Applying exact aggregation conditions to build an integrated macro-micro simulation model for Bolivia.*

David Zavaleta
ePC - Universidad Católica Boliviana

October 2010

Abstract

In this paper we develop a computable general equilibrium (CGE) model for Bolivia, integrating in it a microsimulation model. The integration of both models is based on the exact aggregation hypothesis at the household consumption level.

We apply this integrated macro-micro model to analyze the effects of an increase on natural resources output on poverty and inequality, and to compare different redistributive policies.

Our experience shows that this model behaves very well for this kind of exercise, and shows results about poverty and inequality that are unobservable when using only a CGE model, with few representative households. The considered simulations show also some interesting results about the effects of a natural resource boom on poverty and inequality.

*The present work was extracted and modified from the Ph.D. dissertation of the author, "Measuring the impact of economic policies on poverty and inequality: the case of Bolivia" (Zavaleta, 2010)

Contents

1 Introduction

Lowering poverty and inequality levels in developing countries is among the main priorities of local governments and international institutions. Although policy makers and researchers have paid special attention to predicting the impact of economic policies on poverty and inequality, devising reliable techniques to increase accuracy in evaluation has proven to be a difficult task (Bourguignon et al, 2008).

Over the last decades, computable general equilibrium (CGE) models have been largely preferred by economists to predict the effects of economic policies. As these models reflect the interdependence of markets and the fact that "everything depends on everything" (Hertel et al., 2007), they are able to capture not only the direct effects of economic policies, but also their indirect effects. They have been widely developed in the last few decades, and currently are perhaps the most commonly used tool by researchers to simulate economic shocks and economic policies in a given country's economy.

In recent years, research attention has shifted from the evaluation of aggregate welfare gains to the analysis of income distribution. CGE models have been widely used to simulate the impact of economic policies on poverty and inequality. Typical examples are studies on the impact of trade liberalization on poverty and income inequality within each country. A crucial aspect of these studies lies in the way households are represented in the simulation model.

In a typical CGE model, households are represented by a limited number of representative households, usually differentiated by their endowments of production factors (Robilliard et al., 2008). This aggregation of households into a small number of representative household groups has been largely criticized. In this context, changes in income distribution can be simulated only *between* representative household groups, and nothing can be said about distributional changes *within* a household group. Moreover, the hypothesis that a representative household behaves like the aggregate of the households it represents has also been severely criticized (see Decaluwé (1999) and Kirman (1992) for a more detailed discussion on this topic).

This failure of typical CGE models has led researchers to develop more sophisticated tools which allow a better description of income distribution changes generated by economic policies. These new tools include top-down models (Chen and Ravallion, 2004; Bocanfuso and Savard, 2007; Robilliard et al., 2008) and fully integrated CGE models (Cockburn, 2002; Rutherford and Tarr. 2008).

In the case of top-down models, a CGE (top) model simulates the impact of economic policies on prices of factors and commodities. Then, these new prices are used in a microeconomic (bottom) model, where all households available in the survey are included. The obvious criticism of this approach is the lack of feedback from the bottom microsimulation model to the top CGE model.

The fully integrated approach includes in the CGE model as many households as there are in the household survey. Therefore, the problem of lack of feedback of the top-down approach is solved. Savard (2003) and Bourguignon and Savard (2008) propose an iterative top-down/bottom-up solution to the fully integrated approach. It consists of an extension of top-down models, which allows feedback from the bottom model to the CGE model. However, this approach requires consistency between national accounts and survey information, and to deal with large size data, which may be challenging (Bourguignon and Savard, 2008; Chen and Ravallion, 2004). Nevertheless, the fully integrated approach "appears as the only one based on rigorous theoretical framework that consider all the observed heterogeneity of the population of households" (Bourguignon and Savard, 2008).

In this work, we propose a different approach - labeled the "exact aggregated representative household" (EARH) approach — which addresses the main problems of the top-down and the fully integrated models. The EARH approach applies exact aggregation conditions of household behavior, which allows to link the CGE and microsimulation models through a limited number of representative households. By contrast to the earlier representative household models, our approach accounts for substantial heterogeneity in household behavior and is consistent with econometrically estimated demand functions at the micro level.

This work is organized as follows: Section 2 reviews how the problem of income distribution is currently treated in simulation models, and then we present our approach, the exact aggregations representative households (EARH) model. Section 3 explains the main features of the model presented in this work. Section 4 describes the different data sources used in this model, and how the survey data and the social accounting matrix (SAM) are reconciled. Section 5 describes the application of the EARH model to simulate the expansion of the hydrocarbons sector in Bolivia, and to compare different redistribution policies in the country. Section 6 presents the results at the aggregated level; and the inequality, social welfare and poverty analysis. Section 7 concludes.

2 Types of macro-micro simulation models

Whenever a CGE model is built, three major aspects have to be considered: the nature of the question we are looking to answer, the type and quality of the data, and the mechanism by which individuals are affected (Bourguignon et al, 2008).

For many years, data availability has been a major constraint for developing highly disaggregated models. Nevertheless, data bases have evolved in recent years, allowing researchers to construct more sophisticated tools nowadays. These tools have permitted the introduction of more and more available data. Besides, along with data sources development, technology has also evolved, allowing the treatment of larger sets of information.

Current models can be divided into three major groups: representative household (RH) models, top-down (TD) models and fully integrated (FI) models. A brief description of these types of models is presented below. At the end of the section we describe the EARH model proposed in this paper.

2.1 Representative household models (RH)

These models adopt the simplest approach to household disaggregation, and often only need data on national accounts for a given year. In these models, a limited number or *representative households* stand for the entire set of households in the economy. RH models have the following characteristics concerning households:

1. Through representatives households, individual households are assumed to behave similarly, and share behavioral characteristics in consumption and sources of income.
2. The CGE captures only the aggregate behavior of these groups, like labor supply and consumption.
3. Many strong assumptions are imposed for the aggregation.
4. Income distribution analysis is made between representative households, leaving the within distribution unaffected.

Concerning the last point, some authors have tried to address the problem of intra-household distribution (e.g., de Janvry et al., 1991; Decaluwé et al., 1999). All of them assume that, within a household group, income is distributed according to an exogenous distribution function. De Janvry et al. (1991) chose Lognormal distribution and Decaluwé et al. (1999) a Beta distribution. Nevertheless, the fact of fixing a given intra-household distribution remains very restrictive. Although the chosen distribution function may indeed represent the "within" distribution fairly accurately in the base scenario, this will probably not hold true any longer in the counterfactual scenario.

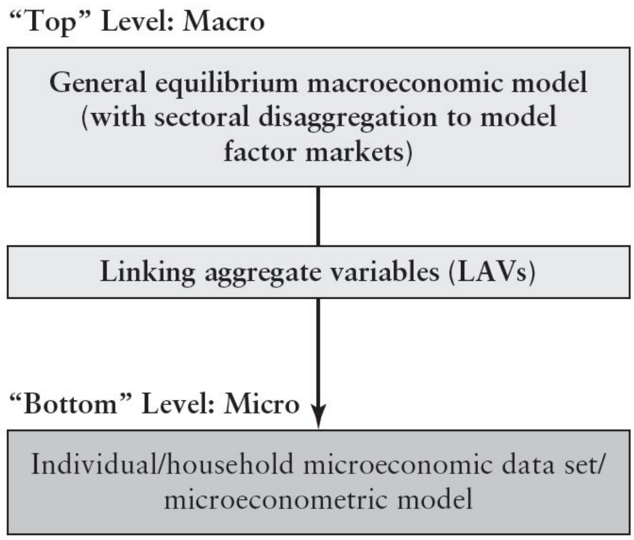
2.2 Top-down (TD) models

The greater availability of data and the development of computer technology have led economists to consider including all information in household surveys to create more sophisticated models. In this context, recursive models, top-down models, or micro simulation sequential models; have been developed to improve the existing CGE models. Top-down models consider as many households as there are in the survey in a micro bottom model. Works in this domain include Robilliard et al. (2008), Chen and Ravallion (2004), Boccanfuso and Savard (2007), Hérault (2007 and 2009).

Top-down models are composed of a macro CGE (top) model and a micro (bottom) household model. The top model is in general similar to RH models, where household disaggregation, if any, is made in a few representative households. The bottom model instead, considers all the households in the survey, and models their behavior according to the researcher's interest and considerations.

These models work in general as follows: In the top model, the macroeconomic shock is introduced. As a result, new equilibrium prices of commodities and factors are obtained. These new prices are introduced in the bottom model, where individual household income, consumption, factor supply and other behavioral variables are obtained. Figure 1 presents an expemple of a TD model mechanism.

Figure 1: Top-down models structure.



Source: Bourguignon et al. (2008).

Besides their quality of including all households of the survey, these models have two main advantages. First, no full consistence between the macroeconomic and the survey data is required. Second, top-down models provide richness in household behavior and are very flexible for behavior specification (Savard, 2003).

Robilliard et al. (2008) for example, exploit this second advantage, including occupational choices of households in the bottom model, which are discrete choices. It is done through labor supply equations, and microeconometric tools are applied.

In some cases, as in Boccanfuso and Savard (2007), no econometric household behavior is included: factor endowment is fixed, and consumption is modeled with a Cobb-Douglas expenditure function, where consumption shares are constant for all commodities.

2.3 Fully integrated (FI) models

This family of models is an extension of the RH models. Fully integrated models include in the CGE model as many households as there are in the household survey. Works in this domain include Cockburn (2002), Annabi et al. (2006), Cororaton and Cockburn (2007), Rutherford and Tarr (2008).

Although FI models should be preferred if the focus is on poverty impact of macroeconomic policies, from the strongest basis of micro observations (Bourguignon et al. 2008), this approach suffers from some serious complications with regard to its implementation. First, since FI models replace representative households in RH models by the entire set of households in the survey, reconciliation between both data sources is required, which might be problematic (Bourguignon and Savard, 2008). Second, the numerical solution can also be a challenge (Rutherford et al., 2005; Chen and Ravallion, 2004; Boccanfuso and Savard, 2007).

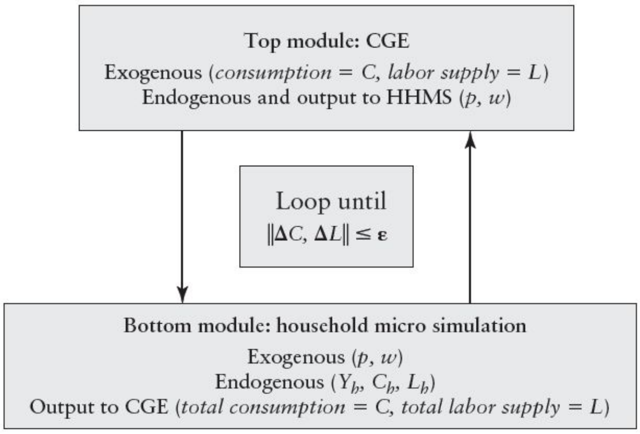
Concerning the first point, several methods for reconciling national accounts and the household survey might be applied. Rutherford and Tarr (2008) for instance, adjust factor shares in the social accounting matrix (SAM). Annabi et al. (2006), make a mixed adjustment in both SAM and survey data. Finally, Robilliard and Robinson (2003) propose a cross entropy estimation method to reconcile household surveys and national accounts. It consists in recalculating the survey weights to obtain aggregate values identical to the SAM values.

2.4 Top-down/bottom-up (TD/BU) models

Savard (2003) and Bourguignon and Savard (2008) propose an alternative TD/BU model, which is an extension of the TD approach. TD/BU models provide some feedback from the bottom model to the top model. This technique brings variables from the top CGE model (e.g., goods and factor prices) into the bottom microeconomic model. Once this bottom model is solved, its output (e.g., consumption and labor supply) is brought to the top model and so on, until a convergence criterium is satisfied. The structure of TD/BU models is illustrated in Figure 2.

Although top-down/bottom-up techniques share most qualities of top-down models (e.g., there is no need for consistence between national accounts and survey data), several assumptions have to be imposed. For instance, Bourguignon and Savard (2008) assume that the proportion of differences between consumption in both data sources is independent from the prices of other

Figure 2: Top-down/bottom-up models structure.



Source: Bourguignon and Savard (2008).

goods and wages. Besides, Savard (2003) finds that convergence of the iterative top-down/bottom-up system may be difficult in some situations.

Table A.1 in the Appendix presents an overview of the main studies that use combined CGE-microsimulation models.

2.5 Exact aggregated representative household (EARH) models

In this section we describe an alternative approach to simulate economic policy effects in a CGE-microsimulation framework. It consists in using exact aggregation conditions to create a relatively small number of representative households in the CGE.

One of the main criticisms of RH models is their underlying hypothesis that a representative household behaves as the aggregate of the households it represents. Nevertheless, under some conditions, this hypothesis can be verified, and we refer to it as *"exact aggregation"*condition (Deaton and Muellbauer, 1980). Exact aggregation is one of the the key aspects of the EARH approach.

In the next subsection we introduce the aggregation problem, and the condition required to obtain *"exact aggregation"* for consumption in the linear case. Subsection 1.2.2 applies *"exact aggregation"* to describe the EARH approach.

2.5.1 Exact linear aggregation

This subsection studies exact aggregation for consumption in the linear case, as in Deaton and Muellbauer (D&M) (1980).[1] To determine how a representative household can behave as the aggregate of the households it represents, first we explore under what conditions we can write aggregate demand as a function of prices and the aggregate expenditure.

Let's begin with the individual household demand for a given commodity:

$$c_{ih} = g_{ih}(y_h, p) \qquad i = 1, ..., I; h = 1, ..., H \tag{1}$$

where c_{ih} is the quantity of commodity i demanded by household h, which is a function g_{ih} of household income, y_h, and of the vector of commodity prices p.

Exact aggregation of all H households is possible for any commodity i, if its average demand can be written as

$$\bar{c}_i = g_i(\bar{y}, p) \qquad i = 1, ..., I \tag{2}$$

where $\bar{y}$ denotes average income of all H households.

