty. The literature that follows this approach is quite flourishing in recent years: there are, among others, the important contributions by Decaluwé et al. (1999a) and (1999b), Cogneau and Robilliard (2001 and 2004), Cockburn (2001), Cogneau (2001), Bourguignon, Robilliard and Robinson (2003b), Boccanfuso *et al.* (2003) and Savard (2003).

The aim of the paper is to give an assessment of recent developments in this field, with a special concern for the different types of linking that are currently used in the literature.

In particular, we will link the microdata from a survey to a CGE model in three different ways: through a full integration of the survey data into a CGE framework, as it is done for instance in Cockburn (2001); by linking a behavioural microsimulation model to a CGE through a set of specific equations, which is the so called Top-Down method, as it is developed in Bourguignon et al. (2003b), and finally through a method which was developed by Savard (2003), also known as Top-Down/Bottom-Up (TD/BU) model.

We will build all the three types of models using the same data from a fictitious economy. After this, by running an identical policy reform in the three models, we will analyse the differ-

tant to consider the behaviour around the poverty line, nothing really demonstrates that the average of aggregated behaviour will be representative of the households around the poverty line.

[4] More generally, this current of the literature develops the use of micro-data drawn from household surveys in the context of a general equilibrium setting, which is usually but not necessarily a CGE model.

ent outcomes deriving from different types of linking. The choice for the use of fictitious data describing a simple economy is made with the aim of being able to understand better the differences that are observed in the results of the models, and to try to "go behind" these differences and look for the causes that generate them. Of course, this is of more difficult realization when using true data of a real and thus more complex economy, which naturally shows more a complex structure in its economic relationships.
Finally, we will analyse in more detail the TD/BU approach as developed by Savard (2003) and propose an alternative way of taking into account feedback effects from the micro level of analysis into the CGE model.

## 2. THE INTEGRATED APPROACH

The main intuition behind this approach is to simply substitute the Representative Household Groups inside a standard CGE model with the real households that are found in the survey[5]. This way, one passes from a model with, for instance, ten representative agents to a model with thousands of agents, thus increasing the computational effort, but leaving substantially unchanged the modelling hypothesis of a standard CGE model. Basically, this approach does not include a true microsimulation module in the modelling framework, but it tries to incorporate the data from the household survey into the CGE model.
The first step to build such a model is to pass from the representative households' data of the survey to population values; to do this, one should weight each variable at the household level with the weights usually given in the survey, thus obtaining population values for each variable.
After this, we need a procedure to reconcile these population data coming from the survey (incomes and expenditures) with the accounts contained in the social accounting matrix (SAM). The literature on data reconciliation offers different alternatives. One may choose to keep fixed the structure of the SAM and adjust the household survey, or otherwise to adjust the SAM in order to meet the totals of the household survey. Another alternative would be that of

[5] The first attempt in this direction was made by Decaluwé et al. (1999b). Among the models following this approach there are the works by Cockburn (2001) for Nepal, by Boccanfuso et al. (2003) for Senegal, and by Cororaton and Cockburn (2005), who studied the case of Philippine economy.

using an intermediate approach. Whatever the method used, however, one necessarily loses the structure of the original data, which is one of the main drawbacks of the integrated approach. Our choice was for the first alternative, and we kept the original composition of households' incomes and expenditures unchanged.

After these changes in the SAM, one encounters the problem of re-balancing it (row totals must be equal to column totals). To do this, we used an appropriate program that minimizes least squares[6].

The CGE model is the one described in section 3.2, except for the fact that we have added an index which refers to households[7].

A thing should be noted at this point: certain types of equations that are commonly included in a behavioural model, such as occupational choice equations, are not easily modelled within standard CGE modelling softwares[8], so that CGE-MS that follow the fully integrated approach are not always able to capture the behavioural responses of the agents to the policy reforms that are implemented. Instead, micro-econometric behavioural modelling provides much more flexibility in terms of the modelling structure used, and is more suitable to describe the complexity of household and individual behaviour, and the way this may be affected by the changes in the macroeconomic framework that are subsequent to a policy reform or an external shock.

The main point here is that with a CGE model like the one used for the integrated approach we are not able to predict which particular individual will enjoy the reduction (or will suffer from the rise) in the employment level on the basis of some characteristics of the individual or of the household that can be observed; this instead can be done through a behavioural microsimulation model.

Indeed, the main feature that differentiates a microsimulation model from a standard CGE framework (not only one with representative agents, but even one with thousands of households from a survey, as we have seen) is that it works at the individual level, selecting those

---

[6] There exist different principles on which SAM-balancing programs can be based, such as the "Row and Sum" or RAS method (see Bacharach, 1971), least squares minimization principles, known also as Stone-Byron methods (see Stone (1977) and Byron (1978)), or the more recent cross-entropy approach proposed by Robinson et al. (2001) and Robilliard and Robinson (2003).

[7] For example, the consumption demand function in Appendix A becomes: $P_q \cdot C_{mq} = \alpha H_{mq} \cdot CBUD_m$, where *m* is now the index for households.

[8] To this regard, see Savard's (2003) discussion about the limits and advantages of the various approaches of linking.

individuals that show the highest probability of changing their labour market status, on the basis of their personal or family characteristics. This fact could bring above significant differences in the results between the two types of models, even after the same policy simulation, as we will see below.

## 3. THE TOP-DOWN APPROACH

We apply now the sequential or Top-Down approach as described in Bourguignon et al. (2003b).

The basic idea is to develop separately a MS model and then to run the simulation on the basis of changes in consumer/producer prices, wages, and sectoral employment levels as predicted by the CGE model. This approach thus uses the two frameworks in a sequential way: first, the policy reform is simulated with the CGE model, and the second step consists of passing the simulated changes in some variables such as prices, wage rates, and employment levels[9] down to the MS module, as illustrated in Figure 1.

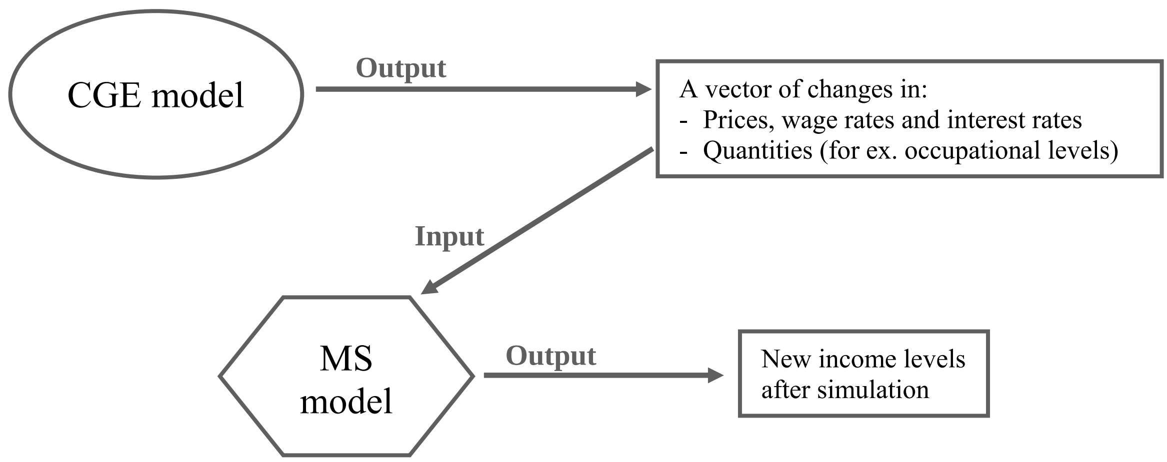


**Figure 1 –** The Top-Down Approach

[9] When the assumption of imperfect labour market is adopted, or when the presence of a formal and an informal sector is predicted, the rationing in the labour market is usually carried out in the macro or CGE model, while the main use of the MS module is to select those households or individuals who will actually be barred out of, or let in, employment, or the formal sector. We will see this in more detail in the simulation section.

### 3.1. The Microsimulation Module

The main role of the microsimulation module in the linked framework is to provide a detailed computation of net incomes at the household level, through a detailed description of the tax-benefit system of the economy, and to estimate individual behavioural responses to the policy change. For instance, through the use of microeconometric equations, we can model behaviours such as labour supply or consumption.

Behavioural Microsimulation (MS) models are developed to capture the possible reactions of the agents to the simulated policies, so that what happens after a reform can be very different from what is predicted by the simple arithmetical computations included in an accounting model.

In this section we will describe in detail a simple behavioural model, following quite closely the discrete labour supply choice model used in Bourguignon *et al.* (2003b). Another description of a similar MS model for labour supply can be found in Bussolo and Lay (2003) with their model for Colombia, and in Hérault (2005), who built a model for the South African economy.

For the building of the model we will use fictitious data describing a very simple economy. In the household survey we have information about some individual characteristics, such as age, sex, level of qualification, education, labour and capital income, the eventual receipt of public transfers, and the activity status. For the sake of simplicity, we have stated that each individual at working age (16-64) can choose between only two alternatives: being a full-time wage worker, or being unoccupied. There are other variables in the survey that are referred to households rather than to individuals, for example the area of residence, the number of household components, the number of adults (over 18 years old) and children (under 18), and so on.

All consumption goods of the economy are grouped in two main categories[10].

We derive income variables referring to households from initial individual data by summing up individual values for each household member; this way, we obtain households' labour and capital incomes, households' public transfers and households' total income:

[10] The focus of our distribution and poverty analysis will be on disposable income, even if an inequality and poverty analysis could also be conducted on expenditure rather than on income levels.

Household $m$'s labour income: $YL_m = \sum_{i=1}^{NC_m} YL_{mi}$

Household $m$'s capital income: $YK_m = \sum_{i=1}^{NC_m} YK_{mi}$

Public transfers to household $m$: $TF_m = \sum_{i=1}^{NC_m} TF_{mi}$

Household $m$'s total income: $Y_m = YL_m + YK_m + TF_m$

where $YL_{mi}$ is labour income of individual $i$ member of household $m$, $YK_{mi}$ his/her capital income, and $TF_{mi}$ are the public transfers he/she receives from government. All these quantities are summed up for each family over all the individuals belonging to the family ($NC_m$ is the number of components of household $m$); then, household $m$'s total income, $Y_m$, is the sum of all incomes received by the family: labour income, capital income, and public transfers.
For the benchmark situation, we assume all initial prices normalized at one.

## *The Model*

The core of the behavioural model is represented by the following two equations:

Regression model for log-wage earnings: $Log(YL_{mi}) = a + b \cdot x_{mi} + c \cdot \lambda_{mi} + v_{mi}$ *(B.1)*

Choice of labour market status: $W_{mi} = Ind[\alpha + \beta \cdot z_{mi} + \gamma \cdot rw_{mi} + \varepsilon_{mi} > 0]$ *(B.2)*

The rest of the MS module is made up by simple arithmetical computations of price indices, incomes, savings and consumption levels. As the parameters entering the following equations (marginal propensity to save $mps_m$, income tax rates $\gamma$, and budget shares $\eta_{mq}$) are constant, this part of the model may be regarded as purely accounting, as it does not contain any possible behavioural response to policy simulations.

