CIRPÉE
Centre interuniversitaire sur le risque, les politiques économiques et l'emploi

Cahier de recherche/Working Paper **03-32**

Poverty, Income Distribution and CGE Modeling: Does the Functional Form of Distribution Matter?

Dorothée Boccanfuso
Bernard Decaluwé
Luc Savard

Août/August 2003

Boccanfuso: CIRPÉE, Université Laval, Dakar, Senegal
dboc@ecn.ulaval.ca
Decaluwé: CIRPÉE, Université Laval, Québec, Canada
bdec@ecn.ulaval.ca
Savard: International Development Research Centre, (IDRC-Dakar, Senegal)
lsavard@idrc.org.sn

Abstract: In this paper, we provide an overview of approaches used to model income distribution and poverty in CGE models. CGE models have started to use income distribution functional forms such as the lognormal, Pareto, beta distribution and Kernel non-parametric methods to apply FGT poverty indices. None of the authors of these papers have gone into much detail to justify the use of one method or functional form over the other, within the context of this type of work. Extensive literature exists on the choice of functional forms to estimate income distribution; however it has not been utilized in the CGE context. Given the fact that the desegregation of groups of households can be important in CGE analysis and the fact that the impact on income of policy simulations are often small in CGE models, we investigate the importance of the choice of the functional form used to estimate the income distribution of groups of households. We compare six functional forms with parametric estimation and on a non-parametric method. Results show that no single form is more appropriate in all cases or groups of households. The characteristics of samples and subgroups play an important role and the choice should be guided by the best fitting distribution.

Keywords: Computable general equilibrium models, Estimation, Personal Income and Wealth Distribution, Measurement and Analysis of Poverty

JEL Classification: I32, D31, C13, C68

1. Introduction

Computable General Equilibrium (CGE) models have traditionally been used to simulate the impact of exogenous shocks (such as changes in international terms of trade, and a recession in importing countries) and changes in policies on the socio-economic system and, in particular, the income distribution. Good examples of such models are those that were built in connection with the OECD research program to explore the impact of structural adjustment on equity (see e.g. Thorbecke, 1991, for Indonesia; de Janvry et al., 1991, for Ecuador; Morrisson, 1991, for Morocco). Still an additional model developed in the context of Africa is that of Chia et al. 1994. These models allowed the impact of counterfactual policy scenarios to be simulated on income distribution. Since CGE models are fully calibrated on the basis of an initial year SAM that provides a set of consistent initial conditions—and the SAM, as such, does not contain information on intra socioeconomic household group income distribution, it follows that conventional CGEs can only simulate the impact of a shock on the representative household in each group. This amounts to the implicit assumption that the variance of income within a group is zero. To the extent that poverty is pervasive and is likely to affect many socioeconomic groups, (albeit to different degrees) it appears essential in any analysis of the impact of a shock on poverty to start with information on intra-group income distribution. Increasingly as more income and expenditure surveys become available, it is possible to generate the within-group income distributions prevailing in the same base year as that of the SAM used to calibrate the general equilibrium model.

During the 1980s and at the beginning of the 1990s, several authors used CGE model to study the impact of economic reforms on the distribution of income. The pioneers in this area were certainly Adelman and Robinson (1979) in Korea, as well as Dervis, de Melo and Robinson (1982) and Gunning (1983) in Kenya. More recently in a series of paper Decaluwé, Patry, Savard, Thorbeke (1998), Decaluwé, Dumond and Savard (1999), Decaluwé, Savard and Thorbecke (2001), Cockburn (2001), Agenor, Izquierdo and Fofack (2001), Cogneau and Robilliard (2000), Bourguignon, Robillard, and Robinson (2002), Boccanfuso, Cissé, Diagne and Savard (2003), Savard (2003) follow these authors by assessing poverty through a computable general equilibrium model.

Initial work in this direction used the mean income changes in the representative households of the sub-categories as an input into changes of the distribution of income of a sub-group of a population. The next step was to apply poverty indicators such as the FTG measure proposed by Foster, Greer and Thorbecke (1984). Among these were de Janvry *et al*. (1991), Chia *et al*. (1994), Decaluwé *et al*. (1998) and Decaluwé *et al*. (1999). This approach is particularly interesting, notably in the way it links policy simulation and external shocks to poverty analysis. These authors use different functional forms to model the income distribution of the groups of households that serve as a basis for calculating poverty indices. de Janvry *et al*. (1991) use the Pareto distribution to characterize the income distribution of different sub-groups of the population of Ecuador, Chia *et al*. (1994) use the lognormal distribution for groups in Ivory Coast and finally, Decaluwé *et al*. (1998) and Decaluwé et al. (1999) use the beta distribution for their African archetype economy.