It can be noted that expression (2) does not depend on h. This means that average demand, and therefore total demand, remains unchanged if there are income transfers among households. This is true only if Engel curves are straight lines and have the same slope for all H households. This implies that equation (1) must be linear in y_h:

$$c_{ih} = \alpha_{ih}(p) + \beta_i(p) \cdot y_h \qquad i = 1, ..., I; h = 1, ..., H \tag{3}$$

The next step is to find a family of utility functions which lead to demand functions that are linear in income. Cobb-Douglas and Stone-Geary utility functions satisfy this condition. In the latter case, maximization of the Stone-Geary utility function leads to the linear expenditure system (LES).

So far, we have shown under which conditions exact aggregation can be obtained for consumption. The next subsection applies these conditions to the EARH approach.

2.5.2 Exact aggregation and the EARH approach

The previous subsection has shown under which conditions the behavior of a representative household equals the aggregate behavior of individual households, in the case of consumption. Whenever those conditions are satisfied,

[1] The nonlinear case and the labor supply aggregation, are also considered in chapter 6 of D&M.

we say that exact aggregation is possible. In the linear case, exact aggregation is possible when individual Engel curves are straight lines, and their slope is the same for all households.

Several problems arise if we suppose that all households satisfy this condition. First, linear Engel curves might seem very restrictive. Marginal consumption of extra income may change when income increases. Second, other factors such as the demographic composition of households are likely to affect consumptions and this can also create nonlinearity in the aggregate relation between demand and income (D&M).

Nevertheless, under some conditions it can be argued that linear Engel curves remain a reasonable approximation for groups of goods (D&M). First, grouping commodities into a restricted number of aggregated goods helps to eliminate the problem of zero demands for some commodities. Next, if the demographic structure of households is a source of Engle curves' nonlinearity, grouping households by their demographic composition appears to be a logical solution to this problem. In this case, we allow households of different demographic structure, to have different slopes of Engel curves. Finally, the intercept α_{ih} in equation (3) allows to absorb some household specific determinants of demand, leaving exact aggregation unaffected.

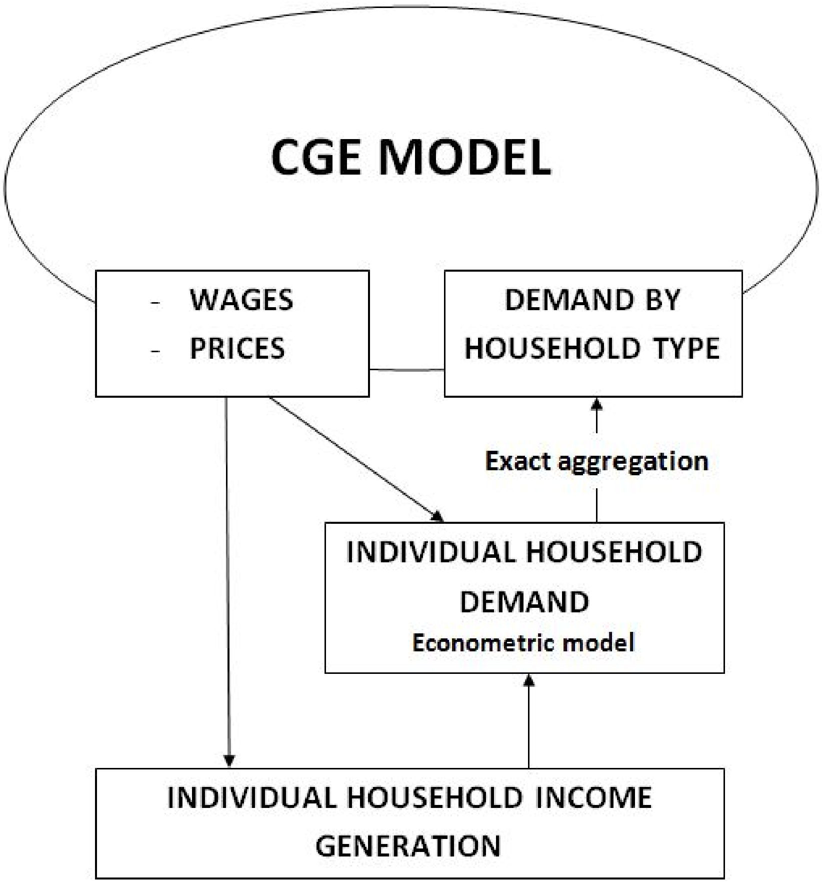
The aggregation of individual households by demographic characteristics, using exact aggregation properties for consumption, returns a limited number of representative households. The number of representative households will depend on the population's heterogeneity and the sample size. The resulting representative households will be introduced in the CGE model, which will work as a classic RH model. As we will see in the case of Bolivia, the about 5000 individual households are grouped into about 300 representative households. This reduction in the number of households eliminates computing problems with large data sizes as found in FI models.

We emphasize that by contrast to the RH models, the EARH approach is based on an econometric model of consumption behavior, which takes individual heterogeneity into account. The exact aggregation property used in this model ensures that the "representative households" in the CGE model reflect all the observable heterogeneity in behavior of the econometric model (i.e. in the marginal budget coefficients, β_i). In other words, there is a representative household for each observed combination of demographic variables in the household sample. In our case, there are about 300 different combinations of demographic variables, thus there are the same number of "household types" or representative households, representing the 5000 households in the survey.

Concerning commodities, the living standard measurement surveys collect consumption data for a large number of commodities. These commodities can be grouped using different classifications, the *"Classification of Individual Consumption According to Purpose" (COICOP)* being the most widely used. This procedure helps to eliminate the possible problem of having consumptions equal to zero for some commodities. It also increases the number of degrees of freedom if demand equations are obtained econometrically.

In the case where households' factor supply is exogenous, the EARH model structure can be represented as in Figure 3.

Figure 3: EARH model structure.



Source: Author's creation.

It is important to note that there is no need of interaction between the CGE model and the microsimulation model to achieve equilibrium for a given simulation. Exact aggregation assures that the representative household in the CGE model behaves as the aggregate of individual households. In this context, the CGE model is closed, unlike the TD or TD/BU models.

Finally, the EARH approach needs reconciliation between national accounts used in the CGE model, and survey data. This can be slightly difficult, but we believe that it is better to have both data sources reconciled, to make the model more consistent between the macro and the micro modules. Besides, several methods and approaches have been developed to make reconciliation more feasible: Robilliard and Robinson (2003), Fofana and Cockburn (2003).

3 The integrated CGE-microsimulation model

The EARH model developed in this work allows for the integration of as many households as there are in the survey through a microsimulation model. The model focuses on the interaction between both the CGE model and the microsimulation model, and on the exact aggregation condition, thereby ensuring that each representative household in the CGE model, behaves exactly as the aggregate of the individual households it represents.

Since the exact aggregation conditions are satisfied in the EARH model, the analysis framework is recursive and enables us to carry out the simulations in two stages. First, we use the CGE model to simulate the desired shock at a macroeconomic level, resulting in changes in prices of commodities and production factors. Next, new prices are transmitted to the microsimulation model in order to determine income and consumption changes at the household level.

It is important to clarify that the modeling itself is not recursive. Households are integrated in the CGE model, represented by *some* households, but satisfying perfect aggregation condition. Thus, the results obtained in the microsimulation model match the results obtained in the CGE model.

We describe in this the integrated CGE-microsimulation model for Bolivia. First, we present the microsimulation model describing the income generation and the consumption behavior of households. Next, the CGE model is briefly introduced, emphasizing the features linking this model to the microsimulation model.[2]

3.1 The microsimulation model

The microsimulation model describes the behavior of each household included in the living standard household survey. In order to ensure consistency with the CGE model, we assume that factor supply is exogenous and represent preferences by a demand system — the linear expenditure system — which allows exact aggregation and can be estimated econometrically.

3.1.1 Household income

Two sources of income are distinguished in the model: factor and non-factor income. Factor income is derived from labor, capital and land, and depends

[2]A detailed description of the CGE model is presented in Zavaleta (2010)

on factor prices and factor endowments of each household. Labor endowments are disaggregated into different skill categories.

Home-produced consumption, or *autoconsommation*, is also considered in the model. *Autoconsommation*, typically composed of food grown or raised on a farm or in gardens, can be considered a commodity produced and bought by the same household. It is properly recorded as both income and consumption; nevertheless it is often difficult to value, especially in economies where markets are not well developed and where home production accounts for an important share of total consumption (Deaton, 1997).

Autoconsommation is modeled in the side of household incomes as a production factor alongside labor and capital.

Factor endowments are assumed to be constant for every household. Factor prices are determined in the CGE model. Total factor income of household h is given by

$$\pi_h = \sum_f WF_f^* \cdot F_{hf} \tag{4}$$

where WF_f^* is the factor f price and F_{hf} is the factor endowment of the household. Note that starred variables represent values transmitted from the CGE model, which are exogenous to the microsimulation model.

Non-factor income comes from all transfers, from national and foreign institutions, and is considered exogenous

$$R_h = \sum_j r_{jh}^* \tag{5}$$

where R_h are total transfers received by household h, and r_{jh}^* is the transfer received by household h from institution j (government, other households, rest of the world, etc.).

Total income net of taxes, transfers and savings is given by

$$y_h = (\pi_h + R_h) - tax_h - T_h - S_h \tag{6}$$

where $\pi_h + R_h$ is gross income, tax_h is the amount of direct taxes, T_h is the exogenous amount of transfers made by the household, and S_h is the savings of household h.

Equation 6 shows taxes and savings as fixed values, which may appear unorthodox since they generally depend on income, and are modeled in that way.

First, there is no income tax in Bolivia. Taxes paid by households are mainly on assets (e.g. housing, land), and, although assets depend up on income, our model does not consider variations on them. Second, the survey does not consider the annual income, but rather monthly income for the month prior to the survey. It does not allow having a reliable saving propensity. Besides, several savings are negative, and instead of manipulating incomes for all these households, we prefer to allow for negative savings, but not as a share of gross income.

Keeping taxes, transfers and savings constant in value for each household is required in order achieve exact aggregation condition.

3.1.2 Household consumption

Since the microsimulation model will be integrated into a complete CGE model, a complete demand system must be estimated to satisfy the budget constraint. The most frequently used system in CGE models is the linear expenditure system (LES) (Sadoulet and de Janvry, 1995). Other demand systems, like the almost ideal demand system (AIDS) and the generalized almost ideal demand system (GAIDS), require much more information than the LES system. Because of limitations in available data, the LES system has been chosen for the model.

To explain our approach, we start with a basic version of the LES, which derives from the Stone-Geary utility function ($U = \prod (c_i - \gamma_i)^{b_i}$):

$$c_{ih} = \gamma_i + \frac{b_i}{p_i} \left(y_h - \sum_j p_j \gamma_j \right) + u_{ih} \tag{7}$$

or, multiplying both sides by p_i:

$$p_i c_{ih} = p_i \gamma_i + b_i \left(y_h - \sum_j p_j \gamma_j \right) + p_i u_{ih} \quad h \in H \tag{8}$$

where c_{ih} is the quantity of commodity i consumed by household h, p_i is its price, and H is the set containing all households. Parameters γ_i and b_i are respectively the subsistence consumption and the marginal budget share of commodity i, and u_{ih} captures unobserved factors at the individual level.

The own and cross price elasticities, ε_{ii} and ε_{ij} respectively, and the income elasticity, η_i, of the LES are respectively:

$$\varepsilon_{ii} = -1 + (1 - b_i)\frac{\gamma_i}{c_i}, \quad \varepsilon_{ij} = -\frac{b_i \gamma_j p_j}{p_i c_i} \; (i \neq j), \quad \eta_i = \frac{b_i}{\omega_i}$$

where ω_i is the budget share of commodity i.

3.1.3 Heterogeneity of behavior and exact aggregation

While the term u_{ih} in equations (7) and (8) captures factors at the individual or household level, a part of it can be imputed to the heterogeneity of household composition. It is clear, for instance, that a larger household will consume more food, but it is also true that a household composed by 5 adults will consume more calories than a household composed by 2 adults and 3 children (Deaton, 1997).

Following this distinction, we subdivide households into different demographic groups or subsets of households. Subsets or groups are noted H_g, where g is the type of household. For instance, if g stands for the type of household “2 adults and 3 children”, then H_g is the set of all households satisfying this characteristic.

Introducing heterogeneity of household composition in equation (8), we obtain equation (9) for each subset $H_g \subset H$.

$$p_i c_{ih} = p_i \gamma_{ig} + b_{ig} \left(y_h - \sum_j p_j \gamma_{jg} \right) + p_i u^*_{ih} \quad h \in H_g; g \in G \tag{9}$$

where G is the set of all possible types of household compositions.

We assume that the following normalization constraint holds:

$$\sum_{h \in H_g} p_i u^*_{ih} = 0 \tag{10}$$

Equation (9) distinguishes LES parameters γ and b for households of different composition. Unobserved factors at the individual level, others than household composition, are represented in u^*_{ih}.

Since all households in each group g have the same marginal propensity to consume each commodity i, exact aggregation is possible. This means that we can aggregate all households of group g into a representative household, whose behavior will be identical to the aggregated behavior of all households of the group.

When applying this aggregation to equation (9), we obtain:

$$p_i C_{ig} = p_i \Gamma_{ig} + b_{ig} \left(Y_g - \sum_j p_j \Gamma_{jg} \right) \qquad g \in G \tag{11}$$

where

$$C_{ig} = \sum_{h \in H_g} c_{ih}$$

$$\Gamma_{ig} = \sum_{h \in H_g} \gamma_{ih}$$

$$Y_g = \sum_{h \in H_g} y_h$$

In the next section, when introducing the CGE model, consumption is based on representative household, and equation 11 will show up to denote households demand for commodities. Given that exact aggregation is satisfied for each group of household, representative households will act exactly as the aggregate of all households represented by it.

3.2 The CGE model

The CGE model used in this work is based upon a standard model built by Lofgren, Harris and Robinson (2002), hereafter LHR, at the International Food Policy Research Institute, based at its time on Dervis et al. (1982). Besides coherence with the microsimulation model, it considers natural gas resources as an input for the oil and gas sector. As we will see further in the illustration of the model, this will enable us to introduce the shock in the model through an increase of this input. A full description of the main characteristics of the model are described in Zavaleta (2010).

3.2.1 Households

The model contains three major groups of representative households: urban, rural and the residual household. The latter was created in order to absorb the large differences between the SAM and the survey for some aggregates, which we suppose are due to the absence of some type of households (rich households) in the survey. This is explained in more detail in the next section.