Household $m$'s income generation model: $$Y_m = \sum_{i=1}^{NC_m} YL_{mi} \cdot W_{mi} + YK_m + TF_m \qquad (B.3)$$

Household disposable (after tax) income: $$YD_m = (1-\gamma) \cdot Y_m \qquad (B.4)$$

Household specific consumer price index: $$CPI_m = \sum_{q=1}^{2} \eta_{mq} \cdot P_q \qquad (B.5)$$

Real disposable income: $$YDR_m = YD_m / CPI_m \qquad (B.6)$$

Savings: $$S_m = mps_m \cdot YD_m \qquad (B.7)$$

Household consumption budget: $$CEBUD_m = YD_m - S_m \qquad (B.8)$$

Consumption expenditure for commodity $q$: $$CE_{mq} = \eta_{mq} \cdot CEBUD_m \qquad (B.9)$$

Consumption level of commodity $q$: $$C_{mq} = \frac{CE_{mq}}{P_q} \qquad (B.10)$$

Household $m$'s capital income: $$YK_m = PK \cdot KS_m \qquad (B.11)$$

**Description of the subscripts:**

| | | | |
|---|---|---|---|
| $m$ | Households | $m = 1, 2, \ldots, 24$ | |
| $i$ | Individuals belonging to household $m$ | $i = 1, \ldots, NC_m$ | $NC_m$: number of components of household $m$ |
| $q$ | Goods | $q = 1,2$ | |

The *first equation* of the model, *(B.1)*, computes the logarithm of labour income (wage) of member $i$ of household $m$ as a linear function of his/her personal characteristics (vector $x_{mi}$ includes the logarithm of age, sex, skill level and educational attainment) and of $\lambda_{mi}$, which represents the inverse Mills ratio estimated for the selection model (for more details on the estimation process see below). The residual term $v_{mi}$ describes the effects of unobserved components on wage earnings.

The *second equation* represents the choice of the labour status made by household members[11]. Each individual at working age has to choose between two alternatives: being a wage worker,

[11] In the literature this kind of equation is known as occupational choice model, or selection model (and also discrete choice model of labour supply). However, it must be specified that this equation is not really intended to explain the individual *choice* between being occupied or unemployed, but rather it tries to find out which characteristics strengthen the *probability* of being in one condition rather than in the other one for each individual, as it is described in more detail in the estimation section below.

or being inactive. The variable $W_{mi}$ is a dichotomic variable taking value one if individual *i* of household *m* decides to be a wage worker, and zero otherwise. The choice is made by each individual according to some criterion, the value of which is specific to the alternative, and the alternative with the highest criterion value is selected. A natural economic interpretation of this criterion value is utility: each individual chooses the alternative with the highest associated utility. Indeed, we will estimate the selection model using a binomial logit specification, which assigns each individual to the alternative with the highest associated probability. In our model we have arbitrarily set to zero the utility of being inactive. Function "*Ind*" is an indicator function taking value one if the condition is verified, and zero otherwise. Vector $z_{mi}$ of explanatory variables includes some personal characteristics of individual *i* of household *m*, that is: age, sex, skill and educational level, the area of residence and the number of children under 6 living in the household. Variable $rw_{mi}$ is the logarithm of real labour income. The equation is defined only for individuals at working age.

The *third equation* is an accounting identity that defines total household income, $Y_m$, as the sum of the wage income of its members $YL_{mi}$, of the exogenous household capital income $YK_m$, and of the total amount of public transfers received by household *m*, $TF_m$. In this equation, variable $W_{mi}$ stands for a dummy variable that takes value one if member *i* is a wage worker and zero otherwise.

The *fourth equation* computes household disposable (after tax) income by applying income tax rates according to the rule reported in Table 1. In order to simplify computations, we have assumed that in this economy direct income taxes are imposed on households' total income $Y_m$, and not on individual incomes.

**Table 1 –** Direct Income Tax Rates

| *Income brackets:* | *Tax rate* |
|---|---|
| Up to 10,000 | 0% |
| Up to 15,000 | 15% |
| Up to 26,000 | 24% |
| Up to 70,000 | 32% |
| Over 70,000 | 39% |

*Equation (B.5)* computes an household specific consumer price index through the consumption shares $\eta_{mq}$. Real disposable income is then obtained by dividing households' disposable income by this index (*equation (B.6)*).

Then, to find out household $m$'s savings level, equation (B.7) multiplies this disposable income by the marginal propensity to save of each household, $mps_m$. The assumption underlying this equation is that household savings behaviour is unvarying, as the savings level is a fixed fraction of household disposable income. Then, subtracting savings from disposable income one obtains the budget that each household spends for consumption (equation (B.8)), which is spent on the two goods of the model according to the budget shares $\eta_{mq}$ by equation *(B.9)*. Again, the assumption in this equation is that consumption behaviour is not flexible, that is, households spend a constant fraction of their consumption budget for each of the two goods.

To get the values of these exogenous parameters (marginal propensity to save $mps_m$ and budget shares $\eta_{mq}$), we use the initial data from the survey in the following way:

Household $m$'s marginal propensity to save: $$mps_m = \frac{S_m}{YD_m}$$

Household $m$'s consumption budget shares: $$\eta_{mq} = \frac{CE_{mq}}{CEBUD_m}$$

*Equation (B.10)* derives then the consumption levels for each household by dividing the expenditure for each good by its price.

Finally, income from capital is obtained by multiplying capital endowment of each family, $KS_m$, by the return to capital, *PK* (equation (B.11)).

The initial values of the variables $C_{mq}$ and $KS_m$ (consumption levels and capital endowments, respectively) are derived from the initial data of the survey by making use of the assumption that in the benchmark situation all prices and returns are equal to one:

Household $m$'s consumption level of commodity $q$: $$C_{mq} = CE_{mq} \quad (B.12)$$

Household $m$'s capital endowment: $$KS_m = YK_m \quad (B.13)$$

Moreover, we assume that public transfers paid to households and household capital endowments are exogenously given. They are fixed at the level reported in the survey, for public transfers, and at the level as computed in equation *(B.13)*, for capital endowment, respectively.

### *Estimation of the Model*

The only two equations in the MS module that need to be estimated are equations *(B.1)* and *(B.2)*.

The former, which expresses the logarithm of wage earnings as a linear function of some individual characteristics and of $\lambda_{mi}$, the inverse Mills ratio, was estimated using a Heckman two-step model (see Heckman (1976) and (1979)). We follow this approach to correct for the selection bias which is implicit in a wage regression, that is, the fact that we observe a positive wage only for those individuals that are actually employed at the moment of the survey.

The results of the estimation are reported in Table 2 below. The estimation was conducted on the sub-sample of individuals at working age (16-64).

**Table 2 –** Heckman selection model, two-step estimates

| *Dependent variable: logarithm of wage* | | | | |
|---|---|---|---|---|
| | *Coefficient* | *Std. Error* | *z* | *P>\|z\|* |
| constant | 7.032117 | 0.3145104 | 22.36 | 0.000 |
| ln(age) | 0.697818 | 0.0833084 | 8.38 | 0.000 |
| sex | -0.466210 | 0.1018222 | -4.58 | 0.000 |
| qualification | 0.396613 | 0.0771516 | 5.14 | 0.000 |
| education | 0.525011 | 0.0871646 | 6.02 | 0.000 |
| Mills ratio | 0.216005 | 0.1473164 | 1.47 | 0.143 |
| *Selection* | | | | |
| ln(age) | 0.338583 | 0.0807227 | 4.19 | 0.000 |
| sex | -1.549158 | 0.2802896 | -5.53 | 0.000 |
| qualification | 1.020388 | 0.2728658 | 3.74 | 0.000 |
| children under 6 | 0.168214 | 0.2368365 | 0.71 | 0.478 |
| region | -0.751549 | 0.2980307 | -2.52 | 0.012 |
| rho | 0.762760 | | | |
| sigma | 0.283187 | | | |

The interpretation of the coefficients for the wage equation thus follows that of a simple linear regression. As we can observe in Table 2, age, schooling and skill level have a positive effect on the wage, while being a woman shows a negative effect.

It is important to say that the aim of the wage equation within the model is that of obtaining an efficient estimate for an eventual wage income only for those individuals that are observed to be inactive in the survey, in the case that, after a policy reform, one or more of them will change their labour market status and become wage workers. In this case, through these estimates, we will be able to assign an estimated wage to the individual that has changed his/her labour market status after the simulation run.

For all the other individuals that are observed to receive a wage in the survey, we use instead the observed wage level and not the estimated one.

Parameters of equation *(B.2)* were obtained through the estimation of a binomial logit model, assuming that the residual terms $\varepsilon_i$ are distributed according to the Extreme Value Distribution – Type I[12]. The estimation was conducted on the sub-sample of individuals at working age (16-64).

Our explanatory variables include individual characteristics such as the logarithm of predicted real wage, sex, skill and education level, the region of residence and a variable accounting for the presence or not of children under 6 years old in the household. The model is estimated by Maximum Likelihood. Results are presented in Table 3.

**Table 3 –** Binary logit model for labour status' choice

| *Dependent Variable: Activity Status* | | | | |
|---|---|---|---|---|
| | *Coefficient* | *Std. Error* | *z-Statistic* | *Prob.* |
| ln(real wage) | 0.197215 | 0.046458 | 4.245037 | 0.0000 |
| sex | -1.894812 | 0.407759 | -4.646894 | 0.0000 |
| qualification | 1.440805 | 0.425709 | 3.384482 | 0.0007 |
| region | -0.718504 | 0.329501 | -2.180586 | 0.0292 |
| children under 6 | 0.269124 | 0.297251 | 0.905378 | 0.3653 |
| education | -0.763275 | 0.671696 | -1.136341 | 0.2558 |
| Mean dependent var | 0.664706 | | S.D. dependent var | 0.473488 |
| S.E. of regression | 0.376673 | | Akaike info criterion | 0.901535 |
| Sum squared resid | 23.26880 | | Schwarz criterion | 1.012210 |
| Log likelihood | -70.63049 | | Hannan-Quinn criter. | 0.946446 |
| Avg. Log likelihood | -0.415473 | | | |

[12] The Extreme Value distribution (Type I) is also known as Gumbel (from the name of the statistician who first studied it) or double exponential distribution, and it is a special case of the Fisher-Tippett distribution. It can take two forms: one is based on the smallest extreme and the other on the largest. We will focus on the latter, which is the one of interest for us. The standard Gumbel distribution function (maximum) has the following probability and cumulative density functions, respectively:

pdf: $f(x) = \exp\left(-x - e^{-x}\right)$

CDF: $F(x) = \exp\left(-e^{-x}\right)$.

A binomial model states that the probability of observing the dependent variable assuming value one, given the explanatory variables ($OCS_{mi} = 1|Z_{mi}$), is equal to the cumulative distribution function of $\varepsilon_i$ (the Extreme Value Type I distribution in our case), evaluated at $\beta\cdot Z_{mi}$, that is:

$$\Pr\left[OCS_{mi} = 1 \mid Z_{mi}\right] = F(\beta \cdot Z_{mi}) = \exp\left(-e^{-\beta \cdot Z_{mi}}\right). \tag{B.14}$$

The effects that the explanatory variables have on the dependent binomial variable are not linear, because they get channelled through a cumulative distribution function. Thus, by observing the values and signs of the estimated coefficients, we can say something about the effect that explanatory variables have on the probability that the dependent binomial variable takes value one (wage worker), relatively to the probability that it takes value zero, but not in a linear way.

For instance, expected real wage and qualification seem to influence in a positive way the probability that the dependent variable takes value one (the more qualified the individual is, the higher is the probability for him/her to be employed), as well as the presence of children under 6 does, which is the opposite of what was expected, but anyway this result is not significant. Moreover, for men the probability of being employed is higher than for women, as the variable *SEX*, which takes value zero for men and one for women, shows a negative coefficient. The same can be said about the region of residence: people living in the first region have a higher probability of being employed than people living in the second one. The variable referring to education, instead, seems to have a negative influence on the probability of being employed, which is the opposite of what we expected, and anyway it is not highly significant.

However, with the estimated coefficients we cannot perfectly predict the true labour market statuses that are actually observed in the survey. Thus, following the procedure described in Duncan and Weeks (1998), we drew a set of error terms $\varepsilon_i$ for each individual from the extreme value distribution, in order to obtain an estimate that is consistent with the observed activity or inactivity choices. From these drawn values, we select 100 error terms for each individual, in such a way that, when adding it to the deterministic part of the model, it perfectly predicts the activity status that is observed in the survey. In other words, the residual term for an individual that is observed to be a wage earner in the survey should be such that:

$$\hat{\alpha} + \hat{\beta}_1 \cdot Log(RW_{mi}) + \hat{\beta}_2 \cdot SEX_{mi} + \hat{\beta}_3 \cdot Q_{mi} + \hat{\beta}_4 \cdot AREA_m + \hat{\beta}_5 \cdot CH6_{mi} + \hat{\beta}_6 \cdot SCH_{mi} + \varepsilon_{mi} > 0,$$

while, for an individual that is observed to be inactive in the survey, the same inequality should be of opposite sign ($\leq$).