In the case of Dervis, de Melo and Robinson (1982), they chose the lognormal distribution because it has interesting characteristics, notably, the two parameters (mean and variance) are linked with a theoretical relationship which allows the authors to use the change in income of a representative group as the new mean of the said group. It then allows them to calculate a new theoretical variance of their function and then plot the new distribution with the CGE calculated mean and theoretical mean. In Adelman and Robinson (1979), a statistical test is also performed on the lognormal, and in some cases the test (skewness and kurtosis) were not satisfactory. They simply eliminated a socio-economic group (by aggregation) to circumvent the problem. The income distribution modelling approach and the statistical literature provide evidence that other functional forms might be more appropriate to represent income distribution (see Bordley, McDonald and Mantrala, 1996). In de Janvry *et al*. (1991), the properties of the Pareto distribution are discussed, and appear to be the most accurate to represent the distribution of groups with higher incomes, whereas the lognormal distribution was more appropriate for groups having a higher concentration of low incomes. Decaluwé *et al*. (1998) argue that when the work requires desegregation with different distributions, a more flexible form should be used such as the beta distribution which allows distributions to better approximate different types of "real income distributions". Chia *et al*. (1994) as Adelman and Robinson (1979), Dervis, de Melo and Robinson (1982) chose the lognormal distribution and Cockburn (2001) used a non-parametric method to estimate the income distribution of different groups of households in Nepal.

As we will discuss later, literature on income distribution has proliferated over the last half century and has not been fully exploited in the context of CGE analysis. This paper analyses a variety of functional forms used to approximate income distribution identified in our literature review and uses them in the framework of CGE analysis. This investigation is, in our view, important, as many researchers using CGE analysis to link policy impact and external shocks are interested in observing disaggregated sub-groups of the population. This division of households into sub-groups can have consequences on the properties of the distribution and therefore on the appropriateness of one functional form or method over another, as we will see later. We think that the higher the degree of desegregation of the household, the more the choice of the functional form of estimation method for income distribution is important in order to have a precise estimation of changes in poverty levels following a policy simulation of external shock.

The second element that makes this analysis important in the context of CGE analysis is that CGE modeling has specific properties that allow the analyst to generate a large vector of prices and factor payment that will be at the origin of income or total expenditure changes for each household of the survey taken individually. The changes are household specifics as the household each have a specific income and expenditure structure. Given these two characteristics the approach proposed is specifically valuable in the CGE context. However, this does not preclude applying the same approach in another context that uses functional forms to approximate income distribution for poverty analysis especially if this approach assumes that functional forms are invariant pre and post simulation. A brief description of the CGE model is provided before the properties of the functional forms. We follow with a presentation of simulation results and the presentation of poverty analysis results to finish with the concluding remarks.

2. Poverty measurement with endogenous poverty line in the CGE context

The procedure used to analyse poverty in this paper is the same as in most papers referenced in the CGE and poverty review. The first step is to define the group's household for the benchmark and after simulation by household's characteristics such as the regional zone or the level of education. The next step consists of estimating the parameters of the income distribution function of each group for the both income vectors. This procedure allows us to compare the poverty levels obtained in the post-simulation case with those prevailing in the pre-simulation case using Foster, Greer and Thorbecke'(F-G-T) P_a measures. The FGT P_a class of additively decomposable poverty measures allows us to measure the proportion of poor in the population (the headcount ratio) but also the depth and severity

of poverty. The P_a measure expressed in terms of the statistical distribution becomes:

$$P_{\alpha} = \int_{mn}^{z} \left(\frac{z-y}{z}\right)^{\alpha} I\left(y;\hat{\Theta}\right) \mathrm{dy} \qquad \textbf{(2-1)}$$

where a is a poverty-aversion parameter, z is the poverty line and mn the minimum (intra-group) income and $\hat{\Theta}$, the estimated parameter's vector of a statistical distribution as defined in the following section.

When a = 0, the headcount ratio is derived from the equation (2-1). In this case, the P_a yields the proportion of the population within a group below the poverty line. With a = 1, the relative importance accorded to all individuals below the poverty line is proportional to their incomes which is the income poverty gap. As a increases, more importance is given to the shortfalls of the poorest households and the measure becomes more distributionally-sensitive; society becomes more averse to poverty. In the case of $\alpha = 2$, this index assumes that each poor household is assigned a weight equal to its shortfall from the poverty line. For further discussion on this measure see Ravallion, (1994).

The poverty line itself (z in equation 2-1) is determined endogenously within the CGE model as in Decaluwé *et al.* (1999). We postulate that the poverty line is determined by a basket of quantities of commodities reflecting basic needs (*BN*) consistent with Ravallion's (1994) approach to estimating absolute poverty. We denote this basket as ϖ_i^p. This basket remains invariant from one simulation to another and is the same for the population and then for all groups of households. In turn, the monetary poverty line is obtained by multiplying the *BN* commodity basket by their respective prices (Pq_i) and aggregating across commodities:

$$z = \sum_i \varpi_i^p Pq_i \qquad \textbf{(2-2)}$$

Since commodity prices are endogenously determined within the model, so is the nominal value of this basket, i.e. the poverty line. If commodity prices rise following an external shock, the poverty line will increase (shift to the right) and poverty will rise *ceteris paribus*.