Urban and rural households are disaggregated in representative households, each representing individual households of identical demographic composition. In this section, representative households are simply called households.

The SAM contains two different disaggregations of commodities, one by productivity sectors, in 35 groups, and another considering the International Comparisons Project (ICP), in ten groups, as in Reimer and Hertel (2004). Thus, household consumption is done in two stages. First, urban and rural households demand ICP commodities, and then, ICP commodities demand by sectors commodities.

For the first stage, each household is assumed to maximize an utility function. As mentioned in the previous section, the utility function used in this work is the Stone and Guery function, which leads to the linear expenditure demand system (LES). Here we recall equation (11) in the previous section:

$$p_i C_{ig} = p_i \Gamma_{ig} + b_{ig} \left(Y_g - \sum_j p_j \Gamma_{jg} \right) \qquad g \in G \tag{12}$$

where each representative household g represents all households in the survey having the same demographic characteristics. The econometric estimation of parameters b_{ig}, and the following calibration of parameters Γ_{ig} are explained in the next section.

In the second stage of household consumption, ICP commodities demand productive sector commodities using a Cobb-Douglas demand system:

$$p_j^* q_j^* = \sum_i a_{ij} \left(\sum_g p_i C_{ig} \right) \tag{13}$$

where the starred symbols refer to the price and quantity of demanded productive sector commodity j, as a function of the expenditure in ICP commodities.

For the residual household, demand is made in only one stage, directly to productive sector commodities, using a Cobb-Douglas demand system with shares equal to the observed values. Since this household must act as neutrally as possible, a Cobb-Douglas demand systems seems appropriate for this purpose. The residual household demand is then added to the household demand (13), to obtain the total household demand.

Households' total income is the sum of its income from factors, transfers from other domestic non-government institutions, net transfers from the government (indexed to the CPI), and net transfers from the rest of the world.

Total consumption spending of households is defined as the income that remains after direct taxes, transfers to other domestic non-government institutions and savings. All outcomes (taxes and transfers), and savings, are fixed. As mentioned in the previous section, this is an imposed condition in order to achieve the exact aggregation condition.

4 Data and calibration

The integrated CGE-microsimulation model presented in the preceding section includes a large number of parameters that are quantified using various data sources. While the CGE model is based on a social accounting matrix (SAM) for Bolivia, the behavioral parameters of the microsimulation model are estimated using household survey data. To ensure consistency between the CGE and microsimulation models, these data sources have to be reconciled. This section describes the main data sources and the adjustments that were carried out in order to obtain a consistent data base for the simulation model.

4.1 National accounts and the social accounting matrix (SAM)

The construction of a SAM using the available data on national accounts often represents an arduous task. In the case of Bolivia, national accounts taken from the Bolivian National Institute of Statistics (INE) do not provide enough information to build the matrix. For instance, disaggregation of value added is not available and the breakdown by factor has to be carried out using additional sources.

In the final SAM, we distinguish the following accounts: 35 activities and 35 commodity groups (shown in Table 1), value added, households, government, other institutions (enterprises), imputed bank services (which are the services borrowers purchase from banks in order to obtain a loan); savings (fixed capital and changes in stocks), investment, and rest of the world. The structure of the SAM is given in Table A.3 in the Appendix.

Table 1: Activities and commodities of the SAM

1.	NIA Non-industrialized agriculture	19.	OPR Processed oil products
2.	IAG Industrialized agriculture	20.	MPR Non-metallic minerals products
3.	COC Coca	21.	BME Base metals
4.	LSP Livestock production	22.	MAC Machinery and equipment
5.	HUN Timber hunting and fishing	23.	MAN Other manufactures
6.	OIL Oil and natural gas	24.	ENR Electricity gas and water
7.	MIN Mining	25.	CON Construction
8.	MEA Meat and processed meat	26.	TRD Trade
9.	DAI Dairy products	27.	TRS Transport and storage
10.	MIL Baking and grain mill products	28.	COM Communications
11.	SUG Sugar and confiture	29.	FIN Financial services
12.	FOO Food products	30.	ENT Company services
13.	BEV Beverages	31.	HOU Property
14.	TAB Processed Tobacco	32.	SER Local, social and personal services
15.	TEX Textiles	33.	RES Restaurants and hotels
16.	WOO Wood and wood products	34.	DOM Domestic services
17.	PAP Paper and paper products	35.	PUB Public sector
18.	CHM Chemical products		

In the official national accounts, value added is not disaggregated by production factors for the different activities. Labor, capital and indirect taxes are available only for total value added. We have used historical data to disaggregate value added for each activity.

Furthermore, capital income has been disaggregated in three different kinds: land, natural resources and physical capital, the last hereafter referred to as capital. This is important since the first two are specific to the agricultural and oil and gas sectors, respectively. For the non-industrialized agriculture (NIA) activity, total farmers revenue is classified as land revenue; while for industrialized agriculture (IAG), shares of land and capital in total capital income have been obtained from estimations from Hertel et al. (2008). These shares are 53% for land and 47% for capital.

The value imputed to income from natural resources is calculated on the basis of the royalties paid to the government for their exploitation in 2000. These royalties were 18% of the negotiated prices between the Bolivian government and the enterprise that held the right to exploit and commercialize the natural resource. An approximate value has been calculated using data of gas exports from the INE Bolivia.

Value added structure of the SAM is given in Table A.2 in the Appendix.

4.2 The living standard measurement survey

The survey used in this work is the living standard measurement survey, elaborated under the MECOVI (mejoramiento de encuestas de condición de vida) program in 2000, available at http:\\www.ine.gov.bo. It contains detailed information for about 5000 households, including demographic characteristics, income and consumption. In this section we describe the MECOVI survey, and how variables are chosen and treated for this work.

The Program for the Improvement of Living Standard Measurement Surveys, MECOVI, was organized and executed by the Inter American Development Bank (IADB), the World Bank, and the Economic Commission for Latin America and the Caribbean, and has been carried out in several countries of Latin America, Bolivia among them. Its objective is to support the involved countries in collecting reliable data regarding living standards of their populations for the development and evaluation of policies.

In the case of Bolivia, the MECOVI survey was carried out every year from 1999 to 2002. Since the social accounting matrix used here is for the year 2000, we use the survey for the same year.

The survey includes demographic information, education, employment, health, expenditures, income, house characteristics and basic needs. The survey is particulary rich in income and expenditure, which is very important for our work.

Similar to other household survey, the MECOVI survey has some limitations which should be kept in mind when analyzing the results. Information is collected through interviews with household members, who answer the questions based on current information (age, education level, employer, etc.) as well as historical information (last month's food expenditure, last year's acquisition of durable goods, etc.). In general, historical data is based on the respondent's recollection, which may compromise the accuracy of the information.

Questions about consumption expenditure are asked mainly on the basis of the "last month" for most articles. For other expenditures, like health, questions are asked for the last three months, which might be too short a period to be extrapolated to an annual basis. Moreover, the chosen survey period, which is before Christmas, may entail a bias when converting monthly and trimestral data to annual data.

The survey was obtained through the Bolivian National Institute of Statistics (INE), whose final version of the survey is already adjusted for some missing values. The data obtained from the INE contains 4994 households. After elimination of observations with erroneous and missing values, a final sample of 4833 households was retained, of which 2736 are in the urban sector and 2097 in the rural sector.

Three groups of variables have been retained for this work, regarding income, consumption and demographic characteristics. In this section, we describe these variables and the adjustment of the data in view of the reconciliation with the SAM.

4.2.1 Demographic variables

The MECOVI survey includes a large number of demographic variables, from which we have retained for our work: sector (urban or rural), households size, age and education of households members, the latter measured in years of schooling. Household size and age are variables crucial for consumption modeling, while the education level is central to labor supply (skilled, unskilled).

Demographic variables have been used to build a new set of variables, mainly grouping the original variables. Hence, a new ordinal education variable has been created with four modalities: Education 1 (no schooling at all), Education 2 (1 to 6 years of schooling), Education 3 (7 to 12 years), and Education 4 (more than 12 years of schooling). Age has also been grouped into four modalities (categories): children (up to 12 years old), teenagers (13 to 17 years old), adults (18 to 65), and elders (more than 65 years of age).

Table 2 shows all demographic variables and modalities retained for this work.

Table 2: Demographic variables

Variable	Modalities	
Sector	Urban	
	Rural	
Household size	Quantitative variable	
Education	Education1	no schooling
	Education2	1 to 6 years of schooling
	Education3	7 to 12 years of schooling
	Education4	more than 12 years of schooling
Age	Child	younger than 12 years
	Teenager	13 to 17 years old
	Adult	18 to 65 years old
	Elders	older than 65 years

4.2.2 Expenditures and *autoconsommation*

The survey also contains a wide number of expenditure variables. Two main groups of expenditure are distinguished: consumption and other expenditures. Consumption has been aggregated in 10 categories following the ICP commodity aggregation as in Reimer and Hertel (2004). Table 3 shows this aggregation. Rural households show a lot of zeros for consumption of commodities 4 through 10, therefore for these households, commodities 4 to 10 are grouped in only one category.

Table 3: ICP Commodity Aggregation

Code and aggregate
1. Grains and other crops
2. Meat, dairy and fish
3. Processed food, beverages and tobacco
4. Apparel, footwear
5. Rent and housing utilities
6. Home furnishings and appliances
7. Medical products and services
8. Transport and communication
9. Recreation and education
10. Other goods and services

Source: Reimer and Hertel (2004).

Non consumption expenditures include taxes and transfers to other households.

4.2.3 *Autoconsommation*

According to Deaton (1997), the consumption of home-produced items, *autoconsommation*, is an issue that "tends to compromise the quality of consumption data in developing countries". Such items represent both an income and a consumption, and should be recorded in that way. Nevertheless, a major problem arises when assigning a value to home-produced items, especially when markets are not well developed.

The MECOVI survey also contains data on home-produced items, and asks households how much they would pay if they had to buy these items at the market. The MECOVI survey uses this information to value *autoconsommation*, and so do we. Thus, *autoconsommation* is considered in this work, and is recorded as both consumption and income.

We consider *autoconsommation* not only for food, but also for housing, where a rent is valued for households living in their own houses. The rent value is also obtained from respondents' own valuation.

Autoconsommation has shown to be very important, especially for food in rural households, and for housing in urban and rural households. Table 4 shows the importance of *autoconsommation* in the survey. It can be seen that the share for housing in the total is identical for both regions.

Table 4: Total consumption and *Autoconsommation* (in millions of Bolivianos)

	ICP Aggregate			
	1	2	3	5
All households				
Total consumption	6333	5753	2769	9148
Autoconsommation	871	622	152	6233
% of total	13.8	10.8	5.5	68.1
Urban households				
Total consumption	4692	4688	2063	8633
Autoconsommation	274	180	89	5881
% of total	5.8	3.8	4.3	68.1
Rural households				
Total consumption	1641	1065	706	516
Autoconsommation	598	442	62	351
% of total	36.4	41.5	8.8	68.1

Source: Author's calculations.

4.2.4 Income and savings

The MECOVI survey is very rich in income information. Income is collected in the survey at the individual level, which permits to identify all incomes by their nature (labor and non labor income), by the source of economic activity (formal or informal market), and by the type of human capital endowment.

Labor income includes income of both employees and self-employed workers. Income of employees includes wages, bonuses, premiums and in-kind payments; while income of self-employed workers is given by net income. Labor income of employees is considered as **formal labor**, while self-employed workers' income is assumed to belong to **informal labor**. [3]

Non labor income includes property rent and income from assets (interest and dividends), transfers from government, transfers from other households, transfers from the rest of the world and *autoconsommation*.

Labor income is disaggregated by sector of activity and by education level. Concerning the activity sector, the MECOVI survey classifies labor income according to the Bolivian Classification of Economic Activities (CAEB), which was used to classify labor income by activity sector as in the SAM, see Table 1. Moreover, within each sector of activity, labor income was assigned to four skill levels (Table 2), depending on the education level of the individual.

Finally, savings is calculated as the difference between income and expenditures.

Individual data is added up by household, and the resulting income generation structure for households is shown in Table 5.

4.3 Reconciliation between the survey and the SAM

In order to reconcile the survey with the SAM, household aggregates in the survey must add up to their corresponding values in the SAM. To achieve this reconciliation we have used the cross entropy method described in Robilliard and Robinson (2003) and other adjustment techniques described in Fofana and Cockburn (2003) and Cloutier and Cockburn (2002).

The cross entropy method proposed by Robilliard and Robinson (2003) relies on the assumptions that (i) income and consumption aggregates in the SAM and (ii) household data in the survey is measured accurately (or have been adjusted to be accurate). Robilliard and Robinson propose to adjust household weights in the survey so that the survey aggregates are consistent with the national accounts data. The cross entropy method consists in calculating new weights that are "close" to the prior weights, but satisfying the condition that the survey aggregates match the SAM data. Prior weights are the original weights of the survey.

[3] In the model, self-employed labor of farmers is attributed to land remuneration.

Table 5: Households' income generation

Labor	Formal sector by activity sector (1...35)	Education 1 Education 2 Education 3 Education 4
	Informal sector by activity sector (1...35)	Education 1 Education 2 Education 3 Education 4
Non labor	Asset income	
	Interest and dividends	
	Transferts	Government Other hhls ROW
	autoconsommation	Food Housing

Source: Author's creation.

More formally, the cross entropy method consists in minimizing a distance function between prior and new weights $f(\omega_1, ..., \omega_n, \overline{\omega_1}, ..., \overline{\omega}_n)$, where ω_i is the prior weight for household i, and $\overline{\omega}_i$ is the new weight for household i; subject to the consistency constraints $\overline{\omega}' x_t = y_t$, where $\overline{\omega}$ is the vector of new weights, x_t is the vector of observed characteristic t in the survey for all households, and y_t is the corresponding aggregated value, usually those found in the SAM.