After a policy change, only the deterministic part of the model is recomputed. Then, by adding the random error terms previously drawn to the recomputed deterministic component, a probability distribution over the two alternatives (being a wage worker or being inactive) is generated for each individual. This implies that the model does not assign every individual from the sample to one particular choice, but it gives the individual probabilities of being in one condition rather than in the other. This way, the model does not identify a particular choice for each individual after the policy change, but generates a probability distribution over the different alternatives[13].

## 3.2. The CGE Model

The CGE model for the fictitious economy is characterized by a representative household who maximizes a Cobb-Douglas utility function with three arguments: leisure and two consumption goods. These commodities are also used as inputs, together with capital and labour, in the production process, which is operated by two firms following a Leontief technology in the aggregation of value added and the intermediate composite good, a Constant Elasticity of Substitution (CES) function for assembling capital and labour into value added, and a Leontief function in the aggregation of intermediate goods. Both factors of production, capital and labour, are mobile among sectors. The capital endowment is exogenously fixed, while labour supply is endogenously determined through household's utility maximization (subject to fixed time endowment). The wage elasticity of labour supply is estimated from the household survey, in order to have consistency in labour supply behaviour between the two models. Investments are savings-driven, while government maximizes a Cobb-Douglas utility function to buy consumption goods and uses labour and capital. The public sector also raises taxes on household's income and tariffs on imported goods, while it pays transfers to the representative household. For the foreign sector we have adopted the Armington assumption of constant elasticity of substitution for the formation of the composite good (domestic production delivered to domestic market plus imports) which is sold on the domestic market. Domestic production is partially delivered to the domestic market and partially exported, according to a Constant Elasticity of Transformation (CET) function. The small country hypothesis is assumed (the economy is price taker in the world market).

[13] This procedure is also described in Creedy and Kalb (2005). See also Creedy et al. (2002).

**Table 4 –** SAM of the Economy

| | $C_1$ | $C_2$ | $S_1$ | $S_2$ | *K* | *L* | *H* | *G* | *SI* | *RoW* | *Total* |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $C_1$ | | | 57.5 | 15.5 | | | 95.2 | 61.2 | 30.3 | 23.5 | 283.3 |
| $C_2$ | | | 17.1 | 23.5 | | | 312.8 | 48.5 | 14.2 | 76.5 | 492.5 |
| $S_1$ | 283.3 | | | | | | | | | | 283.3 |
| $S_2$ | | 492.5 | | | | | | | | | 492.5 |
| *K* | | | 72.2 | 23.0 | | | | 13.1 | | | 108.3 |
| *L* | | | 83.2 | 353.8 | | | | 116.4 | | | 553.4 |
| *H* | | | | | 108.3 | 553.4 | | 39.8 | | | 701.5 |
| *G* | | | 12.3 | 17.7 | | | 249.0 | | | | 279.0 |
| *SI* | | | | | | | 44.5 | | | | 44.5 |
| *RoW* | | | 41.0 | 59.0 | | | | | | | 100.0 |
| *Total* | 283.3 | 492.5 | 283.3 | 492.5 | 108.3 | 553.4 | 701.5 | 269.9 | 44.5 | 100.0 | |

$C_q$: consumption of good *q*; $S_q$: sector *q*; ***K***: capital account; ***L***: labour account; ***H***: representative household account; ***G***: public sector; ***SI***: savings-investments account, ***RoW***: Rest of the World account.

**Table 5 –** Values of Parameters for CGE Model

| | Sector 1 | Sector 2 |
|---|---|---|
| Elasticity of substitution in production function (aggregation of capital and labour) | 0.7 | 0.5 |
| Elasticity of substitution for Armington composite good | 0.7 | 1.2 |
| Elasticity of transformation for exports and domestic production delivered to the domestic market | -2.0 | -3.0 |
| Initial tariff rates on imports | 0.3 | 0.3 |
| Initial time endowment | 656.69 | |
| Wage elasticity of labour supply (estimated from the household survey) | -0.18665 | |

In the model there are in total 49 variables and 41 equations, which, with the 8 exogenous variables (capital endowment, *KS*, time endowment, *TS*, public transfers, *TF*, the four world prices $PWE_q$ and $PWM_q$, and the numeraire, *PC*), fully determine the model and allow for satisfaction of Walras' law (we have a redundant equation).

The calibration of the parameters of the CGE model is done on the basis of a Social Accounting Matrix (SAM) for the economy, in such a way that the benchmark situation is consistent with that of the microsimulation module (for instance, in the benchmark of the two models we have the same average income tax rate, the same average marginal propensity to save, the same budget shares for consumption of the two goods, and so on).

The SAM for the economy under study and the initial values of some other variables are reported in Tables 4 and 5, while the equations of the model can be found in Appendix. The data in the SAM are in millions of the monetary unit we have used for the survey.

### 3.3. Linking the Models

The basic difficulty of this approach is to ensure consistency between the micro and macro levels of analysis. For this reason, one may introduce a system of equations to ensure the achievement of consistency between the two models[14]. In practice, this consists in imposing the macro results obtained with the CGE model onto the microeconomic level of analysis. In particular:

1) changes in the commodity prices, $P_q$, must be equal to those resulting from the CGE model;
2) changes in average earnings with respect to the benchmark in the micro-simulation must be equal to changes in the wage rate obtained with the CGE model;
3) changes in the return to capital of the micro-simulation module must be equal to the same changes observed after the simulation run in the CGE model;
4) changes in the number of wage workers in the micro-simulation model must match those observed in the CGE model.

For our model, these consistency conditions translate into the following set of constraints, which could be called linking equations:

Consumption levels:
$$C_q = \frac{CE_q}{\left(1+\Delta P_q^{CGE}\right)} \qquad (M.1)$$

Logarithm of wage earnings:
$$Log\left(YL_{mi}\right) = Log\left[\hat{YL}_{mi} \cdot \left(1+\Delta PL^{CGE}\right)\right] \qquad (M.2)$$

Capital income:
$$YK_m = KS_m \cdot \left(1+\Delta PK^{CGE}\right) \qquad (M.3)$$

Employment level:
$$\frac{\sum_{m=1}^{24}\sum_{i=1}^{NC_m}\hat{W}_{mi}}{\sum_{m=1}^{24}\sum_{i=1}^{NC_m}WA_{mi}} \cdot 100 = \Delta EMP^{CGE} \qquad (M.4)$$

---

[14] This way, what happens in the MS module can be made consistent with the CGE modelling by adjusting parameters in the MS model, but, from a theoretical point of view, it would be more satisfying to obtain consistency by modelling behaviour identically in the two models.

The variables with no superscripts are those coming from the microsimulation module; those with the ^ notation correspond to the ones that have been estimated: in particular, $Log(\hat{YL}_{mi})$ is the wage level resulting from the regression model for individual *i*, member of household *m*, while $\hat{W}_{mi}$ is the labour market status of individual *i* of household *m* deriving from the estimation of the binomial choice model.

$\Delta P_q^{CGE}$, $\Delta PL^{CGE}$ and $\Delta PK^{CGE}$ indicate, respectively, the change in the prices of goods, the change in the wage rate and in the return to capital deriving from the simulation run of the CGE model, while parameter $\Delta EMP^{CGE}$ is the employment level percentage change from the CGE.

$WA_{mi}$ is a dummy variable taking value one if individual *i* of household *m* is at working age (16-64), and zero otherwise. From equation *(M.4)*, the number of employed over the total number of individuals at working age resulting from the MS model must be equal to the change in the employment level observed after the CGE run. This implies that the CGE model determines the employment level of the economy after the simulation, and that the MS model selects which individuals among the inactive persons have the highest probability of becoming employed (if the employment level is increased from the CGE simulation result), or either who, among the wage workers, has the lowest probability of being employed after the policy change (if the employment level is decreased)[15].

One possible way of imposing the equality between the two sets of parameters of system of equations *(M)* is through a change in the parameters of the selection and regression models. Following Bourguignon et al. (2003b), we restrict this change in the parameters to a change in the intercept of the two functions *(B.1)* and *(B.2)*. The justification for this choice is that it implies *neutrality* of the changes, that is, changing the intercepts *a* of equations *(B.1)* just shifts proportionally the estimated wages of all individuals, without causing any change in the ranking between one individual and the other. The same applies for the activity status choice equation: we choose to change the intercept *α* of equation *(B.2)*, and this will shift proportionally all the individual probabilities of being a wage worker, without changing their relative positions in the probability distribution, only to let some more individuals to become employed (or some less if the employment rate of the CGE model is decreased), irrespectively of their personal characteristics. This change in the intercept will be of the amount that is necessary to reach the number of wage workers resulting from the CGE model. Thus, this choice preserves

[15] And, in this case, his/her new wage level will be determined by the regression model of wage earnings.

the ranking of individuals according to their *ex-ante* probability of being employed, which was previously determined by the estimation of the binomial model. For this reason the change in the intercept parameter satisfies this neutrality property.

# 4. THE TOP-DOWN/BOTTOM-UP APPROACH

This approach was developed by Savard (2003). It allows overcoming the problem of the lack of consistency between the micro and macro levels of the Top-Down approach by introducing a bi-directional link between the two models: this is the reason why this approach is also called “Top-Down/Bottom-Up”. According to this method, indeed, aggregate results from the MS model (such as consumption levels and/or labour supply) are incorporated into the CGE model, and a loop is used to run both models iteratively until the two produce convergent results.

The value added of this approach is that it takes into account the feedback effects that come from the micro level of analysis, which are instead completely disregarded by the Top-Down model. The basic assumption behind this approach is that the microeconomic effects provided by the MS model run do not correspond to the aggregate behaviours of the representative households used in the CGE model, and that it is thus necessary to take these effects back into the CGE model to fully account for the effects of a simulated policy. A stylized scheme of the way in which this approach works can be observed in Figure 2.