In this case, the function to minimize is the Kullback-Leibler cross entropy measure of the distance, and the problem can be written as:

$$Min \sum_i^n \omega_i \log \frac{\omega_i}{\overline{\omega}_i}$$

subject to T constraints

$$\sum_i^n \overline{\omega}_i x_t = y_t; \ t = 1, ..., T$$

Before applying the cross entropy method to calculate new weights for households, a descriptive analysis of both data sources, the survey and the SAM, has first been done in order to measure the gap between the two sources. While the SAM and the survey show similar aggregates for wages, the gap is very important for consumption aggregates and income from capital. Both are much larger in the SAM. This can lead us to consider that rich households tend to underreport their income and consumption (Deaton, 1997; Akee, 2007), or, they simply have not participated in the survey or are underestimated in it.

Therefore, applying the cross entropy method in order to match the survey and the SAM aggregates, for both income and consumption, may lead to sharp changes from prior to new weights, and may lead to erroneous results if we consider that new weights are calculated to overcome erroneous information, or to fill the absence of some groups of households who have not participated in the survey.

In view of these considerations, we chose to keep wages aggregates from the SAM to reconcile the survey and the SAM, and to create a "residual" household whose income stems from capital income and dividends, and whose expenditure fills the gap of consumption between the two data sources. As discussed in the previous section, the behavior of this "residual" household is modeled in a neutral manner, thereby minimizing the influence of this data adjustment on the results of the simulation.

Other adjustments have been made in the SAM to impute missing values. We use survey ratios to disaggregate wages by sector, urban and rural, and by skill level in the four categories shown in Table 5. Survey ratios are also used to distinguish gross operating surplus and income from self-employment (informal sector), non-industrialized agriculture land and capital. Informal sector income is also disaggregated by skill level.

Several adjustments have also been made to the household survey data before applying the cross entropy method. First, for rural households, total income has been adjusted to equalize total expenditure. Indeed, consumption is a better measure for permanent income than current income because the income of rural self employed workers, who are mainly farmers, fluctuates with seasons. Consumption, on the other hand, has been shown to be more constant throughout the year. Moreover, since rural households are mainly poor, the assumption that income matches consumption through the year, thus there are no savings, seems quite realistic.

Urban households also show negative savings for several observations. Nevertheless the assumptions made for rural household, especially the one concerning dependance of income on seasons, are not likely to hold for them. Other reasons may apply to these households to explain negative savings; as temporary unemployment, or underemployment; or unusual expenditures in the year.

As the CGE model represents a long run equilibrium of the economy, it seems reasonable to assume that savings of representative households should not be negative in the base equilibrium. However, allowing a few individual households to have negative savings for one year is not necessarily problematic for the model, as long as the aggregate for each representative household in the CGE model is nonnegative. Representative households in the CGE are determined as follows. First we distinguish representative households by sector, urban or rural. Urban and rural households have been shown to have very different behavior in consumption, and factor income structure is also very different. Next, households in each sector are grouped by demographic structure. Characteristics which determine the demographic structure are: number of children, number of teenagers, number of adults and number of elders. The number of members in each age group has been limited, as shown in Table 6, in order to avoid extreme households and to limit the number of possible combinations.

Table 6: Representative households' demographic characteristics

Sector	Dem. Characteristic	Range
Urban	Number of children	0 to 3
	Number of teenagers	0 to 3
	Number of adults	0 to 4
	Number of elders	0 to 2
Rural	Number of children	0 to 4
	Number of teenagers	0 to 3
	Number of adults	0 to 4
	Number of elders	0 to 2

The resulting number of representative households for the urban and rural sectors are 164 and 174 respectively.

Once the representative households are defined, some adjustments are made to representative households where savings is negative. Since households belonging to a same group might have all of them negative savings, reconciliation would assign weights equal to zero to all households belonging to this group. Thus, before calculating new weights, households' income in each groups with negative aggregate savings are proportionally increased until the mean saving of the group equals zero.

At this point we have all the elements to proceed with the reconciliation through the cross entropy method. The resulting adjusted weights turn out to be quite close to the original weights. Finally, to complete the match between the SAM and the survey, a transition matrix between the commodities classification in the SAM and the classification in the survey has been constructed. This matrix has been created using information of "supply-use equilibrium correspondence" data from the INE Bolivia, and can be found in Zavaleta (2010).

4.4 Estimation of the Linear Expenditure Demand System (LES)

Let's consider equation (9). In order to account for heterogeneity in household behavior, parameters γ_{ig} and b_{ig} depend on demographic characteristics (k) of households. Demographic characteristics have been defined in previous section, and are further described in Table 6.

To construct the econometric equation of household consumption for its later estimation, we consider each characteristic (k) as qualitative variables, with m_k modalities each. Thus, γ_{ig} and b_{ig} are written as:

$$\gamma_{ig} = \gamma_{i0} + \sum_k \gamma_{ik} d_{kg} \tag{14}$$

$$b_{ig} = b_{i0} + \sum_k b_{ik} d_{kg} \tag{15}$$

where γ_{i0} and b_{i0} are respectively the committed consumption and the marginal consumption share of the reference household (a household with 2 adults and no other members, for instance); k is a demographic characteristic (number of adults, of children, of teenagers, of elders); γ_{ik} and b_{ik} are vectors of length $m_k - 1$, where each term of the vector is associated to a modality of variable k, other than the reference, and each value is the contribution of the associated modality in γ_{ig} and b_{ig} respectively. Finally, d_{kg} is the "modalities vector" of length $m_k - 1$, where each value is equal to 1 if the corresponding modality applies to group g, and 0 otherwise.

Introducing equations 14 and 15 in equation 9, for household h and commodity i we obtain:

$$\begin{aligned} p_i c_{ih} \quad = \quad & p_i \left(\gamma_{i0} + \sum_k \gamma_{ik} d_{kh} \right) \\ & + \left(b_{i0} + \sum_k b_{ik} d_{kh} \right) \left(y_h - \sum_j p_j \left(\gamma_{j0} + \sum_k \gamma_{jk} d_{kh} \right) \right) \\ & + p_i u^*_{ih} \qquad h \in H \end{aligned} \tag{16}$$

where $d_{kh} = d_{kg}$ if $h \in H_g$.

To estimate the parameters of equation 16, some constraints must be imposed. First, even in the absence of demographic components of parameters γ_i and b_i, cross sectional data isn't enough to estimate a LES demand system (Deaton, 1997), because no changes in prices are observed. We need additional information, for instance an elasticity, or another restriction to be able to estimate the parameters.

Second, the introduction of the demographic dependent parameters in the committed consumption parameter γ_i, also needs some restrictions on the term $\sum_i \gamma_{ik}$, for each demographic characteristic k.

If we consider equivalent expenditures rather than gross expenditures, using an equivalence scale, this second gap may be fulfilled. In this case, we can impose $\sum_i p_i \gamma_{ik} d_k = 0$ for all k. We can interpret this restriction considering that committed consumption is equal for all equivalent households, but distributed differently among commodities depending on the household's demographic structure.

The equivalence scale used in this study is the square root scale, which is a widely used scale. Besides, the size elasticity of this scale, 0.5, is close to the value found by Lanjouw and Ravallion (1995), 0.6, for which correlation between poverty and family size vanishes.

Concerning the lack of prices, another restriction must be introduced. Because of the lack of estimated elasticities for Bolivia, in this work we will use a two steps procedure to find the parameters of the demand system. The first step consists on estimating econometrically the parameter of the marginal propensity to consume, b_i, for each commodity i, using the OLS method. Then, to find the rest of the parameters, the total committed consumption is calibrated so that it satisfies the LES properties. This procedure is described in detail in the following paragraphs.

Since we have no information on prices in the survey, but only on expenditures, and considering the restrictions on demographic gammas discussed above, equation 16 can be written as follows

$$c_{ih}^p = \gamma_{i0}^p + \sum_k \gamma_{ik}^p d_{kh} + \left(b_{i0} + \sum_k b_{ik} d_{kh} \right) \left(y_h - \sum_j \gamma_{j0}^p \right) + \epsilon_{ih} \tag{17}$$

where $c_{ih}^p = p_i c_{ih}$, $\gamma_{i0}^p = p_i \gamma_{i0}$, $\gamma_{ik}^p = p_i \gamma_{ik}$ and $\epsilon_{ih} = p_i u_{ih}^*$.

To estimate the b_i parameters, an arbitrary value for the total committed consumption $\sum_i \gamma_{i0}^p$ is assigned, because the resulting estimated betas don't depend on it. The estimation is done trough weighted OLS, where each observation is weighted by its corresponding expansion factor in the survey. This ensures that the regression passes by the weighted means of the sample, making aggregated values coherent between the CGE model and the microsimulation model.

Obtained results allow negative betas for some households. Negative betas violate the negativity property of direct price effects of demand. In fact, the expression of the diagonal elements of the Slutsky matrix are

$$s_{ii} = \frac{\partial q_i}{\partial y} q_i + \frac{\partial q_i}{\partial p_i}$$

for the LES demand this expression is

$$s_{ij} = \frac{\beta_i(\beta_i - 1)}{p_i^2}\left(y - \sum_j p_j\gamma_j\right)$$

.

Thus, negativity of s_{ii} holds, if betas are non negative and smaller than 1; and if total expenditure y is greater than total committed expenditure, $\left(y - \sum_j p_j\gamma_j\right)$.

Keeping negative betas results in having demands that don't satisfy the negativity property of demand functions. Some adjustments have been done in order to avoid negative betas. For instance, aggregated commodity of medical products and services (ICP 7), has been removed from the estimations for urban households. This because several households show no consumption of this commodity, resulting in an important number of negative betas; besides, estimated parameters for this commodity are statistically non significant. Thus, the marginal consumption propensity for this commodity is set to zero.

For rural households, ICP commodities 4 to 10 present many zeros in the survey, and are aggregated in one category.

Tables A.4 and A.5 in the appendix show the estimated b_i parameters, for urban and rural households respectively. The reference is a household with two adults.

Some negative parameters b_i remain, but the number of households concerned is very small, less than 5%, and we consider this wont have a big effect in the results.

The next step consists in calibrating the γ_i parameters for each household. To do this we make two considerations. First, total expenditure must be at least equal to the sum of committed expenditures, to satisfy once again the negative property discussed above. Second, resulting price elasticities must be close to estimated price elasticities for countries of the region. The elasticities used as indicators, are those found in Dimaranan et al. (2002), for Peru and Colombia.

The calibration is made by fixing different values for total committed consumption until finding a compromise between both conditions. From equation 17, we have for any household h and for commodity i:

$$\gamma_{ih}^p = \gamma_{i0}^p + \sum_k \gamma_{ik}^p d_{kh} = c_{ih}^p - \left(b_{i0} + \sum_k b_{ik} d_{kh}\right)\left(y_h - \sum_j \gamma_{j0}^p\right) \tag{18}$$

Replacing b_{i0} and b_{ik} by its estimated values, and c_{ih}^p and y_h by the observed values in the survey, the estimated value of γ_{ih}^p for household h becomes:

$$\hat{\gamma}^p_{ih} = c^p_{ih} - \left(\hat{b}_{i0} + \sum_k \hat{b}_{ik} d_{kh}\right)\left(y_h - \sum_j \gamma^p_{j0}\right) \tag{19}$$

Estimated errors of the OLS regression are implicitly introduced in equation (19), since $\hat{\gamma}^p_{ih}$ can also be obtained from the regression estimated parameters.

$$\hat{\gamma}^p_{ih} = \hat{\gamma}^p_{i0} + \sum_k \hat{\gamma}^p_{ik} d_{kh} + \hat{\epsilon}_{ih} \tag{20}$$

Thus, all households have different values for $\hat{\gamma}^p_{ih}$. Errors are interpreted in this case as unobserved characteristics of each household, in the allocation of committed consumption. It can be shown that $\sum_i \hat{\epsilon}_{ih}$ equals 0 for all household h.

The first value assigned to total committed consumption is, arbitrarily, one half of the median total expenditure of the survey. This first value gives many households not satisfying the condition of total consumption higher than total committed consumption.

We decrease this value until the first consideration is satisfied. Price elasticities are closer to the reference.

The values of total committed expenditure kept for the rest of this work are 4000 bolivianos for urban households, and 800 bolivianos for rural households. Sensibility analysis shows small reaction on price elasticities for values around those retained.

Finally, grouping households by demographic characteristics ($d_{ih} = d_{ig}$, and thus $b_{ih} = b_{ig}$), equation (9) for household $h \in H_g$ is replaced by:

$$p_i c_{ih} = p_i \hat{\gamma}_{ih} + \hat{b}_{ig}\left(y_h - \sum_j p_j \hat{\gamma}_{jh}\right) \tag{21}$$

where $\hat{b}_{ig} = \hat{b}_{i0} + \sum_k \hat{b}_{ik} d_{kg}$ and $p_i \hat{\gamma}_{ih} = p_i \hat{\gamma}_{ig} + p_i u^*_{ih}$. Equation (11) still holds to obtain the demand function for representative households g, since the normalization constraint (10) is satisfied. It is important to note that we don't need prices to estimate and calibrate consumption for households and representative households.

5 Illustration: hydrocarbons, a booming sector in Bolivia.

Since the late 1990s and mainly during the last decade, hydrocarbons have become the most important and dynamic product in the economy. Its increase in reserves, production and exports, and also the increase of its price, makes the hydrocarbons the main exported product (Table 7).

Increase in gas reserves has been accompanied by an increase of the investment in the sector, which shifted its production to much higher levels. After the *capitalization* process in the mid-1990s, cumulated foreign direct investment in hydrocarbons sector between 1995 and 2003 was almost US$ 3.5 billion, about 10% of the average GDP for that period (Cerutti and Mansilla, 2008). Table 3.1 shows that production of the sector has increased by a multiple of five in the selected period, and exports by a multiple of ten in volume. Combining the increase in volume with the increase of international prices, the value of exports has increased from 36 million US$ in 1999 to almost 3.2 billion US$ in 2008. This represents an increase by a multiple of 87.

Its share in total exports has increase from 3.4% of total exports to more than 45%, and from less than 1% to more than 20% of GDP. Exports of non traditional products have almost remained constant for the first half of the period. At the end of the period they emerge again, part of it because of the increase in soja prices.

The government has also benefitted from this boom. Its income from the hydrocarbons sector rose from 385 US$ to 1777US$ (Table 8), coming from royalties and other taxes on hydrocarbons; making the government heavily dependent on the sector. About one third of its total income comes from hydrocarbons, nevertheless this information differs across data sources. In fact, since 2006 the hydrocarbons sector has been nationalized, and data sources differ on the share of the hydrocarbons company (YPFB) incomes received by the government. Public deficit has also reduced and even changed sign for 2006 and 2007.

Table 8 also shows that net international reserves (NIR) increase from 1.1 US$ billions to more than 7.7US$ billions. In fact most of hydrocarbons income has been used to increase the net international reserves (NIR) and to increase public savings, which has sterilized the spending impact of natural gas income (Cerutti and Mansilla, 2008).

Table 7: Evolution of the gas sector

Year	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
Natural gas proven reserves[1]	5	18	24	27	29	28	27	n.a.	n.a.	n.a.
Proven and possible reserves[1]	14.1	49.8	70	77.2	79.1	76.4	63.9	n.a.	n.a.	n.a.
Total gas production[2]	92.2	127	186.3	226.7	261.3	362.2	442.7	474.4	505	526
Gas exports[2]	36.2	74.2	129.9	133.2	178.3	297.1	367	394.1	418.8	427.1
Average gas prices[3]	0.98	1.64	1.84	2	2.19	2.09	2.96	4.23	4.71	7.4
Total value of exports[4]	1042	1246	1226	1320	1590	2195	2867	4088	4822	6899
Natural resources[4]	472	604	644	693	875	1308	1989	3122	3682	5473
Of which gas exports[4]	36	121	239	266	390	620	1087	1668	1971	3159
Other products[4]	570	643	583	627	715	887	879	967	1140	1426
Gas exports as % of tot. exports	3.4	9.7	19.5	20.2	24.5	28.2	37.9	40.8	40.9	45.8
Gas exports as % of GDP	0.6	1.8	3.9	4.4	6.3	8.8	13.5	17	17.1	20.6

[1] Trillions of cubic feet.
[2] Billions of cubic feet.
[3] $us per 1000 cubic feet.
[4] Millions of current US$.

Source: INE Bolivia.

Table 8: Importance of the gas sector in government income (in millions of $us)

Year	1999	2000	2001	2002	2003	2004	2005	2006[1]	2007[1]	2008[1]
Gov. income from hydrocarbons	384.8	429.6	411.9	364.6	370.1	438.8	857.7	1337.3	1493.5	1777.3
% on total Gov. Income	17.9	20.3	20.8	19.3	19	18.9	29.4	35.4	34.6	32.3
Royalties	109.1	194.2	214.2	181.6	230.4	294.2	335.1	395.5	424.6	501.5
Tax on hydrocarbons	70.8	15.3	0.3	0	0	0	288.4	690.6	763.4	924
VAT, TT and IDH[2]	204.9	220.1	197.4	183	139.7	144.6	234.3	251.2	305.5	351.9
Public savings	-1632	-2035	-3902	-5044	-4779	-3947	-1779	3154	2262	-12
NIR	1114	1085	1077	854	976	1123	1714	3178	5319	7722

[1] Government from hydrocarbons differs from one source to the other from 2006 and later, year of the "Nationalization Decree". Some sources give larger values for these three years.

[2] VAT = Value added tax, TT = Tax on transactions and IDH = Direct tax on hydrocarbons.
Source: INE Bolivia, UDAPE.

Source: INE Bolivia.

The roll played by the hydrocarbons sector in Bolivia in the last ten years is quite important, and although the country has undergone several social and political conflicts because of it, the economic question of how to use this resource to improve Bolivia's economic situation persists. So far, a big share of its income has been confined to increase net international reserves, and to implement some redistribution policies in the form of transfers to the population, such as the "Bono Dignidad" and the "Bono Juancito Pinto", to elders and children respectively. The effects of the hydrocarbon sector boom, and of the implemented redistribution policies, on Bolivians' welfare also remain as a study subject.

5.1 The Renta Dignidad

Since 1997, under the presidency of Gonzalo Sánchez de Lozada, a transfer to people older than 65 years has been introduced. It was first named "Bonosol", and has changed its name and its conception since. Recently, with the government of Evo Morales, it has changed its name to "Renta Dignidad", and is payed to people older than 60 rather than 65.

When created, the Bonosol had to be paid through a fund created with the capital received from the privatized companies. This has also changed. Recently, president Evo Morales has nationalized some of those companies, like the telecommunications and the oil companies. Oil companies were though nationalized in part, and direct tax on oil and gas extraction, called IDH has been fixed. Now, resources to pay the current Renta Dignidad come from the IDH.

In spite of its positive reception among the population, several questions have arisen about the benefits and the sustainability of this transfer. Hugo Banzer, who has succeeded Sánchez de Lozada in the presidency of the country, has suppressed the "Bonosol" because of the lack of resources to pay it. Later, he has created the "Bolivida", a more modest version of the "Bonosol". When Sánchez de Lozada was elected president again in 2002, the "Bonosol" was established again as it was originally. During the current government, as mentioned before, president Evo Morales has replaced the "Bonosol" with the "Renta Dignidad".

5.2 The Bono Juancito Pinto

The National Development Plan of the government, implemented in 2006 the Social Protection and Communitarian Development policy. Its objective was to eradicate extreme poverty and social exclusion. In the frame of this policy, the Bolivian government created the "Bono Juancito Pinto", a transfer destined to children of public elementary schools. Its objectives are to incentive the school enrollment and to reduce children desertion from school.

This transfer started paying 200 Bs. (about 28 US$) to all students enrolled in public schools, from first to fifth grade, and who have not deserted, once a year at the end of the academic period. At present, the target population has expanded, and the transfer is made to students from first to eight grade.

The transfer is mainly financed by the direct tax on hydrocarbons (IDH). The total amount of the transfer has passed from Bs. 220 millions in 2006 to Bs. 360 millions in 2008 (YPFB).

The "Renta Dignidad" and the "Bono Juancito Pinto" are not the only transfers in Bolivia. The "Bono Juana Azurduy", partly financed by the World Bank, was created as an incentive to pregnant women to attend their regular medical controls, and has as target to reduce infant mortality. There are also several projects to create new transfers, one of them to handicap people.

6 Simulations and results

This section presents the simulations and the results from the CGE and the microsimulation models.

Five simulations are considered in this work and are labeled as follow: **SIM1** introduces a shock of a 60% increase on extracted natural gas, **SIM2** and **SIM3** raises the tax on natural gas to the level of the base scenario (25%), and transfers government income from this tax to elders and to children respectively, **SIM4** and **SIM5** increases the tax on natural gas to 80% and distributes government surplus to elders and children respectively.

The simulated 60% increase on the volume of extracted gas is less than the actual increase in the last decade, but we don't seek to describe the evolution of the gas sector and the economy during the last decade. We rather want to find the mechanisms through which a natural resource boom affects poverty and inequality. Thus, we choose this increase of 60%, which corresponds to the increase between the year 2000 and somewhere between the years 2001 and 2002 (Table 7).

Concerning the taxes on natural resources, **SIM1** decreases the tax because of the closure of the CGE model. The government closure in our study fixes the deficit while taxes are assumed flexible. **SIM2** and **SIM3** restores tax to the base scenario level, and SIM4 and SIM5 rices taxes on natural resources to 80%. This is consistent with the nationalization process of the gas sector.

The simulations are summarized in Table 9.

6.1 Main results

The aggregate results of the five simulations are shown in Table 10.

Table 9: Simulations description.

Simulation	Shock
SIM1	Increase of 60 % on natural gas reserves
SIM2	SIM1 + a transfer to elderly people
SIM3	SIM1 + a transfer to children
SIM4	SIM1 + increase on natural gas taxes to 80% + higher transfer to elders
SIM5	SIM1 + increase on natural gas taxes to 80% + higher transfer to children

Table 10 Aggregate Results (percentage change respect to the *base* simulation).

Change in	SIM1	SIM2	SIM3	SIM4	SIM5
GDP	1.9	1.9	1.9	2.1	2.0
Households consumption	1.8	1.9	1.9	2.4	2.4
urbans	2.1	2.3	2.3	3.3	3.2
rurals	-0.1	0.7	0.7	4.5	4.5
residual	1.9	1.2	1.1	-2.4	-2.5
Investment	1.0	0.6	0.6	-1.9	-2.0
Government income	1.7	2.8	2.8	8.1	8.2
Exports	3.9	3.8	3.8	3.3	3.4
Oil and gas	86.4	86.3	86.3	86.2	86.2
Mining	-35.0	-35.4	-35.3	-37.4	-36.8
Food products	-4.2	-4.2	-4.2	-4.5	-4.5
Base metals	-47.6	-48.1	-47.9	-50.3	-49.3
Industrialized agr.	1.1	1.1	1.1	1.3	1.2
Textiles	-8.1	-8.1	-8.1	-8.1	-8.2
Factors prices					
formal urb 1	-4.8	-4.9	-4.8	-5.3	-5.2
formal urb 2	0.4	0.4	0.4	0.1	0.2
formal urb 3	0.4	0.3	0.4	0.1	0.3
formal urb 4	4.1	4.1	4.1	3.9	4.0
informal urb 1	0.5	0.4	0.4	0.2	0.4
informal urb 2	-0.6	-0.7	-0.7	-0.9	-0.7
informal urb 3	1.2	1.2	1.2	1.0	1.2
informal urb 4	2.2	2.2	2.2	2.0	2.1
formal rur 1	-5.7	-5.8	-5.7	-5.8	-5.6
formal rur 2	-2.2	-2.3	-2.3	-2.7	-2.5
formal rur 3	-0.3	-0.4	-0.3	-0.6	-0.4
formal rur 4	3.1	3.1	3.1	2.9	3.1
informal rur	-1.3	-1.3	-1.2	-1.2	-1.0
capital	0.0	0.0	-0.1	-0.1	-0.2
land non ind	-0.1	0.0	0.1	0.5	0.7

Continued on next page

Table 10 – continued from previous page

Change in	SIM1	SIM2	SIM3	SIM4	SIM5
land ind	-4.9	-4.9	-4.9	-5.2	-4.9
natural resource	-12.3	-12.3	-12.3	-12.6	-12.5
Real exchange rate	-1.1	-1.2	-1.2	-1.3	-1.3

As mentioned in the model description, natural resources are modeled as a production factor in the top of the nested production function, along value added and intermediate inputs.

GDP increases about 2% in all five simulations. Private consumption increases even in the absence of transfers, about 1.9% for the first three simulations, and 2.4% for the last two simulations. The residual household, which receives its income only from capital and profits from enterprizes, has its consumption lowered in the last two simulations. The residual household accounts for about 18% of total household consumption.

Investment first increases, but transfers pull it down. Investment in the last two simulations are even below the base simulation. Government income, increases of 1.7% in the absence of transfers, and it reaches 8.2% in order to pay transfers in the last two simulations.

The real exchange rate appreciates in all simulations. We assume that capital required to export new gas reserves comes from the rest of the world thus, remuneration to this capital is paid abroad. This specification has two main consequences: First, capital inflow neutralizes the effects of the increasing demand for capital of the gas sector, such as an increase on its price and a migration of capital form other sectors. Second, capital remuneration outflow depreciates real exchange rate. We might believe that without this specification, appreciation of real exchange rate would be higher.

In spite of the large increase in gas exports (more than 86%), the overall export level increases only a little, about 3.9% for the first three simulations and 3.4% for the last two. This because almost all other exports decrease, specially mining and base metals sectors, suggesting Dutch disease effects. The structure of exports in the base scenario and in the different simulations are shown in Table 11.

As we can see, the structure of exports is also affected. The share in total exports of the oil and gas sector increases from 22.3% to more than 40%. Mining diminish from 12% to about 7.4%. Base metals from 5.2% to 2.5%. For other sectors, changes are much smaller. Changes in activities production are shown in Table A.6 in the Appendix.

Table 11: Export's Structure (In Percentage of Total Exports)

	BASE	SIM1	SIM2	SIM3	SIM4	SIM5
Oil and gas	22.3	40.1	40.1	40.1	40.3	40.2
Food products	14.9	13.8	13.8	13.8	13.8	13.8
Mining	11.9	7.5	7.4	7.4	7.2	7.3
Commerce	7.0	6.3	6.3	6.3	6.3	6.3
Base metals	5.2	2.6	2.6	2.6	2.5	2.5
other manufactures	4.8	0.0	0.0	0.0	0.0	0.0
Industrialized agriculture	4.7	4.6	4.6	4.6	4.6	4.6
Transport	4.2	4.0	4.0	4.0	4.0	4.0
Textiles	4.0	3.5	3.5	3.5	3.5	3.5
Wood and wood products	3.5	3.1	3.1	3.1	3.1	3.1
Others	17.5	14.7	14.7	14.6	14.6	14.6
Total	100.0	100.0	100.0	100.0	100.0	100.0

In Table 10 it is shown that factors prices has changed after the shock, being skilled labor the most beneficed and unskilled labor the harmed sector in urban and rural sectors. Considering that household endowed with skilled labor factor are wealthier than households endowed with unskilled labor, this change in factors prices can explain at least part of the increase of total consumption; nevertheless it also suggests an increase in income inequality, as is confirmed hereafter.

6.2 Inequality and poverty analysis

We have presented the aggregate results of the five simulations. Although results show an increase in GDP, overall households' consumption, exports and government income after a natural resource boom (SIM1 in Table 10); results also show changes in factors prices suggesting an increase in inequality. For example, unskilled labor wages in the formal sector decreases for both the urban and rural regions, and skilled labor wages increases in all sectors.

Thus, it is important to analyze the changes in inequality after a natural resource boom. If inequality indeed increases a redistribution policy might be desirable to neutralize it. We have shown two examples of social transfers carried out in Bolivia: the "Renta Dignidad" and the "Bono Junacito Pinto", which are transfers to the elderly and to the children respectively.

Next, since poverty reduction is a major goal for Government and international institutions, the analysis of the effects on poverty of redistribution policies following a natural resource boom is also pertinent.

This section seeks to determine the effects of a boom in the natural gas sector on inequality and poverty at a disaggregated level. To do this, we use the results obtained in the microsimulation model to obtain the equivalent variation for each household and for each simulation. We also analyze the consequences of a redistribution policy to compensate for the potential harmful effects of a natural resource boom on equity and poverty.