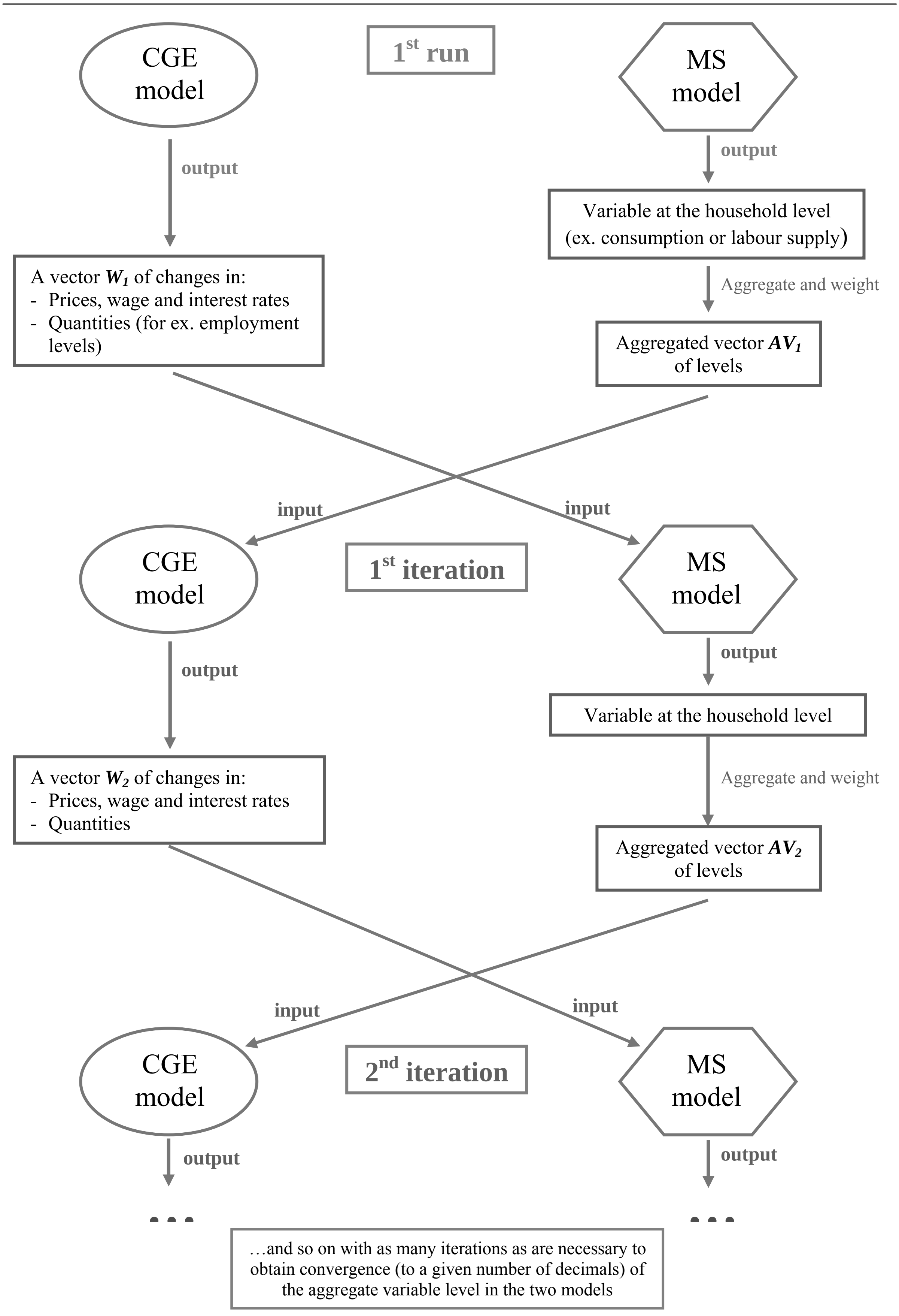


**Figure 2 –** The Top-Down/Bottom-Up Approach

The bilateral communication between the two levels of analysis is achieved through a set of vectors of changes, as in the Top-Down approach: from the macro to the micro level of analysis the communication is guaranteed by the changes in the price, wage and return vector and in the employment levels, as before, while from the micro to the macro level the communication we apply two different strategies: in one version, we will use as input for the CGE model a vector of changes in the aggregate consumption and in the labour supply levels from the MS model[16]; in another version of the same model, only the change in the labour supply level which results from the MS model will be used as input for the CGE model[17]. The process is then iterated as many times as it is necessary to come to a convergent point, that is, when convergence (at a certain number of decimals) is obtained in the aggregate variable levels of the two models.

# 5. SIMULATION

We will now run a policy simulation with each of the three models. The simulation will be an exogenous shock on the world price level of the good exported by sector 2, which is the labour intensive sector in our stylized economy. The world price of good 2 is reduced of 64 % from its initial value.

[16] The choice for consumption and labour supply as communicating variables is made following Savard (2003). However, as both consumption and labour supply are not exogenous in the CGE model, we have to change some of the initial hypothesis of the model. First, we remove the equations determining consumption demand by the representative household (equation *C.1* in Appendix A), substituting them with the following single equation: $CBUD = \sum_{i=1}^{2} P_i \cdot C_i$ . In the initial hypothesis (endogenous consumption) we had 2 endogenous variables ($C_i$) and 2 equations. Now we have 2 exogenous variables and one equation. As we need to insure the balancing of the household's budget constraint, a variable needs now to be endogenized in the following equation: $CBUD = (1 - mps) \cdot (1 - ty) \cdot (PK \cdot KS + PL \cdot LS + PC \cdot TF)$. Following Savard, we choose to endogenize the marginal propensity to save, *mps*, which is now a variable that changes in order to satisfy the budget constraint.

In addition, we introduce an exogenous level of labour supply into the CGE model, and just leave out the equation that determines the demand for leisure (equation *C.2* in Appendix A). This way, equation *C.3* will now yield the demand for leisure as the time remaining after having supplied an exogenous level of labour.

[17] In this case, we only introduce an exogenous level of labour supply into the CGE model, just leaving out the equation that determines the demand for leisure (equation *C.2* in Appendix A).

The simulation results for the most relevant macroeconomic variables are reported in percentage changes in Tables 6 and 7. In the table, also the two different strategies adopted for the TD/BU approach are taken into account, so that we will compare the results coming from the introduction into the CGE model of, respectively, the consumption level and the labour supply coming from the microsimulation module, and only the labour supply.

**Table 6 –** Simulation Results: Percentage Changes (CGE Model)

| | *Integrated Approach* | *Top-Down Approach* | *TD/BU Approach (Cons. and LS)* | *TD/BU Approach (Labour Supply)* |
|---|---|---|---|---|
| *Government Surplus* | 0.00 | 0.00 | 0.00 | 0.00 |
| *Wage Rate* | -14.87 | -14.67 | -14.42 | -14. 64 |
| *Capital return* | 19.70 | 19.30 | 17.91 | 19.13 |
| *Consumer Price Index (num.)* | 0.00 | 0.00 | 0.00 | 0.00 |
| *Exchange rate* | 53.83 | 53.76 | 53.83 | 53.70 |
| *Labour Supply* | -1.00 | -1.18 | -1.32 | -1.32 |
| *Government Use of Labour* | 4.82 | 4.23 | 3.72 | 4.06 |
| *Government Use of Capital* | -25.45 | -25.45 | -24.72 | -25.43 |
| *Income** | -9.50 | -9.39 | -9.50 | -9.48 |
| *Disposable Income** | -9.50 | -9.39 | -9.50 | -9.48 |
| *Consumption Expenditure** | -9.50 | -9.39 | -7.90 | -9.48 |
| *Marginal Propensity to Save* | 0.00 | 0.00 | -16.22 | 0.00 |
| *Savings** | -9.28 | -9.39 | -24.18 | -9.48 |
| *Tax Revenues* | -9.28 | -9.48 | -9.63 | -9.58 |

[*] For the integrated model, these changes are computed as average percentage changes across households.

**Table 7 –** Simulation Results: Percentage Changes (CGE Model)

| | *Integrated Approach* | | *Top-Down Approach* | | *TD/BU Approach (Cons. and LS)* | | *TD/BU Approach (Labour Supply)* | |
|---|---|---|---|---|---|---|---|---|
| | ***Sector 1*** | ***Sector 2*** | ***Sector 1*** | ***Sector 2*** | ***Sector 1*** | ***Sector 2*** | ***Sector 1*** | ***Sector 2*** |
| *Commodity Prices* | -0.99 | 0.30 | -1.23 | 0.38 | -1.70 | 0.52 | -1.27 | 0.39 |
| *Domestic Sales* | -8.69 | -12.52 | -8.81 | -12.54 | -10.21 | -12.05 | -8.88 | -12.64 |
| *Domestic Production* | 27.81 | -14.20 | 27.91 | -14.31 | 26.77 | -13.86 | 27.84 | -14.43 |
| *Labour Demand* | 43.52 | -13.22 | 43.05 | -13.36 | 41.08 | -12.94 | 42.88 | -13.48 |
| *Capital Demand* | 13.07 | -26.82 | 13.14 | -26.72 | 12.72 | -25.84 | 13.15 | -26.76 |
| *Consumption** | -8.60 | -9.78 | -8.26 | -9.73 | -6.58 | -8.30 | -8.32 | -9.84 |
| *Investments* | -7.65 | -8.84 | -8.26 | -9.73 | -22.87 | -24.57 | -8.32 | -9.84 |
| *Imports* | -32.92 | -47.63 | -33.11 | -47.57 | -34.37 | -47.21 | -33.16 | -47.60 |
| *Exports* | 207.36 | -78.38 | 209.23 | -78.53 | 209.10 | -78.48 | 209.11 | -78.59 |

[*] For the integrated model, these percentage changes are computed as average percentage changes across households.

In general, we can say that we have very similar results for most of the macro variables in all the four simulations. The shock has negative effects on the economy. Indeed, as we can observe in Table 6, the fall in the price of the exported good for sector 2 causes a reduction of the production level for this sector, which reduces its demand for both factors of production. However, due to the depreciation of local currency, the reduction in the local price of the exported good is lower than the 64% world price reduction. For the same reason, exports for the other production sector become convenient, so that for this sector we observe an increase in the level of the exported good, an increase in the production level, and in the demand for capital and labour. The depreciation of local currency has a negative effect on the level of imports, which contributes to a decrease of the amount of goods sold on the domestic market.
The lower level of labour demand as a whole (the second sector is labour-intensive, as can be observed in the SAM, Table 4) generates a reduction in the wage rate, which causes a decrease in labour supply. The opposite is observed for capital, as the first sector is more capital-intensive. As a consequence of the change in the price of the factors, government increases its demand for labour input and decreases the demand for capital, as the latter has become relatively more expensive.
As the income of the representative household is based chiefly on the supply of labour, we observe a reduction in nominal income and, as a consequence, of savings and consumption expenditure. The amount of consumption goods always decrease, but the percentage change varies according to the change in their relative price: the commodity produced by the second sector has become relatively more expensive, due to the negative shock that hit the sector.
As investments are savings-driven, we observe also a reduction in the demand for investment goods (again, the investment good produced by the second sector is now relatively more expensive, so we observe a higher reduction for the demand of this good).

However, a particular result needs further explanations: savings and investments in the TD/BU-C&LS model decrease much more than in the other three models. The reason for this lays in the fact that, in order to be able to introduce exogenous consumption levels into the CGE model, we must endogenize one variable in the households' budget constraint to keep the equilibrium in this constraint. Savard's choice is for the marginal propensity to save, and we follow his approach. But the consequence of this will be a change in the household behaviour with respect to the initial assumptions made for the benchmark. Indeed, the marginal propensity to save of the household will decrease, and thus also households' savings. As in our model investments are savings-driven, this will generate a further reduction of investments. We will analyse this aspect further in the next subsection (5.1).

With respect to the microeconomic results, and mainly the changes in poverty and inequality, we can observe in Table 8 and 9 that the differences are generally significant only for the case of the integrated model.

The underlying variable for the computation of the indices is per-capita real disposable income, obtained by dividing disposable income by the household specific consumer price index[18], and then dividing it again by the number of adult equivalents resulting by the "Oxford" or "Old OECD" scale (see OECD, 1982). This equivalence scale calculates the number of adult equivalents living in a household by assigning a value of 1 to the first household member, of 0.7 to each additional adult and of 0.5 to each child:

$AE = 1 + 0.7 \cdot (\#Adults - 1) + 0.5 \cdot (\#Children)$.

**Table 8 –** Inequality Indices on Disposable per Adult Equivalent Real Income (MS Model)

| | ***Benchmark Values*** | ***Integrated Approach**** | ***Top-Down Approach**** | ***TD/BU Approach (C & LS)**** | ***TD/BU Approach (LS)**** |
|---|---|---|---|---|---|
| *Gini Index* | 33.96 | 2.81% | 1.62% | 1.47% | 1.60% |
| *Atkinson's Index, ε = 0.5* | 9.60 | 4.51% | 2.73% | 2.48% | 2.70% |
| *Coefficient of Variation* | 71.80 | 3.13% | 2.29% | 2.14% | 2.27% |
| ***Generalized Entropy Measures:*** | | | | | |
| *I(c), c = 2* | 25.78 | 6.36% | 4.64% | 4.32% | 4.60% |
| *Mean Logarithmic Deviation, I(0)* | 19.93 | 3.85% | 2.05% | 1.81% | 2.02% |
| *Theil Coefficient, I(1)* | 20.55 | 5.17% | 3.38% | 3.11% | 3.34% |

* Percentage deviations from benchmark values.

First of all, we observe that the Top-Down and the TD/BU-Labour Supply approach show almost identical results for what concerns both poverty and inequality indices.

The TD/BU-C&LS model we observe a smaller effect on inequality, but in the same direction as for the other two models, and the same is true for poverty.

[18] The household specific price index is computed using households' consumption shares and the change in prices deriving from the CGE model, as follows: $CPI_m = \sum_{q=1}^{2} \eta_{mq} \cdot \left(1 + \Delta P_q^{CGE}\right)$.

The biggest difference in the microeconomic results is to be detected in the integrated approach, where we observe a higher increase both in the inequality and poverty indices. The increase in inequality for the integrated approach is also confirmed by the higher level of the Severity of Poverty Index, which measures the degree of inequality among the poor, while a higher Poverty Gap Index indicates that the gap between the income of the poor and the poverty line has increased (see Appendix B for more details on poverty indices).

**Table 9 –** Poverty Indices on Disposable per Adult equivalent Real Income (MS Model)

| | Benchmark Values | Integrated Approach* | Top-Down Approach* | TD/BU Approach (C & LS)* | TD/BU Approach (LS)* |
|---|---|---|---|---|---|
| ***General Poverty Line*** | | | | | |
| *Headcount Index, $P_0$* | 39.34 | 16.67% | 8.33% | 8.33% | 8.33% |
| *Poverty Gap Index, $P_1$* | 9.88 | 40.09% | 28.48% | 28.07% | 28.42% |
| *Poverty Severity Index, $P_2$* | 0.00 | 39.99% | 29.42% | 28.98% | 29.36% |
| ***Extreme Poverty Line*** | | | | | |
| *Headcount Index, $P_0$* | 4.92 | 33.33% | 33.33% | 33.33% | 33.33% |
| *Poverty Gap Index, $P_1$* | 0.96 | 3.34% | 3.18% | 3.04% | 3.15% |
| *Poverty Severity Index, $P_2$* | 0.00 | -0.36% | -0.34% | -0.27% | -0.34% |

* Percentage deviations from benchmark values.

## 5.1. More on the TD/BU Approach

In this subsection we want to investigate further what happens within the TD/BU approach in general, and in particular we will try to understand which is the main cause of the unusual deviation that is observed in the level of savings under the TD/BU-C&LS approach.

At a first intuition, such a deviation could be generated either by a problem of initial data inconsistency between the two datasets (the SAM and the survey), or by what we will refer to as "feedback effects" from the microeconomic level of analysis. With this concept we intend to incorporate all the effects that derive from a response (behavioural or not) of the agents in the MS model that is different from the one observed in the CGE model for the Representative Household (RH). This difference could be due either to a different way of modelling a particular behaviour in the two models (for instance, in the case of labour supply, the MS model uses a discrete and individualized concept of labour supply, while in the CGE model we have a continuous labour supply defined for the RH), or simply to the fact that in the MS model we

consider single households as the unit of modelling, while in the CGE model we have a unique RH (as for consumption and savings, for instance).
In order to check whether the problem derives from an initial data inconsistency, we will run the same model using a new Social Accounting Matrix, which has been built in such a way that it is fully consistent with the data observed in the survey appropriately aggregated. As we can observe in Table 10, the variables that were adjusted to survey data are those in the grey cells, while all the other columns and rows were then rebalanced to obtain full consistency[19]. By comparing this SAM with the original one in Table 4, we can observe that in our case initial data inconsistencies were not very big (the biggest inconsistency is observed in the savings level).

**Table 10 –** SAM of the Economy made consistent with the Household Survey

| | $C_1$ | $C_2$ | $S_1$ | $S_2$ | K | L | H | G | SI | RoW | Total |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $C_1$ | | | 57.8 | 15.6 | | | 95.4 | 62.6 | 28.1 | 23.6 | 283.0 |
| $C_2$ | | | 17.1 | 23.5 | | | 313.2 | 48.8 | 13.6 | 76.6 | 492.8 |
| $S_1$ | 283.3 | | | | | | | | | | 283.0 |
| $S_2$ | | 492.5 | | | | | | | | | 492.8 |
| K | | | 73.4 | 23.2 | | | | 13.2 | | | 109.8 |
| L | | | 81.7 | 353.8 | | | | 117.5 | | | 552.6 |
| H | | | | | 109.8 | 552.6 | | 38.7 | | | 701.2 |
| G | | | 12.3 | 17.7 | | | 250.8 | | | | 280.8 |
| SI | | | | | | | 41.7 | | | | 41.7 |
| RoW | | | 40.8 | 59.4 | | | | | | | 100.2 |
| Total | 283.0 | 492.8 | 283.0 | 492.8 | 109.8 | 552.6 | 701.2 | 280.8 | 41.7 | 100.2 | |

$C_q$: consumption of good $q$; $S_q$: sector $q$; ***K***: capital account; ***L***: labour account; ***H***: representative household account; ***G***: public sector; ***SI***: savings-investments account, ***RoW***: Rest of the World account.

With the SAM shown in Table 10, we will run the shock on the export price of sector 2 as before (-64%). Results are reported in Tables 11 and 12 for the TD/BU-C&LS (consumption and labour supply levels are reported from the MS model into the CGE model) and the TD/BU-LS (only labour supply is reported from the micro level) approaches. Observing the result for savings in the TD/BU-C&LS approach, we can see that in our case data inconsistencies were responsible only for a 2% change in the marginal propensity to save and in the savings level. This means that the remaining change of around 13% (the difference between the change ob-

[19] To rebalance the SAM a least square minimization method was used.

served in the other approaches, around 9%, and the one observed in this approach, 22.24%) is to be attributed to the feedback effects from the MS model.

Observing the results for the TD/BU-LS approach we discover instead that the change in labour supply that was observed after the first iteration (-1.32% instead of -1.18% of the first iteration) was due only to a problem of data inconsistency and not to feedback effects from the MS model. This means that modelling labour supply as a discrete choice and individually in the MS model does not affect the results of the macro model in a significant way, at least for what concerns our particular case.

**Table 11 –** Simulation Results with Consistent Data: Percentage Changes

| | *TD/BU Approach (Cons. and LS)* | *TD/BU Approach (Labour Supply)* |
|---|---|---|
| *Government Surplus* | 0.00 | 0.00 |
| *Wage Rate* | -14.63 | -14. 81 |
| *Capital return* | 18.36 | 19.37 |
| *Consumer Price Index (num.)* | 0.00 | 0.00 |
| *Exchange rate* | 53.90 | 53.80 |
| *Labour Supply* | -1.18 | -1.18 |
| *Government Use of Labour* | 4.13 | 4.42 |
| *Government Use of Capital* | -24.89 | -25.48 |
| *Income* | -9.45 | -9.43 |
| *Disposable Income* | -9.45 | -9.43 |
| *Consumption Expenditure* | -8.14 | -9.43 |
| *Marginal Propensity to Save* | -14.13 | 0.00 |
| *Savings* | -22.24 | -9.43 |
| *Tax Revenues* | -9.57 | -9.52 |

**Table 12 –** Simulation Results with Consistent Data: Percentage Changes

| | *TD/BU Approach (Cons. and LS)* | | *TD/BU Approach (Labour Supply)* | |
|---|---|---|---|---|
| | *Sector 1* | *Sector 2* | *Sector 1* | *Sector 2* |
| *Commodity Prices* | -1.44 | 0.44 | -1.07 | 0.33 |
| *Domestic Sales* | -9.86 | -12.06 | -8.89 | -12.55 |
| *Domestic Production* | 26.77 | -13.80 | 27.65 | -14.27 |
| *Labour Demand* | 41.65 | -12.85 | 43.17 | -13.30 |
| *Capital Demand* | 12.70 | -25.99 | 13.05 | -26.76 |
| *Consumption* | -7.13 | -8.45 | -8.45 | -9.73 |
| *Investments* | -21.11 | -22.58 | -8.45 | -9.73 |
| *Imports* | -34.12 | -47.30 | -33.10 | -47.63 |
| *Exports* | 207.50 | -78.34 | 207.46 | -78.43 |

Once we have established that in the case of the TD/BU-C&LS approach most of the deviation in the savings level (13% against a 2% due to data inconsistencies) is to be attributed to feedback effects coming from the micro level of analysis, we want now to understand which is the variable or the parameter that affects mostly this deviation. Intuitively, as we have already seen with the TD/BU-LS approach that the different way of modelling labour supply does not have big effects, then this deviation in the savings level must be due to the fact that in the MS model we have expenditure shares and tax parameters that are specific to every single household, while in the CGE model there is only one RH group with "average" shares and parameters (in this sense ours is an extreme case, as we have only one RH in the CGE model). In order to understand which is the parameter that particularly affects the deviation in the savings level, we run the MS model using for all the households the RH's shares taken from the CGE model, instead of the shares and parameters that are observed in the survey for each household. The communicating variables from the MS model to the CGE model will remain the ones used in the TD/BU-C&LS approach, that is consumption levels and labour supply. Results in Table 13 clearly indicate that the main cause of difference between the two models is to be detected in the income tax rate, while labour supply and expenditure shares account only for a small part of it (the change in the savings level remains at 22% in these cases). When we use all the parameters from the CGE model (labour supply change, income tax rate, *mps* and consumption shares), the deviation in the savings level is almost reduced to zero, as it was to be expected.

**Table 13 –** TD/BU-C&LS approach with consistent data: RH shares from CGE model used in the MS model (Percentage Changes, CGE Model Results)

| | ***only ty*** | ***only ΔLS*** | ***only $\eta_i$ & mps*** | ***ΔLS, ty, mps & $\eta_i$*** |
|---|---|---|---|---|
| *Marginal propensity to save* | 2.92 | -14.82 | -14.47 | 0.12 |
| *Savings* | -6.78 | -22.87 | -22.55 | -9.33 |

These results are not surprising, as the income tax rate in the MS model is modelled in a way that is not linear with respect to the income level, as the rate depends on the income brackets to which household income belongs. Of course this feature is not captured at all in the CGE model, where we have a unique tax rate for the RH that is merely proportional to his income. Under the TD/BU-C&LS approach, while transmitting the consumption level from the MS to the CGE model, we were implicitly transmitting a level of disposable (after tax) income that was incompatible with the one of the CGE model[20].

As a consequence of our modelling choices (made following Savard, 2003), all the effect of the mismatching between the disposable income levels of the two models is going into the change in the marginal propensity to save, then into the savings and investments levels as a consequence, but it was not transmitted in a significant way to the rest of the economy. Indeed, if we observe the results in Tables 6 and 7, we would be tempted to say that, except for these big deviations in savings and investments levels (and a lower difference in the level of consumption), for the rest feedback effects do not appear to bring about significant differences in the results. This is even more evident once we have eliminated the effects coming from data inconsistencies (see Tables 11 and 12 compared with the columns for the Top-Down approach of Tables 6 and 7).

But the deviation in the savings level is quite big[21], even after having eliminated the problem of data inconsistency, and it allows us to believe that all the effects from the micro level of analysis are absorbed by the change in savings (and consequently of investments), and only in a very small part they are transmitted to the rest of the economy. Thus, a doubt arises: is consumption in our case[22] the right variable to pass the feedback effects onto the CGE model? And then, the choice of letting the marginal propensity to save free to vary in the CGE model was the best channel to transmit these feedback effects to the whole economy?

Which is the parameter we have seen to be driving the biggest change between the micro and the macro level? It is income the tax rate, which is in our case the main determinant of disposable income. So let us try to use this parameter (conveniently "aggregated" into a representative one), together with the change in aggregate labour supply, as communicating variable from the MS model to the CGE model. We will try to use not only the income tax rate from the MS model, but also the marginal propensity to save and the consumption shares.

---

[20] In both our models, consumption and savings are simply modelled as fixed proportions of disposable income.

[21] In the paper by Savard (2003), where he analyses the case of Philippines using a TD/BU-C&LS approach, «… results of variation of this adjustment variable [the marginal propensity to save, n.d.a.] have shown to be relatively small» (page 21). This probably means that the feedback effects in that case are not particularly important for the results of the model.