Changes in individual welfare is measured using the equivalent variation concept. For each scenario j, the individual welfare for household h is given by:

$$y_h^j = y_h^0 + EV_h^j \tag{22}$$

where y_h^0 is the expenditure of household h in the base scenario, and EV_h^j is the equivalent variation of household h in the j scenario respect to the base scenario. Savings, taxes and transfers from households are considered fixed in this model, so there is a direct relation between income and household consumption and in this section we use both terms indistinctly.

We also adjust income applying and square-root equivalence scale, which means that household income is divided by the square root of the size of the household. Each household is also weighted by the same scale for the aggregation (See Cowell (1984) and Zavaleta (2010) for a detailed discussion about this topic).

6.2.1 Inequality

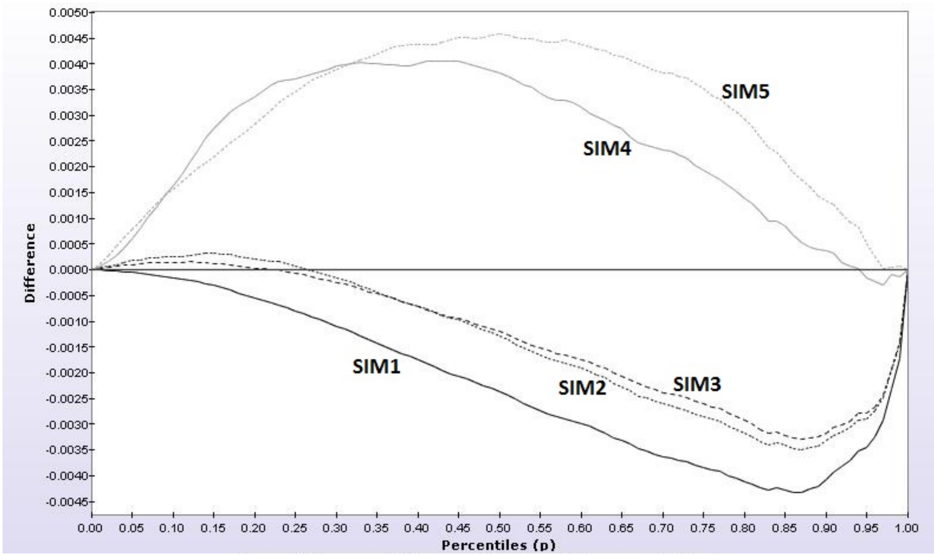
Comparison of inequality for the five scenarios is presented through differences of Lorenz curves for the five simulations with respect to the base scenario. Simulations SIM1 to SIM5 are those presented in Table 9.

Lorenz curves are very important for characterizing the robustness of inequality measures (Deaton, 1997). Whenever two Lorenz curves do not cross, the upper curve represents a more egalitarian distribution without ambiguity, and will show a lower inequality level for any inequality index that respects the principle of transfers of Pigou and Dalton. In fact, the upper Lorenz curve can be obtained from the lower one by a series of transfers from richer to poorer individuals, thus reducing inequality, which is the the principle of transfers. When two Lorenz curves do not cross we say that the upper curve Lorenz dominates the lower one.

In this section we calculate difference between Lorenz curves for all simulations respect to the base scenario, instead of the Lorenz curves themselves. We also present the main inequality indices to analyze inequality changes whenever there is no Lorenz dominance.

Our first result is that an increase in natural resources increases inequality. This result is shown in Figure 4, where the difference between Lorenz curves for simulation SIM1 and the base scenario (bottom curve) is negative everywhere.

Figure 4: Difference of Lorenz curves between simulations and the base scenario (PEAE-EA combination).



Source: Author's creation.

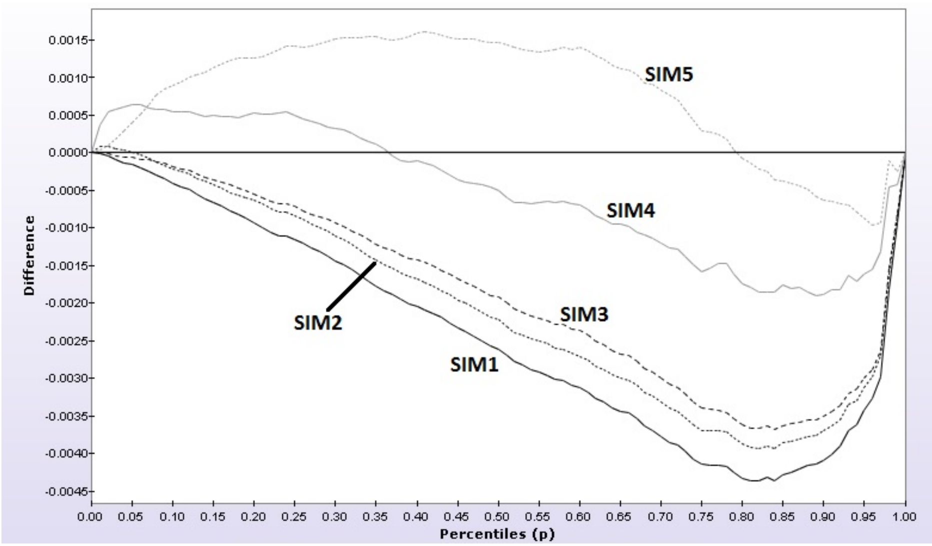
The increase in inequality is mainly due to the changes in prices of factors, as presented in Table 10. As shown, unskilled labor remuneration drops in almost all sectors and regions, while skilled labor wages increase for both the formal and informal sectors. A redistribution policy seems necessary to compensate for this increase in inequality.

Inequality decreases with respect to SIM1 when applying a transfer equal to government income from natural resources (SIM2 and SIM3). Figure 4 shows both SIM2 and SIM3 curves above SIM1 meaning that they both Lorenz dominates SIM1; nevertheless, the increase in inequality generated by the natural resource boom is not compensated for by these transfers. An increase on taxes to raise transfers to the actual current values, eliminates the inequality generated by SIM1 only when the transfer is made to the children (SIM5). In the case of the transfer to the elderly (SIM4), the curve lies upon the horizontal zero line up to the 90th percentile.

Although SIM5 is the only distribution that Lorenz dominates the base scenario, it does not dominates the curve of SIM4 because they cross twice in the first half of the distribution. We will compare inequality indices for both simulations further in this section.

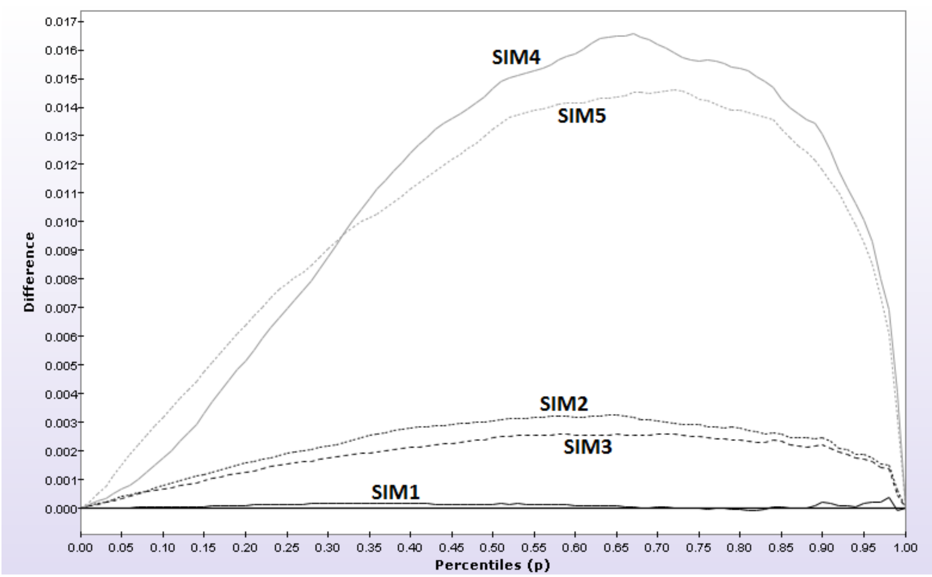
Results for urban and the rural regions are presented separately in Figures 5 and 6. They show that the natural resource boom more greatly affects urban households than rural households. A consequence is that for rural households any transfer decreases inequality with respect to the base scenario, while for urban households no transfer compensates completely for the increase in inequality created by the natural resource boom.

Figure 5: Difference of Lorenz curves between simulations and the base scenario. Urban households.



Source: Author's creation.

Figure 6: Difference of Lorenz curves between simulations and the base scenario. Rural households.



Source: Author's creation.

Figures 5 and 6 also show that for urban households, a transfer to the children Lorenz dominates a transfer to the elderly for most of the distribution except at the very beginning, where the curve of SIM4 is above the curve of SIM5. For rural households the opposite happens: SIM5 dominates SIM4 up to the percentile 30 and then it is SIM4 that dominates. It suggests that elderly people predominate in the very poor households in the urban region, while children predominate in the very poor households in the rural region.

Since curves for SIM4 and SIM5 cross each other for both urban and rural households, we can not conclude about which transfer is better for reducing inequality. We pass then to analyze the inequality measures for both scenarios.

To compare inequality indices, three indicators were chosen: Gini, Theil and Atkinson. For the latter, three different parameters of inequality aversion are considered. Table 12 shows all the indices for the base scenario for the whole population, and also for the urban and the rural regions separately.

Table 12: Inequality indices.

Index	Comb.	BASE	SIM1	SIM2	SIM3	SIM4	SIM5
Gini	All	0.488	0.493	0.491	0.491	0.483	0.482
	Urban	0.418	0.423	0.422	0.422	0.419	0.416
	Rural	0.476	0.476	0.472	0.473	0.455	0.456
Theil	All	0.440	0.449	0.446	0.446	0.433	0.431
	Urban	0.325	0.333	0.332	0.332	0.327	0.324
	Rural	0.461	0.461	0.454	0.455	0.425	0.427
Atk. ϵ=0.5	All	0.197	0.200	0.198	0.199	0.192	0.191
	Urban	0.142	0.146	0.145	0.145	0.143	0.142
	Rural	0.193	0.193	0.189	0.190	0.1773	0.1774
Atk. ϵ=1	All	0.355	0.360	0.356	0.357	0.343	0.342
	Urban	0.252	0.257	0.256	0.256	0.251	0.250
	Rural	0.332	0.331	0.325	0.326	0.305	0.304
Atk. ϵ=2	All	0.594	0.599	0.590	0.592	0.568	0.565
	Urban	0.414	0.420	0.414	0.418	0.402	0.408
	Rural	0.525	0.524	0.512	0.514	0.488	0.479

Table 12 shows that for the entire population, income distribution for SIM5 has lower inequality values for all the indices considered. This leads us to prefer a transfer to children rather than a transfer to the elderly. Nevertheless, when considering the urban and the rural regions separately, results differ depending on the considered inequality indicator. For urban households for instance, a transfer to the children (SIM5) reduces inequality more than a transfer to elderly (SIM4), for all indices except for the Atkinson index when $\epsilon = 2$.

For the rural households, inequality indicators for SIM4 are lower than indicators for SIM5, except for the Atkinson index when $\epsilon = 1$ and when $\epsilon = 2$.

6.2.2 Poverty analysis

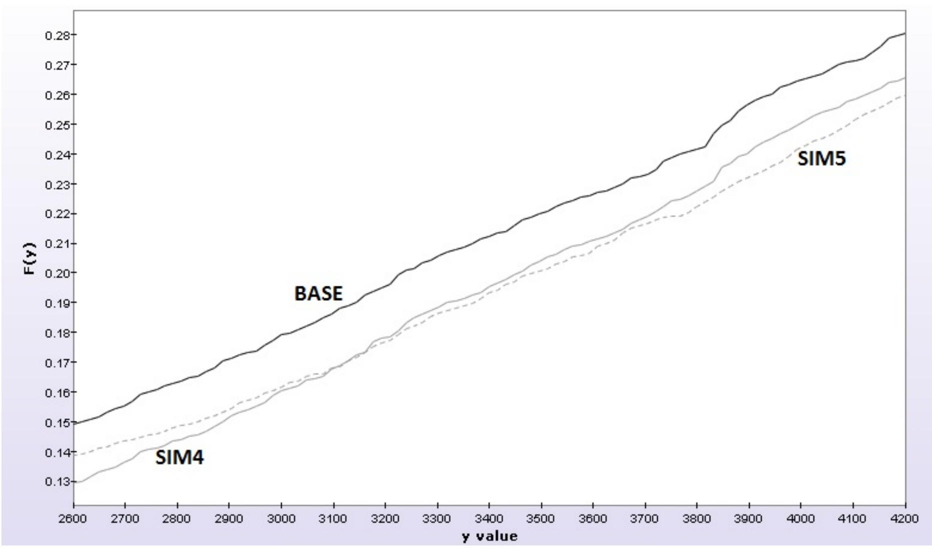
To analyze poverty, we apply the stochastic dominance analysis, or ordinal analysis, in a restricted version[4]. For the poverty analysis, simulations SIM1, SIM2 and SIM3 have shown being very close to the base scenario. Thus, we have limited the ordinal analysis to compare the base scenario to simulations SIM4 and SIM5 to avoid cluttering in the figures.

A major problem when aggregating urban and rural households is that they have different poverty lines, making it difficult to compare poverty between two distributions. Therefore, besides presenting the results for the whole population, urban and rural households' results are also presented separately.

Main poverty indicators are also calculated. These indicators include the generalized Foster-Greer-Thorbecke index for different poverty aversion parameters and the Sen index. The poverty lines considered in this work are those presented in Landa (2002) for the year 2000, which are 3955 and 2780 Bolivianos respectively for the urban and the rural regions.

Figure 7 shows the distribution functions, limited to values close to the poverty lines mentioned above.

Figure 7: Poverty incidence.