[22] We remember that in our case consumption is not modelled in a significantly different way in the two models. However, there could be other cases where the level of consumption can be an important carrier of feedback effects from the micro level of analysis.

Results are shown in Tables 14 and 15. As we can see by comparing these results with the ones in Table 6 and 7 for the Top-Down approach, feedback effects from the micro level of analysis can be important. In particular, in our case, we observe a different path for disposable income and tax revenues (due to the reduction of the income tax rate), and for savings and consumption, whose percentage changes are now closer to the ones of the MS model (see Table 16). Anyway, full consistency between the CGE and the MS model results is only obtained when working with consistent data and when all the parameters (change in labour supply, tax rates, marginal propensity to save and consumption shares) are transmitted to the CGE model. However, if we report all these parameters from the MS model into the CGE model without having previously adjusted the data, we can see in Tables 14 and 15 that the problem of data inconsistency comes out again and distorts the results of the CGE model, and especially the level of savings (and that of investments as a direct consequence)[23].

Here we would like to focus also on another important fact: the Top-Down approach suffers not only from the problem of a lack of feedback effects from the micro level of analysis, but it is not even exempt from the problem of data inconsistency. Indeed, the fact that the results of the two models (the micro and the macro model) do not coincide, as it is in our case, could be due either to a problem of initial data inconsistency or to a different microeconomic behaviour of the agents in the MS model. In any case, one has to decide which results are the most reliable ones in the case they do not coincide.

We report also results on income inequality and poverty changes after the simulation of the shock, for the three models described above (Tables 17 and 18).
As we can see, no big differences are observed with respect to the results reported in Tables 8 and 9. This means that, at least in our case, the fact of taking into account feedback effects does not have a strong influence on the results on income distribution and on poverty change.
In any case, these values confirm once again the fact that the integrated approach tends to overestimate the effects of the shock on income inequality and poverty change, even though at the macro level we do not observe significant deviations in the main macroeconomic variables (see Tables 6 and 7).

[23] Indeed, if we observe the two SAMs (Table 4 and Table 10, respectively), we can see that the level of savings is one of the biggest sources of data inconsistency between the SAM and the survey.

**Table 14 –** Simulation Results TD/BU Approach: Percentage Changes (CGE Model)

| | *ΔLS & ty (inconsistent data)* | *ΔLS, ty, mps & $\eta_l$ (inconsistent data)* | *ΔLS, ty, mps & $\eta_l$ (consistent data)* |
|---|---|---|---|
| *Government Surplus* | 0.00 | 0.00 | 0.00 |
| *Wage Rate* | -14.70 | -14. 62 | -14.84 |
| *Capital return* | 19.43 | 18.95 | 19.46 |
| *Consumer Price Index (num.)* | 0.00 | 0.00 | 0.00 |
| *Exchange rate* | 53.90 | 53.95 | 54.02 |
| *Labour Supply* | -1.18 | -1.18 | -1.18 |
| *Government Use of Labour* | 2.26 | 2.13 | 1.62 |
| *Government Use of Capital* | -26.96 | -26.69 | -27.55 |
| *Income* | -9.39 | -9.40 | -9.44 |
| *Disposable Income* | -8.47 | -8.48 | -8.12 |
| *Consumption Expenditure* | -8.47 | -7.93 | -8.14 |
| *Marginal Propensity to Save* | 0.00 | -5.53 | 0.25 |
| *Savings* | -8.47 | -13.54 | -7.89 |
| *Tax Revenues* | -10.95 | -10.97 | -11.60 |

**Table 15 –** Simulation Results TD/BU Approach: Percentage Changes (CGE Model)

| | *ΔLS & ty (inconsistent data)* | | *ΔLS, ty, mps & $\eta_l$ (inconsistent data)* | | *ΔLS, ty, mps & $\eta_l$ (consistent data)* | |
|---|---|---|---|---|---|---|
| | *Sector 1* | *Sector 2* | *Sector 1* | *Sector 2* | *Sector 1* | *Sector 2* |
| *Commodity Prices* | -1.21 | 0.37 | -1.38 | 0.42 | -1.09 | 0.33 |
| *Domestic Sales* | -8.75 | -12.00 | -9.27 | -11.77 | -8.92 | -11.73 |
| *Domestic Production* | 28.13 | -13.75 | 27.72 | -13.53 | 27.87 | -13.42 |
| *Labour Demand* | 43.37 | -12.79 | 42.66 | -12.58 | 43.46 | -12.44 |
| *Capital Demand* | 13.28 | -26.30 | 13.11 | -25.93 | 13.20 | -26.07 |
| *Consumption* | -7.35 | -8.81 | -6.90 | -8.24 | -7.45 | -8.35 |
| *Investments* | -7.35 | -8.81 | -12.33 | -13.91 | -6.88 | -8.19 |
| *Imports* | -33.09 | -47.31 | -33.57 | -47.16 | -33.20 | -47.23 |
| *Exports* | 210.17 | -78.31 | 210.17 | -78.27 | 208.79 | -78.11 |

**Table 16 –** Simulation Results TD/BU Approach: Percentage Changes (MS Model)

| | *ΔLS & ty (inconsistent data)* | | *ΔLS, ty, mps & $\eta_l$ (consistent data)* | | *TD Approach (inconsistent data)* | |
|---|---|---|---|---|---|---|
| | *Sector 1* | *Sector 2* | *Sector 1* | *Sector 2* | *Sector 1* | *Sector 2* |
| *Consumption* | -7.23 | -8.28 | -7.45 | -8.35 | -7.21 | -8.28 |
| *Savings* | | -7.78 | | -7.88 | | -7.78 |

**Table 17 –** Inequality Indices on Disposable per Adult Equivalent Real Income (MS Model)

| | Benchmark Values | ΔLS & ty (in-consistent data)* | ΔLS, ty, mps & $\eta_l$ (consistent data)* |
|---|---|---|---|
| *Gini Index* | 33.96 | 1.63% | 1.64% |
| *Atkinson's Index, ε = 0.5* | 9.60 | 2.76% | 2.76% |
| *Coefficient of Variation* | 71.80 | 2.31% | 2.32% |
| ***Generalized Entropy Measures:*** | | | |
| *I(c), c = 2* | 25.78 | 4.68% | 4.68% |
| *Mean Logarithmic Deviation, I(0)* | 19.93 | 2.08% | 2.08% |
| *Theil Coefficient, I(1)* | 20.55 | 3.41% | 3.42% |

* Percentage deviations from benchmark values.

**Table 18 –** Poverty Indices on Disposable per Adult equivalent Real Income (MS Model)

| | Benchmark Values | ΔLS & ty (inconsis-tent data)* | ΔLS, ty, mps & $\eta_l$ (consistent data)* |
|---|---|---|---|
| ***General Poverty Line*** | | | |
| *Headcount Index, $P_0$* | 39.34 | 8.33% | 8.33% |
| *Poverty Gap Index, $P_1$* | 9.88 | 28.54% | 28.92% |
| *Poverty Severity Index, $P_2$* | 0.00 | 29.49% | 29.89% |
| ***Extreme Poverty Line*** | | | |
| *Headcount Index, $P_0$* | 4.92 | 33.33% | 33.33% |
| *Poverty Gap Index, $P_1$* | 0.96 | 3.20% | 3.31% |
| *Poverty Severity Index, $P_2$* | 0.00 | -0.35% | -0.34% |

* Percentage deviations from benchmark values.

# 6. CONCLUSION

In this paper we tried to give an assessment of the recent developments observed in methods that link together CGE and microsimulation models, with a special concern for the different linking approaches existing in the literature. Especially, we have focused our attention only on static models. By using data from a fictitious economy, we have built three models: one that follows the full integrated approach, as in Cockburn (2001); another one that follows the so called Top-Down approach, as it is developed in Bourguignon et al. (2003b), and the last one that follows the method developed by Savard (2003), also known as Top-Down/Bottom-Up model.

On one side we can say that a simple integrated approach like the one we have implemented in this paper is deficient on the side of the microeconomic specification and behavioural responses by individual agents. Anyway, the introduction of microeconometric behavioural equations into a CGE model looks of hard application and cumbersome for computational aspects.

On the other side, a Top-Down approach completely disregards the possible feedback effects coming from the microeconomic side of the economy, which could affect also the macroeconomic variables, as we have seen in subsection 5.1.

In our opinion, indeed, the TD/BU modelling looks the most complete approach, as on one side it can include all the possible microeconometric estimates to account for behavioural responses by individual agents, and on the other side it also takes into account the feedback effects from the micro to the macro level of analysis. «…The value added of this approach comes from the fact that feedback effects, provided by the household model, do not correspond to the aggregate behaviours of the representative households used in the CGE model» (Savard, 2003, page 20).

However, two main problems arise when using this approach. First of all, the way in which these feedback effects are reported into the CGE model can affect results in a fundamental way. In particular, the fact of using shares or parameters instead of absolute levels (as in Savard's approach, 2003, where consumption levels are used), when possible, seems to lead to more consistent results, especially for the fact that when transmitting absolute levels from the MS model one has to change the initial hypothesis of the CGE model (see section 4). Secondly, eventual data inconsistencies between the micro and the macro datasets can also affect results seriously, and this can be overcome only by adjusting either one or the other dataset, thus going back to the problem of data reconciliation encountered with the integrated model (see section 2). However, while with an integrated model we encounter this problem when building the model, when we run a TD/BU model without previously adjusting the data, we have the problem of data inconsistencies that enters the results and we are not able to distinguish which is the part of the change that is due to feedback effects and which is the part due to data inconsistencies.

## BIBLIOGRAPHY


Adelman I., Robinson S. (1978), Income Distribution Policy in Developing Countries: A Case Study of Korea. Oxford, Oxford University Press.

Agénor P.R., Izquierdo A. and Fofack H. (2001), IMMPA: A Quantitative Macroeconomic Framework for the Analysis of Poverty Reduction Strategies. Washington, World Bank.

Ahuja V., Bidani B., Ferreira F., Walton M. (1997), Everyone's miracle? Revisiting poverty and inequality in East Asia. Directions in Development, Washington D.C., World Bank.

Armington P.S. (1969), A Theory of Demand for Products Distinguished by Place of Production, IMF Staff Paper n° 16, pp.159-176.

Arntz, M., Boeters S., Gürtzgen N., Schubert S. (2006), Analysing Welfare Reform in a Microsimulation-AGE Model: The Value of Disaggregation, ZEW Discussion Paper No. 06-076, Mannheim.

Atkinson A.B. (1970), On the measurement of inequality. «Journal of Economic Theory», 1970, 2: 244-263.

Bacharach M. (1971), Bi-proportional Matrices and Input-Output Change, Cambridge University Press, Cambridge.

Bargain O. (2005), On Modelling Household Labour Supply with Taxation, IZA DP No. 1455.

Boccanfuso D., Cissé F., Diagne A. and Savard L. (2003), Un modèle CGE-Multi-Ménages Intégrés Appliqué à l'économie Sénégalaise, miméo, CREA, Dakar.

Bourguignon F., Pereira da Silva L., eds. (2003a), The impact of economic policies on poverty and income distribution: evaluation techniques and tools, The World Bank.

Bourguignon F., Robilliard A.S. and Robinson S. (2003b), Representative versus real households in the macro-economic modelling of inequality, DT/2003/10, DIAL/Unité de recherche CIPRE.

Bourguignon F. and Savard L. (2008), Distributional Effects of Trade Reform: An Integrated Macro-Micro Model Applied to the Philippines, in Bourguignon F., Pereira da Silva L. A. and Bussolo M. (eds.), The Impact of Macroeconomic Policies on Poverty and Income Distribution Macro-Micro Evaluation Techniques and Tools, Palgrave-Macmillan Publishers Limited, Houndmills.

Bussolo M. and Lay J. (2003), Globalisation and Poverty Changes: A Case Study on Colombia, Working Paper No. 226, OECD Development Centre, Paris.

Byron R.P. (1978), The Estimation of Large Social Accounting Matrices, «Journal of the Royal Statistical Society», Series A, 141 (3).

Cockburn J. (2001), Trade Liberalization and Poverty in Nepal: A Computable General Equilibrium Micro-simulation Approach. Working Paper 01-18, CRÉFA, Université Laval.

Cogneau D. (2001), Formation du revenue, segmentation et discrimination sur le marché du travail d'une ville en développement: Antananarivo fine de siècle, DIAL DT/2001-18.

Cogneau D., Robilliard A.-S. (2001), Growth, Distribution and poverty in Madagascar: Learning from a Microsimulation Model in a General Equilibrium Framework, IFPRI TMD Discussion Paper No. 61 and DIAL DT/2001-19.

Cogneau D., Robilliard A.-S. (2004), Poverty Alleviation Policies in Madagascar: A Micro-Macro Simulation Model, DT/2004/11, DIAL/Unité de Recherche CIPRÉ.

Colatei D., Round J. (2000), Poverty and Policy: Some Experiments with a SAM-Based CGE Model for Ghana, mimeo, Warwick University.

Cororaton C.B., Cockburn J. (2005), Trade Reform and Poverty in the Philippines: a Computable General Equilibrium Microsimulation Analysis, CIRPÉE, Working Paper 05-13, Université Laval.

Creedy J., Duncan A., Harris M. and Scutella R. (2002), Microsimulation Modelling of Taxation and The Labour Market: The Melbourne Institute Tax Transfer Simulator, Cheltenham: Edward Elgar.

Creedy J. and Kalb G. (2005): Discrete hours labour supply modelling: specification, estimation and simulation, «Journal of Economic Surveys» 19, 697-734.

Davies J.B. (2004), Microsimulation, CGE and Macro Modelling for Transition and Developing Economies, WIDER Discussion Paper No. 2004/08.

Decaluwé B., Patry A., Savard L. and Thorbecke E. (1999a), Poverty Analysis Within a General Equilibrium Framework, Working Paper 99-09, African Economic Research Consortium, CRÉFA 99-06.

Decaluwé B., Dumont J.C. and Savard L. (1999b), Measuring Poverty and Inequality in a Computable General Equilibrium Framework, Université Laval, Cahier de recherche du CRÉFA n° 99-20.

Dervis K., de Melo J., Robinson S. (1982), General Equilibrium Models for Development Policy. New York, Cambridge University Press.

Duncan A. and Weeks M. (1998), Simulating transitions using discrete choice models, «Papers and Proceedings of the American Statistical Association» 106, pp.151-156.

Ferreira F.H.G. and Litchfield J.A. (2001), Education or inflation?: The Micro and Macroeconomics of the Brazilian Income Distribution during 1981-1995, «Cuadernos de Economia», 38, pp. 209-238.

Foster J., Greer J., Thorbecke E. (1984), A Class of Decomposable poverty Measures, «Econometrica», 52(3), pp. 293-312.

Heckman J. J. (1976), The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models. «Annals of Economic and Social Measurement», 5(4): 475-492.

Heckman J. J. (1979), Sample Selection Bias as a Specification error, «Econometrica», Vol. 47, No.1, pp.153-162.

Hérault N. (2005), A Micro-Macro Model for South Africa: Building and Linking a Microsimulation Model to a CGE Model, Melbourne Institute Working Paper Series, WP No. 16/05.

Lysy F.J., Taylor L. (1980), The General Equilibrium Income Distribution Model, in Taylor L., Bacha E., Cardoso E., Lysy F.J. (eds.), Models of Growth and Distribution for Brazil, Oxford: Oxford University Press, Ch.6.

Magnani R. and Mercenier J. (2007), On linking microsimulation and applied GE by exact aggregation of heterogeneous discrete-choice making agents, mimeo, THEMA - University de Cergy-Pontoise.

Mookherjee D., Shorrocks A. (1982), A Decomposition Analysis of the Trend in UK Income Inequality, «Economic Journal», Vol. 92, n° 368, 886-902.

Müller T. (2004), Evaluating the economic effects of income security reforms in Switzerland: an integrated microsimulation-computable general equilibrium approach, paper presented at the IIOA-Ecomod Conference, Brussels, September 2-4 2004.

OECD (1982), The OECD List of Social Indicators, Paris.

Orcutt G. (1957), A new type of socio-economic system, «Review of Economics and Statistics», Vol. 58, pp. 773-797.

Robilliard A.-S. and Robinson S. (2003), Reconciling Household Surveys and National Accounts Data Using a Cross Entropy Estimation Method, «Review of Income and Wealth», vol. 49, No. 3, pp. 395-406.

Robinson S., Cattaneo A., El Said M. (2001), Updating and Estimating a Social Accounting Matrix Using Cross Entropy Methods, «Economic Systems Research» 13 (1): 49-64.

Savard L. (2003), Poverty and Income Distribution in A CGE-Household Micro-Simulation Model: Top-Down/Bottom-Up Approach, CIRPÉE Working Paper 03-43, Université Laval.

Savard L. (2004), Poverty and Inequality Analysis Within a CGE Framework: a Comparative Analysis of the Representative Agent and Micro-Simulation Approaches, CIRPÉE, Working Paper 04-12, Université Laval.

Shoven J.B., Whalley J. (1984), Applied General Equilibrium Models of Taxation and International Trade. «Journal of Economic Literature» 22: 1007-51.

Stone R. (1977), "Forward" to G. Pyatt, A. Roe, Social Accounting for Development Planning with Special Reference to Sri Lanka. Cambridge University Press, Cambridge, pp. xvi-xxxi.

Van Soest A. (1995), A Structural Model of Family Labour Supply: A Discrete Choice Approach, «Journal of Human Resources», n.30, pp.63-88.

## Appendix A – Equations for the CGE Model

| | | |
|---|---|---|
| Demand for consumption goods | $P_q \cdot C_q = \alpha H_q \cdot CBUD \qquad q = 1,2$ | *C.1* |
| Leisure | $C_l = \left[(1-ty)\cdot PL\right]^{-1} \cdot \frac{\alpha H_l}{(1-\alpha H_l)} \cdot CBUD$ | *C.2* |
| Labour supply | $LS = TS - C_l$ | *C.3* |
| Savings | $S = mps \cdot (1-ty)\cdot Y$ | *C.4* |
| Consumer price index | $PC = \prod_{q=1}^{2} P_q^{\alpha H_q}$ | *C.5* |
| CES production function | $XD_q = aF_q \cdot \left[\gamma F_q \cdot K_q^{\frac{(\sigma F_q - 1)}{\sigma F_q}} + (1-\gamma F_q)\cdot L_q^{\frac{(\sigma F_q - 1)}{\sigma F_q}}\right]^{\frac{\sigma F_q}{(\sigma F_q - 1)}}$ | *C.6* |
| CES FOC for capital | $K_q = \frac{XD_q}{aF_q} \cdot \left(\frac{\gamma F_q}{PK}\right)^{\sigma F_q} \cdot \left[\gamma F_q^{\sigma F_q} \cdot PK^{(1-\sigma F_q)} + (1-\gamma F_q)^{\sigma F_q} \cdot PL^{(1-\sigma F_q)}\right]^{\frac{\sigma F_q}{(1-\sigma F_q)}}$ | *C.7* |
| Demand for commodity *q* as investment good | $P_q \cdot I_q = \alpha I_q \cdot S$ | *C.8* |
| Price of imports in local currency | $PM_q = (1+tm_q)\cdot PWM_q \cdot ER$ | *C.9* |
| Price of exports in local currency | $PE_q = PWE_q \cdot ER$ | *C.10* |
| Armington function | $X_q = aA_q \cdot \left[\gamma A_q \cdot M_q^{\frac{(\sigma A_q - 1)}{\sigma A_q}} + (1-\gamma A_q)\cdot XDD_q^{\frac{(\sigma A_q - 1)}{\sigma A_q}}\right]^{\frac{\sigma A_q}{(\sigma A_q - 1)}}$ | *C.11* |
| Armington FOC for imports | $M_q = \left(X_q / aA_q\right)\cdot\left(\gamma A_q / PM_q\right)^{\sigma A_q} \cdot \left[\gamma A_q^{\sigma A_q} \cdot PM_q^{(1-\sigma A_q)} + (1-\gamma A_q)^{\sigma A_q} \cdot PDD_q^{(1-\sigma A_q)}\right]^{\frac{\sigma A_q}{(1-\sigma A_q)}}$ | *C.12* |
| CET function | $XD_q = aT_q \cdot \left[\gamma T_q \cdot E_q^{\frac{(\sigma T_q - 1)}{\sigma T_q}} + (1-\gamma T_q)\cdot XDD_q^{\frac{(\sigma T_q - 1)}{\sigma T_q}}\right]^{\frac{\sigma T_q}{(\sigma T_q - 1)}}$ | *C.13* |

| | | |
|---|---|---|
| CET FOC for exports | $E_q = \left( {XD_q}/{aT_q} \right) \cdot \left( {\gamma T_q}/{PE_q} \right)^{\sigma T_q} \cdot$ $\left[ \gamma T_q^{\sigma T_q} \cdot PE_q^{(1-\sigma T_q)} + (1-\gamma T_q)^{\sigma T_q} \cdot PDD_q^{(1-\sigma T_q)} \right]^{\frac{\sigma T_q}{(1-\sigma T_q)}}$ | *C.14* |
| Market clearing condition for labour | $\sum_{q=1}^{2} L_q + LG = LS$ | *C.15* |
| Market clearing condition for capital | $\sum_{q=1}^{2} K_q + KG = KS$ | *C.16* |
| Market clearing condition for commodity *q* | $XD_q + M_q \cdot (1 + tm_q) = \sum_{s=1}^{2} io_{qs} \cdot XD_q + CG_q +$ $+ C_q + I_q + E_q$ | *C.17* |
| Income definition | $Y = PK \cdot KS + PL \cdot LS + TF \cdot PC$ | *C.18* |
| Disposable income minus savings | $CBUD = (1 - ty) \cdot Y - S$ | *C.19* |
| Zero profit condition in production function | $PD_q \cdot XD_q = PK \cdot K_q + PL \cdot L_q +$ $+ \sum_{s=1}^{2} io_{sq} \cdot XD_q \cdot PD_s$ | *C.20* |
| Zero profit condition in Armington function | $P_q \cdot X_q = PM_q \cdot M_q + PDD_q \cdot XDD_q$ | *C.21* |
| Zero profit condition in CET function | $PD_q \cdot XD_q = PE_q \cdot E_q + PDD_q \cdot XDD_q$ | *C.22* |
| Demand of commodity *q* by government | $P_q \cdot CG_q = \alpha CG_q \cdot (TAXREV - TF \cdot PC)$ | *C.23* |
| Demand of capital by government | $PK \cdot KG = \alpha KG \cdot (TAXREV - TF \cdot PC)$ | *C.24* |
| Demand of labour by government | $PL \cdot LG = \alpha LG \cdot (TAXREV - TF \cdot PC)$ | *C.25* |
| Tax revenues | $TAXREV = ty \cdot Y + \sum_{q=1}^{2} (tm_q \cdot PWM_q \cdot ER)$ | *C.26* |
| | | |
| Exogenous variables:<br>- capital endowment (*KS*)<br>- time endowment (*TS*)<br>- public transfers (*TF*)<br>- world prices ($PWE_q$ and $PWM_q$)<br>- Numeraire: consumer price index (*PC*) | Number of variables: 49<br>Number of equations: 41<br>Number of exogenous variables: 8<br>Walras' law satisfied<br>Model homogeneous of degree one | |