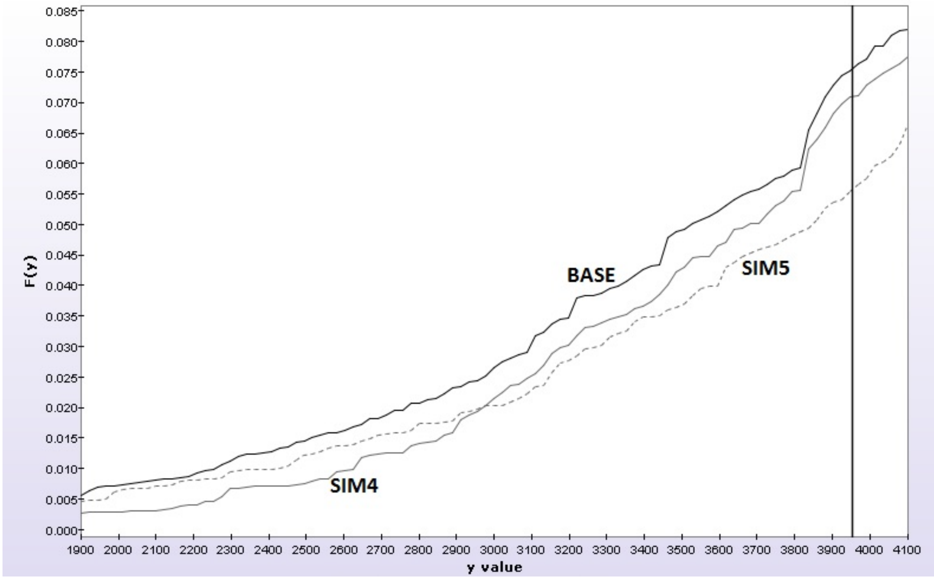
Source: Author's creation.

The curves of simulations SIM4 and SIM5 dominate the base scenario reducing the poverty incidence of about 2% depending on the poverty line. Nevertheless, they cross each other in the selected range.

[4]We call it restricted version as in Deaton (1997), because the stochastic dominance analysis is held over a relevant range of incomes, and not over the whole range.

If we consider the poverty lines for each region, Figure 8 shows that a transfer to children (SIM5) is much more effective reducing poverty for the urban region, while for the rural region, Figure 9 shows that a transfer to the elderly (SIM4) is more effective. Nevertheless, neither for the urban nor the rural households the dominance is absolute in the considered intervals. For instance, for urban households a transfer to children might be more effective in fighting poverty, but a transfer to the elderly seems more effective in fighting extreme poverty.

Figure 8: Poverty incidence curves. Urban region.



Source: Author's creation.

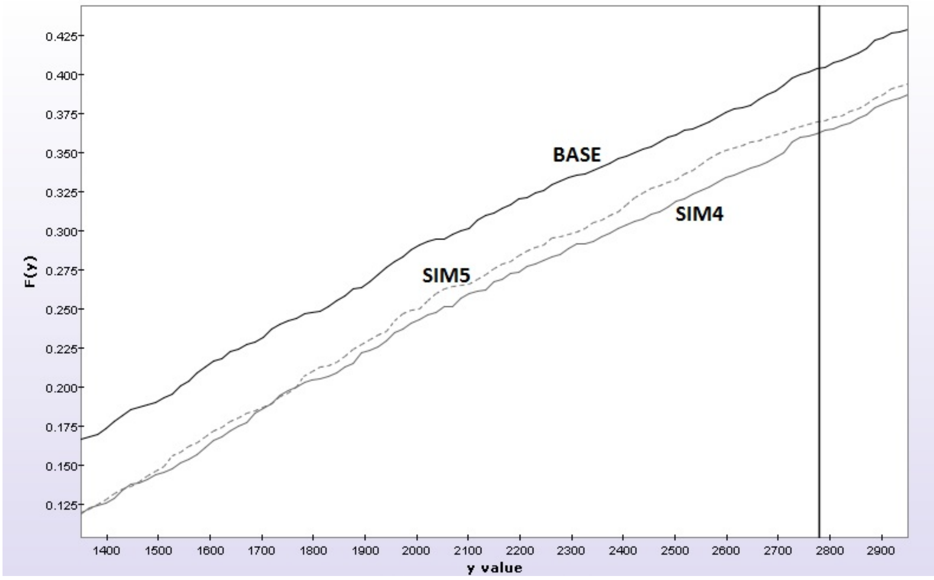
Finally, Table 13 shows the results for the different poverty indices.

7 Conclusions

This study has created an integrated CGE-microsimulation model for Bolivia as a framework to analyze the impact economic policies on poverty and inequality. The model applies exact aggregation conditions for consumption to create representative households in the CGE model that behave exactly as the aggregate of the households they represent. We call this model exact aggregated representative household (EARH) model.

The EARH model was applied to analyze the impact of an increase in natural resource output on poverty and inequality, and to compare different redistribution policies. The model has proved to be an adequate instrument for this type of analysis. It is a model that considers household heterogeneity, and

Figure 9: Poverty incidence curves. Rural region.



Source: Author's creation.

Table 13: Poverty indices.

Index	Comb.	BASE	SIM1	SIM2	SIM3	SIM4	SIM5
P_0	All	0.197	0.198	0.195	0.194	0.179	0.172
	Urban	0.076	0.076	0.076	0.073	0.071	0.055
	Rural	0.404	0.406	0.399	0.400	0.364	0.370
P_1	All	0.074	0.074	0.072	0.072	0.061	0.062
	Urban	0.016	0.016	0.016	0.016	0.013	0.013
	Rural	0.173	0.173	0.167	0.168	0.142	0.144
P_2	All	0.039	0.039	0.037	0.037	0.030	0.030
	Urban	0.006	0.006	0.006	0.006	0.004	0.005
	Rural	0.095	0.095	0.090	0.091	0.074	0.072
Sen	All*	0.101	0.101	0.098	0.098	0.084	0.082
	Urban	0.024	0.024	0.023	0.023	0.019	0.019
	Rural	0.226	0.227	0.218	0.220	0.189	0.188

* Sen indicator is obtained transforming rural expenditures to equalize its poverty line to the urban's poverty line.

in contrast to the top-down and the top-down bottom-up models, the EARH model can be solved easily in one step. Moreover, reconciliation between the national accounts and the survey has permitted not only to disaggregate the social accounting matrix (SAM) in several representative households, but has also permitted to disaggregate the production factors, especially the labor by education level.

Concerning the considered simulations, although a natural resource boom increases GDP, the results show several side effects in the economy. At the aggregate level we can mention three of these effects. First, the natural gas sector crowds out the exports of the other sectors, which may create an overdependence of the economy on natural resources. An appreciation of the real exchange rate is also observed in the simulations. This appreciation together with the deterioration of the other exporting sectors are symptoms of "Dutch disease".

Second, there are very important and unequal changes in factor prices. Skilled labor benefits from the natural gas boom, while unskilled wages decrease. This result suggests an increase in inequality that is confirmed when analyzing the results of the microsimulation model. Third, there is a contraction in the output of the other sectors, mainly because of the migration of factors specific to the gas sector. This is the other mechanism described in the Dutch disease literature that explains how a natural resource boom may harm the other sectors in the economy.

The microsimulation model has shown to be very useful for this study. It has permitted to analyze the effects of the simulated shocks on inequality and poverty at the household level, and to observe changes that are unobservable at the CGE model level. We apply stochastic dominance concepts for robust analysis of poverty and inequality.

Results of the microsimulation model confirm that the expansion of the natural gas sector increases inequality, but it is the urban region that is the most affected. This is due in part to the fact that the gas sector is intensive in urban skilled labor in the formal sector, and the natural resource boom increases its wage. Two different redistribution policies using government income from the gas sector have been compared: a transfer to the elderly and a transfer to children. The transfers modeled in this study have positive effects in reducing inequality, although they do not neutralize completely the increase in inequality created by the natural resource boom.

The EARH model developed in this study can be applied for the analysis of other economic policies. In this study we have used the model to describe the mechanism through which a natural resource boom can affect poverty and inequality. The model could easily be adapted to analyze other policies that might be considered in Bolivia, such as the creation of a new lithium industry in Bolivia, or the possible future reforms in international trade.

Finally, this work has applied exact aggregation conditions for household consumption, but this approach can also be applied in other contexts, for instance in the modeling of labor supply. A first step in this direction is undertaken by Magnani and Mercier (2009) who show how exact aggregation can be used in discrete choice models. This result could be a starting point for the future extensions of this work.

Table A.1: Studies using combined CGE-microsimulation models

	Type of model	Country	Simulation
de Janvry et al. (1991)	RH	Ecuador	Fiscal, monetary and exchange rate adjustments
Decaluwé et al. (1999)	RH	African country	Change in world prices and trade reforms
Chen and Ravallion (2004)	TD	China	Access to WTO
Robilliard et al. (2008)	TD-microeconometric	Indonesia	Financial crisis
Boccanfuso and Savard (2007)	TD	Mali	Cotton subsidies
Herault (2007)	TD	South Africa	Trade librealisation
Herault (2009)	TD	South Africa	Trade liberalisation [a]
Savard (2003)	TD/BU-microeconometric	Philippines	Trade reforms and wage increase
Bourguignon and Savard (2008)	TD/BU-microeconometric	Philippines	Trade reforms
Cockburn (2002)	FI	Nepal	Trade liberalization
Rutherford et al. (2005)	FI	Russia	Trade liberalization
Annabi et al. (2006)	FI	Senegal	Trade liberalization
Cororaton and Cockburn (2007)	FI	Philippines	Trade reforms
Rutherford and Tarr (2008)	FI	Russia	Access to WTO

[a] Herault (2009) compares a behavioral microsimulation model, as in Robilliard et al. (2008), and a more simple reweighting approach.

Table A.2: Value added structure of the SAM (in millions of Bolivianos)

	Value[a]	% of total
Wages	18648	39.4
Gross operating surplus	28553	60.3
Land in industrialized agriculture	430	0.9
Land in non-industrialized agriculture	1909	4.0
Natural resources	944	2.0
Self-employed and capital[b, c]	25270	53.3
Production taxes	165	0.3
Total value added	**47367**	**100.0**

[a] *Autoconsommation* is excluded.
[b] Section 3.2 explains it in more detail about self-employed workers.
[c] Capital includes land in non-industrialized agriculture. Its value is obtained when reconciling the SAM and the survey in section 3.3

Source: Author's calculations.

Table A.3: Social Accounting Matrix (SAM) for Bolivia

	Expenditures										
Receipts	**Activities**	**Commodities**	**Factors**	**Natural Resources (NR)**	**Households**	**Enterprises**	**Government**	**Bank Services**	**Savings Investment**	**Rest of the World (ROW)**	**Total**
Activites		Marketed output			Home-consumed output						Activity income (gross output)
Commodities	Intermediate inputs	Transaction costs			Private consumption		Government consumption	Bank Services absortion	Investment	Exports	Demand
Factors	Value-added									Factor income from ROW	Factor income
NR	NR income										NR income
Household			Factor income to households		Interhousehold transfers	Surplus to households	Tranfers to households			Transfer to households from ROW	Household income
Enterprises			Factor income to enterprises				Transfers to enterprises			Transfers to enterprises from ROW	Enterprises income
Government	Producer taxes, value-added tax	Sales taxes, tarifs, expor taxes	Factor income to government factor taxes	Royalties	Transfers to government, direct	Surplus to government, direct enterprise taxes				Transfer to government from ROW	Government income
Bank Services			Inputed Bank Services								Inputed Bank Services
Savings-Investment					Household savings	Enterprise savings	Government savings			Foreign savings	Savings
Rest of the World (ROW)		Imports	Factor income to ROW			Surplus to ROW	Government transfers to ROW				Foreing exchange outflow
Total	Activity	Supply expenditures	Factor expenditures	Royalties	Household expenditures	Enterprise expenditures	Government expenditures	Bank Services	Investment	Foreign exchange inflow	

Source: Lofgren et al. (2002), with modifications by the author.

Table A.4: Estimated parameters b_i for the LES (Urban households)

Variable	ICP 1	ICP 2	ICP 3	ICP 4	ICP 5	ICP 6	ICP 8	ICP 9	ICP 10
b_0	0.022	0.053	0.021	0.027	0.346	0.169	0.119	0.090	0.153
	(4.8)	(6.74)	(4.29)	(3.84)	(7.4)	(12.54)	(9.58)	(3.85)	(4.45)
Adults=0	0.002	-0.015	-0.001	0.000	0.032	-0.019	0.008	0.023	-0.031
	(0.1)	(-0.58)	(-0.06)	(-0.01)	(0.28)	(-0.43)	(0.3)	(0.54)	(-0.64)
Adults=1	-0.016	-0.034	0.003	0.006	0.086	-0.037	-0.008	-0.020	0.021
	(-2.82)	(-3.72)	(0.44)	(0.72)	(1.02)	(-2.33)	(-0.66)	(-0.66)	(0.3)
Adults=3	-0.003	-0.005	0.004	-0.002	-0.038	-0.026	-0.001	0.124	-0.053
	(-0.66)	(-0.52)	(0.54)	(-0.24)	(-0.87)	(-1.38)	(-0.09)	(2.41)	(-1.54)
Adults=4	-0.007	-0.018	0.000	0.005	0.048	-0.050	0.014	0.091	-0.082
	(-1.04)	(-1.54)	(0.04)	(0.48)	(0.63)	(-2.14)	(0.63)	(1.71)	(-2.84)
Children=1	-0.003	-0.012	-0.006	-0.003	0.006	-0.043	0.005	0.005	0.050
	(-0.54)	(-1.31)	(-1.21)	(-0.32)	(0.12)	(-2.99)	(0.44)	(0.13)	(1.15)
Children=2	0.025	0.034	0.029	0.025	-0.137	-0.050	0.036	-0.033	0.069
	(3.48)	(3.29)	(2.02)	(1.35)	(-2.69)	(-2.6)	(1.52)	(-1.19)	(1.69)
Children=3	0.044	0.047	0.040	-0.009	-0.100	-0.031	0.000	0.022	-0.013
	(5.11)	(1.85)	(1.11)	(-1)	(-2.07)	(-1.67)	(-0.02)	(0.62)	(-0.45)
Teenagers=1	0.017	0.025	0.008	-0.003	-0.009	-0.009	-0.016	0.049	-0.063
	(3.33)	(3.1)	(1.66)	(-0.31)	(-0.16)	(-0.48)	(-1.59)	(1.7)	(-1.86)
Teenagers=2	0.011	0.027	0.010	-0.010	-0.019	-0.024	0.032	0.051	-0.078
	(1.89)	(2.6)	(1.42)	(-1.27)	(-0.38)	(-1.39)	(1.18)	(1.66)	(-2.01)
Teenagers=3	0.033	0.038	0.039	0.019	-0.098	-0.013	0.036	-0.001	-0.053
	(3.28)	(2.64)	(1.36)	(0.62)	(-1.38)	(-0.59)	(0.75)	(-0.03)	(-0.94)
Elders=1	0.011	0.004	-0.005	-0.011	0.106	0.024	-0.025	-0.026	-0.080
	(1.56)	(0.38)	(-1.04)	(-1.47)	(1.36)	(0.98)	(-1.97)	(-0.7)	(-1.75)
Elders=2	0.005	0.004	-0.004	-0.021	0.256	-0.029	-0.053	-0.116	-0.042
	(0.63)	(0.26)	(-0.8)	(-3.02)	(3.18)	(-1.3)	(-2.29)	(-2.67)	(-1.04)

t-stat values are in parenthesis

Table A.5: Estimated parameters b_i for the LES (Rural households)