**Variables:**

| | |
|---|---|
| $PK$ | return to capital |
| $PL$ | wage rate |
| $P_q$ | Armington composite good price |
| $PD_q$ | output price |
| $PM_q$ | import prices in local currency |
| $PE_q$ | export prices in local currency |
| $ER$ | exchange rate (numeraire) |
| $PC$ | consumer price index |
| $KS$ | capital endowment (exogenous) |
| $LS$ | labour supply (endogenous) |
| $TS$ | time endowment (exogenous) |
| $X_q$ | domestic sales-Armington composite |
| $XD_q$ | domestic output |
| $M_q$ | imports |
| $E_q$ | exports |
| $K_q$ | capital demand by firms |
| $KG$ | capital demand by government |
| $L_q$ | labour demand by firms |
| $LG$ | labour demand by government |
| $I_q$ | demand for investment goods |
| $C_q$ | demand for consumption goods |
| $C_l$ | demand for leisure |
| $CG_q$ | government commodity demand |
| $Y$ | RH's income |
| $S$ | RH's savings |
| $CBUD$ | RH's disposable income |
| $TF$ | public transfers to RH (exogenous) |
| $PDD_q$ | price of domestic production delivered to domestic market |
| $XDD_q$ | domestic production delivered to domestic markets |
| $PWE_q$ | export prices in foreign currency (exogenous) |
| $PWM_q$ | import prices in foreign currency (exogenous) |
| $TAXREV$ | tax revenue |

**Parameters:**

| | |
|---|---|
| $ty$ | direct income tax rate |
| $tm_q$ | tariff rate on imports |
| $mps$ | RH's marginal propensity to save |
| $io_{qs}$ | technical coefficients |
| $aF_q$ | efficiency parameter of firm $q$'s production function |
| $\gamma F_q$ | share parameter in CES production function |
| $\sigma F_q$ | elasticity of substitution in CES production function |
| $\alpha H_q$ | C-D power of commodity $q$ in RH's utility function |
| $\alpha H_l$ | C-D power of leisure in RH's utility function |
| $\alpha I_q$ | C-D power of good $q$ in Bank's utility function |
| $\alpha CG_q$ | C-D power of commodity $q$ in gov.'s utility function |
| $\alpha KG$ | C-D power of capital in government's utility function |
| $\alpha LG$ | C-D power of labour in government's utility function |
| $aA_q$ | efficiency parameter in Armington function |
| $\gamma A_q$ | share parameter in Armington function |
| $\sigma A_q$ | elasticity of substitution in Armington function |
| $aT_q$ | efficiency parameter in CET function |
| $\gamma T_q$ | distribution parameter in CET function |
| $\sigma T_q$ | elasticity of transformation in CET function |
| $\varepsilon\_LS$ | wage elasticity of labour supply |

## Appendix B – Inequality and Poverty Indices

Here, we will give details of some inequality and poverty measures used during the analysis.

***Gini index***

The Gini coefficient is one of the most commonly used indicators of income inequality. It is defined as:

$$G = \frac{1}{2\mu N^2} \sum_i \sum_j \left| y_j - y_i \right|$$

where $\mu$ is the arithmetical mean of the incomes, $N$ is the size of the population, and $y_i$ and $y_j$ are the incomes of agents *i* and *j*, respectively. Thus, the second factor at the right hand side represents the sum of the differences (in modulus) computed over all pairs of incomes. In the literature, however, we can also find different (although equivalent) definitions. In particular, it can be derived from the Lorenz curve, which plots the cumulative share of total income earned by households ranked from bottom to top (see below), in the following way:

$$G = 1 - 2\int_0^1 L(p)dp \,, \tag{A.1}$$

where $L(p)$ is the Lorenz curve. The previous formula thus measures the area that is laying between the curve and the diagonal as a fraction of the total area under the 45° line. In terms of Figure A.1 below, this means:

$$G = \frac{A}{A+B} = \frac{\frac{1}{2} - B}{\frac{1}{2}} = 1 - 2B \,.$$

If the Lorenz curve coincides with the 45° line, which represents the situation of perfect equality, then the integral in equation *(A.1)* will take the value of ½, and the Gini index will equal zero.

The Gini index can thus take values between zero (perfect equality) and one (maximum level of inequality, that is, when all the income in the economy is owned by only one individual: $y_{\max} = \mu N$). Thus, the smaller is the index, the smaller is the inequality in the economy.

The Gini index is very useful because it allows the ordering of different income distributions according to their level of inequality.

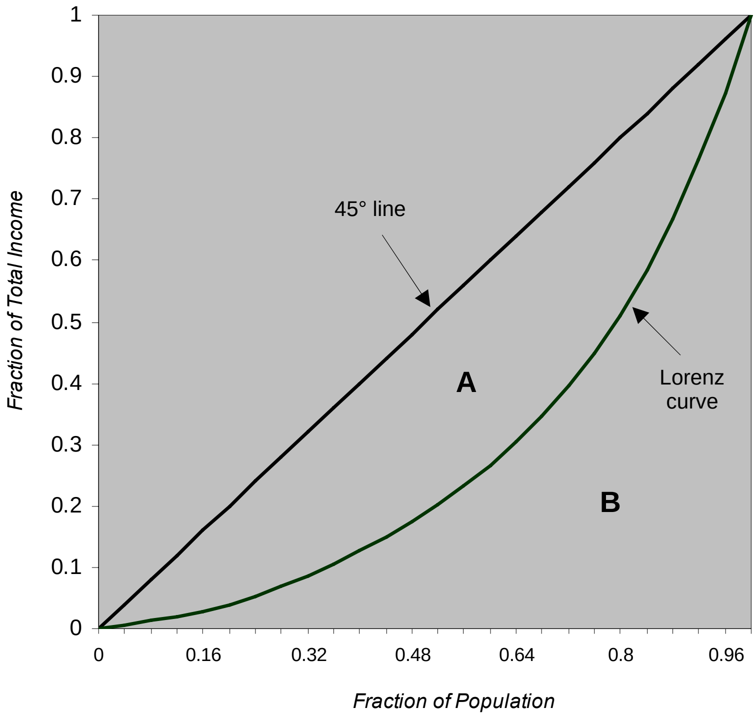


**Figure A.1 –** Lorenz Curve and Gini Coefficient

***Atkinson's index***

Atkinson's index is one of the few inequality measures that explicitly incorporate normative judgments about social welfare (Atkinson, 1970). The index is derived by calculating the so-called equity-sensitive average income ($y_e$), which is defined as that level of per capita income which if enjoyed by everybody would make total welfare exactly equal to the total welfare generated by the actual income distribution. It is sometimes also called equally distributed equivalent income. It is given by:

$$y_e = \left[ \frac{1}{N} \cdot \sum_{i=1}^{N} \left( \frac{y_i}{\mu} \right)^{(1-e)} \right]^{1/(1-e)},$$

where $y_i$ is the proportion of total income received by individual *i*, and *e* is the so-called inequality aversion parameter, which measures the degree of society's inequality aversion. It indeed reflects the strength of society's preference for equality, and can take values ranging from zero to infinity. When $e > 0$, there is a social preference for equality (or an aversion to inequality). As *e* rises, society attaches more weight to income transfers at the lower end of the distribution and less weight to transfers at the top. $e \to 0$ implies neutrality with respect to inequality, so that inequality is not perceived as a problem. Suppose instead that $e \to \infty$, then it means that there are Rawlsian preferences in the society, that is, that individuals have a preference for perfect equality. Typically, in the literature the most common values that are used for $e$ include 0.5 and 2.

The Atkinson index ($I_e$) is then given by:

$$I_e = 1 - \frac{y_e}{\mu},$$

where $\mu$ is the actual mean income. The more equal the income distribution is, the closer $y_e$ will be to $\mu$, and the lower the value of the Atkinson index. For any income distribution, the value of $I_e$ lies between 0 and 1.

***Coefficient of variation***

The coefficient of variation is a measure of the dispersion of data around the mean. It is defined as the ratio of the standard deviation to the mean, that is:

$$CV = \frac{\sigma}{\mu}.$$

The coefficient of variation is a dimensionless number that allows comparison of the variation of populations that have significantly different mean values. It is often reported as a percentage (%) by multiplying the above calculation by 100.

***Generalized Entropy Coefficients***

The family of Generalized Entropy indices satisfies a desirable property for inequality indices, that is, all the indices belonging to this family can be decomposed into a within-group and a between group contribution. The formulas for the indices are:

Generalized entropy index: $I(c) = \frac{1}{Nc} \cdot \frac{1}{(c-1)} \sum_{i=1}^{N} \left[ \left( \frac{y_i}{\mu} \right)^c - 1 \right]$ for $c \neq 0,1$

Mean Logarithmic Deviation: $I(0) = \frac{1}{N} \sum_{i=1}^{N} \ln \left( \frac{\mu}{y_i} \right)$ for $c = 0$

Theil coefficient: $I(1) = \frac{1}{N} \sum_{i=1}^{N} \frac{y_i}{\mu} \cdot \ln \left( \frac{y_i}{\mu} \right)$ for $c = 1$

Parameter $c$ reflects different perceptions of inequality, with lower values indicating a higher degree of inequality aversion. A value of $c$ greater than one means that differences at the higher end of the welfare distribution are assigned more importance than those at the lower end.

For the second index, known as Mean Logarithmic Deviation, a value of zero represents perfect equality and higher values denote increasing levels of inequality, within a given administrative unit. The parameter value 0 means that differences at the low end of the welfare distribution are assigned more importance than those at the high end.

Finally, Theil coefficient (or "information theory" measure) has a potential range from zero to infinity, with higher values (greater entropy) indicating more unequal distribution of income. If instead everyone has the same (i.e., mean) income, then the index equals 0. If one person has all the income, then the index is equal to ln($N$). The parameter value 1 means that differences are equivalently treated at all points in the welfare distribution.

The Theil index has the advantage of being additive across different subgroups or regions in the country. Indeed, it is the weighted sum of inequality within subgroups. For example, inequality within the United States is the sum of each state's inequality weighted by the state's income relative to the entire country.

If the population is divided into $m$ certain subgroups and $s_k$ is the income share of group $k$, $T_k$ is the Theil index for that subgroup, and $\mu_k$ is the average income in group $k$, then the Theil index of the population is:

$$T = \sum_{k=1}^{m} s_k \cdot T_k + \sum_{k=1}^{m} s_k \cdot \ln \left( \frac{\mu_k}{\mu} \right).$$

Therefore, one can say that a certain group "contributes" a certain amount of inequality to the whole.

***Poverty Indices***

Foster, Greer and Thorbecke (1984) have suggested a useful class of poverty indices that takes the following form:

$$P_\alpha = \frac{1}{N} \cdot \sum_{i=1}^{q} \left[ \frac{(Z_p - Y_i)}{Z_p} \right]^\alpha \quad ,$$

where $Z_p$ denotes the poverty line, $Y_i$ the expenditure or income of the $i$-th poor household (or individual), $N$ the total number of households and $q$ the number of households whose expenditures or incomes are below the poverty line. Of course, the choice of the poverty line is of great importance in the determination of the index, and it may reflect different judgements about the researcher's choice for an appropriate level of welfare.

From the general formula above, one can compute different kinds of poverty measures by simply varying the value of $\alpha$:

- If $\alpha = 0 \Rightarrow P_0 = \frac{q}{N}$

$P_0$ is also called "Headcount ratio", as it measures the incidence of poverty as the proportion of total population lying below the poverty line.

- If $\alpha = 1 \Rightarrow P_1 = \frac{1}{N} \cdot \sum_{i=1}^{q} \frac{(Z_p - Y_i)}{Z_p} = IP_0$

This index gives a good measure of the intensity of poverty, as it reflects how far the poor are from the poverty line. Indeed, it quantifies the extent to which the income of the poor lays below the poverty line. Hence the reason why it is also called "Income or Poverty gap ratio".

- If $\alpha = 2 \Rightarrow P_2 = \frac{1}{N} \cdot \sum_{i=1}^{q} \left[ \frac{(Z_p - Y_i)}{Z_p} \right]^2$

This measure is also known as "Poverty Severity Index", as it gives an indication of the degree of inequality among the poor. The greater is the inequality of distribution among the poor and thus the severity of poverty, the higher is $P_2$.