Variable	ICP 1	ICP 2	ICP 3	ICP 4-10
b_0	0.115	0.146	0.105	0.634
	(6.04)	(5.85)	(6.02)	(13.2)
Adults=0	0.142	0.076	0.004	-0.221
	(4.06)	(1.48)	(0.19)	(-2.5)
Adults=1	0.001	-0.011	-0.008	0.017
	(0.05)	(-0.39)	(-0.37)	(0.28)
Adults=3	0.021	0.044	-0.025	-0.040
	(1.12)	(1.69)	(-1.47)	(-0.87)
Adults=4	-0.001	0.015	0.137	-0.151
	(-0.07)	(0.63)	(5.83)	(-3.93)
Children=1	0.023	0.024	-0.018	-0.029
	(1.19)	(0.98)	(-1.38)	(-0.66)
Children=2	0.012	0.027	-0.044	0.005
	(0.55)	(0.82)	(-2.26)	(0.09)
Children=3	-0.041	-0.034	-0.062	0.137
	(-2.18)	(-1.61)	(-3.53)	(3.55)
Children=4	0.116	0.086	0.033	-0.235
	(3.99)	(2.72)	(0.88)	(-3.59)
Teenagers=1	-0.058	-0.066	0.015	0.109
	(-3.6)	(-3.27)	(0.96)	(2.69)
Teenagers=2	0.011	-0.011	-0.023	0.023
	(0.26)	(-0.29)	(-1.69)	(0.28)
Teenagers=3	0.040	-0.019	-0.048	0.026
	(1.3)	(-0.47)	(-1.1)	(0.26)
Elders=1	-0.009	0.004	-0.036	0.040
	(-0.35)	(0.12)	(-3.1)	(0.62)
Elders=2	-0.089	-0.027	-0.040	0.156
	(-2.86)	(-0.79)	(-2.71)	(2.26)

t-stat values are in parenthesis

Table A.6: Changes in activities production (percentage change respect to the *base* scenario)

Activities	sim1	sim2	sim3	sim4	sim5
Non-industrialized agriculture	-0.03	-0.02	-0.02	-0.02	-0.03
Industrialized agriculture	-0.75	-0.76	-0.75	-0.79	-0.75
Coca	1.82	1.79	1.79	1.66	1.64
Livestock production	0.44	0.53	0.54	0.97	1.03
Timber hunting and fishing	-2.35	-2.42	-2.43	-2.73	-2.81
Oil and natural gas	60.00	60.00	60.00	60.00	60.00
Mining	-32.82	-33.20	-33.12	-35.11	-34.61
Meat and processed meat	0.78	0.79	0.81	0.82	0.95
Dairy products	-0.21	-0.20	-0.18	-0.13	-0.03
Baking and grain mill products	-0.39	-0.40	-0.40	-0.46	-0.42
Sugar and confiture	0.61	0.57	0.61	0.35	0.64
Food products	-2.89	-2.93	-2.92	-3.13	-3.04
Beverages	0.76	0.71	0.76	0.47	0.79
Processed Tabacco	0.13	0.10	0.14	-0.10	0.17
Textiles	-2.82	-2.81	-2.81	-2.75	-2.79
Wood and wood products	-3.27	-3.25	-3.28	-3.14	-3.33
Paper and paper products	-0.03	-0.02	0.00	0.01	0.18
Chemical products	-1.65	-1.66	-1.65	-1.70	-1.67
Processed oil products	10.05	10.11	10.12	10.41	10.47
Non-metallic minerals products	-0.97	-1.22	-1.22	-2.50	-2.49
Base metals	-44.11	-44.55	-44.40	-46.78	-45.88
Machinery and equipment	-59.72	-59.87	-59.92	-60.60	-60.91
Other manufactures	-99.95	-99.95	-99.95	-99.96	-99.96
Electricity gas and water	1.98	2.03	1.95	2.29	1.83
Construction	1.25	0.78	0.77	-1.62	-1.68
Trade	-0.60	-0.62	-0.62	-0.72	-0.69
Transport and storage	3.34	3.41	3.42	3.74	3.81
Communications	-0.55	-0.48	-0.45	-0.17	0.01
Financial services	-1.09	-1.10	-1.10	-1.16	-1.18
Company services	8.85	8.83	8.85	8.70	8.80
Property	0.83	0.94	0.89	1.50	1.19
Local, social and personal services	1.38	1.42	1.46	1.62	1.83
Restaurants and hotels	1.36	1.30	1.35	1.01	1.27
Domestic services	2.99	3.42	3.33	5.62	5.07
Public sector	0.01	0.02	0.02	0.05	0.07

References

Akee, R.K.Q. (2007) 'Errors in self-reported earnings: The role of previous earnings volatility.' IZA Discussion Papers 3263, Institute for the Study of Labor (IZA), Dec

Andersen, L. et Faris. (2002a) 'Natural gas and income distribution in Bolivia.' Andean Competitiveness Project. Working paper

Andersen, L. et R. Faris. (2002b) 'Reducing volatility due to natural gas exports: Is the answer a stabilization fund?' Andean Competitiveness Project. Working paper

Annabi, N., F. Cissé J. Cockburn, and B. Decaluwé (2008) 'Libéralisation commerciale, croissance et pauvreté au Sénégal: Une analyse à l'aide d'un megc microsimulé dynamique.' *Economie et Prévision* (186), 117–131

Atkinson, A. B. (1992) 'Measuring poverty and differences in family composition.' *Economica* 59(233), 1–16

Boccanfuso, D., and L. Savard (2007) 'Poverty and inequality impact analysis regarding cotton subsidies: A Mali-based cge micro-accounting approach.' *Journal of African Economies* 16(4), 629–659

Bourguignon, F., A. S. Robilliard, and S. Robinson (2005) 'Representative versus real households in the macroeconomic modeling of inequality.' In *Frontiers in applied general equilibrium modeling: In honor of Herbert Scarf,* ed. T. N. Srinivasan Timothy J. Kehoe and John Whalley Cambridge University Press

Bourguignon, F., and L. Savard (2008) 'Distributional effects of trade reform: an integrated macro-micro model applied to the Philippines.' In *The impact of macroeconomic policies on poverty and income distribution,* ed. M. Bussolo F. Bourguignon and L. A. Pereira da Silva The World Bank

Chen, S., and M. Ravallion (2004) 'Welfare impacts of China's accession to the World Trade Organization.' *The World Bank Economic Review* 18, 29–57

Cloutier, M.H., and J. Cockburn (2002) 'How to build an integrated cge microsimulation model. Step-by-step instructions with an illustrative exercise.' Technical Report, CIPREE, Université de Laval

Cockburn, J. (2002) 'Trade liberalisation and poverty in nepal: A computable general equilibrium micro simulation analysis.' Technical Report

Cororaton, C. B., and J. Cockburn (2007) 'Trade reform and poverty - Lessons from the Philippines: A cge-microsimulation analysis.' *Journal of Policy Modeling* 29, 141–163

Cowell, F. A. (1984) 'The structure of american income inequality.' *Review of Income and Wealth* 30(3), 351–75

Danziger, S., and M. K. Taussig (1979) 'The income unit and the anatomy of income distribution.' *Review of Income and Wealth* 25(4), 365–75

de Janvry, A., E. Sadoulet, and A. Fargeix (1991) *Adjustment and equity in Ecuador* (OECD Development Center)

de Melo, J., and D. Tarr (1992) *A General Equilibrium Analysis of United States Foreign Trade Policy* (MIT Press, Cambridge Mass.)

de Melo, J., and S. Robinson (1989) 'Product differentiation and the treatment of foreign trade in computable general equilibrium models of small economies.' *Journal of International Economics* 27, 47–67

Deaton, A. (1997) *The analysis of household surveys. A microeconometric approach to development policy* (World Bank by The Johns Hopkins University Press)

Deaton, A., and J. Muellbauer (1980) *Economics and consumer behavior* (Cambridge University Press)

Decaluwé, B., A. Patry L. Savard, and E. Thorbecke (1999) 'Poverty analysis within a general equilibrium framework.' CRÉFA Working Paper 9909

Decaluwé, B., J.-C. Dumont, and L. Savard (1999) 'Measuring poverty and inequality in a computable general equilibrium model.' CRÉFA Working Paper 9920

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

Dimaranan, B. V., R. A. McDougall, and T. W. Hertel (2002) 'Behavoral parameters.' In *Global Trade, Assistance, and Production: The GTAP 5 Data Base,* ed. B. V. Dimaranan and R.A. McDougall Center for Global Trade Analysis

Cuclos, J.-Y., and A. Araar (2006) *Poverty and equity: measurement, policy, and estimation with DAD* (Springer/IDRC)

Ebert, U. (1997) 'Social welfare when needs differ: An axiomatic approach.' *Economica* 64, 233–244

—— (1999) 'Using equivalent income of equivalent adults to rank income distributions.' *Social Choice and Welfare* 16, 233–258

F. Bourguignon, M. Bussolo, and L. A. Pereira da Silva (2008) 'Introduction: evaluating the impact of macroeconomic policies on poverty and income distribution.' In *The impact of macroeconomic policies on poverty and income distribution,* ed. M. Bussolo F. Bourguignon and L. A. Pereira da Silva The World Bank

Fofana, I., and J. Cockburn (2003) 'Micro-simulation in computable general equilibrium: procedure for analyzing and reconciling data.' Poverty and Economic Policy Network and Centre Interuniversitaire sur le Risque, les Politique Economiques et l'Emploi, Université de Laval

Gylfason, T. (2001) 'Natural resources, education, and economic development.' *European Economic Review* 45(4-6), 847–859

Hatfield, M. (2002) 'Constructing the revised market basket measure.' Technical Report, Human Resources Development Canada, Apr

Herault, N. (2007) 'Trade liberalisation, poverty and inequality in South Africa: A computable general equilibrium - microsimulation analysis.' *The Economic Record* 83(262), 317–328

—— (2009) 'Sequential linking of computable general equilibrium and microsimulation models.' Melbourne Institute Working Paper No. 2/09

Hertel, T., M. Tsigas, and B. Narayanan (2008) 'Primary factor shares.' In *Global Trade, Assistance, and Production: The GTAP 7 Data Base,* ed. B. Narayanan and T.L. Walmsley Center for Global Trade Analysis

Hertel, T. W., and J. J. Reimer (2005) 'Predicting the poverty impacts of trade reform.' *Journal of International Trade and Economic Development* 14(4), 377–405

Hertel, T. W., R. Keeney M. Ivanic, and L. A. Winters (2007) 'Distributional effects of wto agricultural reforms in rich and poor countries.' *Economic Policy* 22, 289–337

Hicks, J.R. (1956) *A revision of demand theory* (Clarendon Press, Oxford)

Jomini, P., J.F. Zeitsch R. McDougall A. Welsh S. Brown J. Hambley, and J. Kelli (1991) 'Salter: a general equilibrium model of the world economy, vol. 1. model structure, data base and parameters.' Technical Report

Kirman, A. P. (1992) 'Whom or what does the representative individual represent?' *Journal of Economic Perspectives* 6(2), 117–36

Landa, F. (2003) 'Pobreza y distribución del ingreso en bolivia entre 1999 y 2002.' Technical Report, UDAPE, Bolivia

Lanjouw, P., and M. Ravallion (1995) 'Poverty and household size.' *Economic Journal* 105(433), 1415–34

Lofgren, H., R. L. Harris, and S. Robinson (2002) *A standard computable genera equilibrium (CGE) model in GAMS* (International Food Policy Research Institute)

Magnani, Riccardo, and Jean Mercenier (2009) 'On linking microsimulation and computable general equilibrium models using exact aggregation of heterogeneous discrete-choice making agents.' *Economic Modelling* 26(3), 560–570

Ravallion, M. (1992) 'Poverty comparisons. a guide to concepts and methods.' Technical Report, Living Standards Measurement Study. Working paper No. 88

Reimer, J., and T. Hertel (2004) 'International cross section estimates of demand for use in the gtap model.' GTAP Technical Papers 23, Center for Global Trade Analysis, Department of Agricultural Economics, Purdue University

Robilliard, A. S., F. Bourguignon, and S. Robinson (2008) 'Examining the social impact of the Indonesian financial crisis using a micro-macro model.' In *The impact of macroeconomic policies on poverty and income distribution,* ed. M. Bussolo F. Bourguignon and L. A. Pereira da Silva The World Bank

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

Rutherford, T., D. Tarr, and O. Shepotylo (2005) 'Poverty effects of russia's wto accession : modeling"real"households and endogenous productivity effects.' Policy Research Working Paper Series 3473, The World Bank, January

Rutherford, T. F., and D. G. Tarr (2008) 'Poverty effects of Russia's WTO accession: modeling "real" households with endogenous productivity effects.' *Journal of International Economics* 75, 131–150

Sadoulet, E., and A. de Janvry (1995) *Quantitative development policy analysis* (The Johns Hopkins University Press)

Savard, L. (2003) 'Poverty and income distribution in a cge-household microsimulation model: Top-down/bottom up approach.' Technical Report

Zavaleta, D. (2010) 'Modeling the impact of economic policies on poverty and inequality: the case of Bolivia', PhD disseration, University of Geneva.