3. Income distribution and poverty analysis and goodness-of-fit test

3.1. A brief review

> "*The forces determining the distribution of incomes in any community are so varied and complex, and interact and fluctuate so continuously, that any theoretical model must either be unrealistically simplified or hopelessly complicated.*"
>
> *(Champernowne, 1953)*

Research interest on income distribution began at the end of the nineteenth century and one of the objectives was to provide a mathematical description of the size of income distribution to approximate the 'true' distribution. Initially, it was believed that incomes were distributed normally but Pareto (1897) empirically proved that incomes were lognormally distributed and that the skewness to the right had a flat tail, meaning unequal distribution.

Since Pareto, various functional forms have been proposed and can be grouped into three main categories. The first category is composed of forms describing an income distribution generated by a stochastic process (Champernowne, 1953; Rutherford, 1955). The main criticism addressed to this group is that the adepts only take into account the theoretical properties of the income variable, omitting the empirical aspect. The second category concerns the functional forms that provide a good fit to empirical data but which have no theoretical basis (Salem and Mount, 1974; McDonald, 1984). Finally, in the last group we find functional forms as a solution of specified differential equations such that the theoretical foundation is developed on the basis of empirical evidence (Pareto, 1897; Singh and Maddala, 1976).

The literature offers many alternatives of the probability density function to approximate the 'true' distribution of income distribution. Generally, the parameters should be simple to estimate and to interpret. This is true for the lognormal and Pareto distributions where these two desirable properties are respected. But other distributions are recognized to improve the fit, although they are more difficult to interpret, particularly the displaced lognormal or the beta distribution.

The two functions most often used are the Pareto and the lognormal, however it appears that empirically, the first one is appropriate only for the upper tail of the income distribution and the fit over the whole range of income is poor. Nonetheless, this last result seems to be the rule for all the two-parameter income distributions. The lognormal income distribution suggested by Gibrat (1931) and further examined by Aitchinson and Brown (1957), seems to fit well at the lower income levels in the literature, but its fit towards the upper end is not satisfactory. Salem and Mount (1974) reintroduced after Ammon (1895) the gamma distribution. This distribution generally provides a better fit than lognormal only at the two tails. Salem and Mount (1974) found that empirically gamma distribution fits better than lognormal.

Champernowne suggested a three-parameter distribution, which fit better than the two-parameter ones. The limiting form of this Champernowne distribution, two-parameter distribution appears to be an empirical substitute for the normal distribution and asymptotically, this distribution approaches a form of Pareto for the large values of income (Fisk, 1961).

Furthermore, McDonald (1984), McDonald and Xu (1995), Bordley, McDonald and Mantrala (1996) and Gordy (1998) think that even if the beta distribution is flexible and can take a variety of shapes, it is a two-parameter distribution and the precision in fitting data is limited. In the last years, these researchers among others, contributed to generalizing the beta function. This more complex model seems to be more appropriate to reflect the impact of economic fluctuations. With respect to the displaced lognormal, a generalization of the lognormal distribution, most of the literature encountered dealt with difficulties in estimating its parameters but less in terms of its fitting characteristics. The only comparison found was performed by Abdelkhalek and Chaoubi (2000) who state that the displaced lognormal, the lognormal and the Champernowne have similar fitting properties.

Better fits could be obtained with two distributions members of the Burr family: the Singh-Maddala (1976) and the Dagum (1977)[1]. The Singh-Maddala is a generalization of the Pareto and Weibull distributions and in terms of goodness-of-fit, this model outperforms both the lognormal and gamma distributions considering the US income data application done by Singh and Maddala (1976). The Dagum (1977) proposed a theoretical description based on the observed characteristic of regularity of income-elasticity in observed income distribution. There are three types of Dagum distributions (three and four parameters). The Dagum Type I three-parameter, has been chosen in this study since it is considered to be the best to represent the behaviour of employed wage earners. Many other distributions could have been added to this list in this context of income distribution but we chose to look at seven as they are quite representative of the usual choices made. Moreover, they are relatively tractable and allow us to achieve our objective without too much technical complexity.

3.2. The distribution income: estimation and poverty measures

We used two methods to evaluate poverty with the Foster, Greer and Thorbecke (FGT) indices for a base year and after simulation, with an endogenous poverty line and for each group. The first one is a non-parametric method and can easily be calculated with the use of DAD software[2]. The second is the parametric method. For the latter, we present seven continuous distributions and estimate their respective parameters. In this section, we present some technical elements, which are useful in understanding the properties of each functional forms used in the calculation of the FGT poverty indices.

3.2.1. The Kernel method (DAD)

The histogram is a graphical way to summarize the relative frequency of occurrences of the value for the income variable X. This function has the characteristic of having jumps at the points, even if the data represent realizations of continuous random variables. As Rosenblatt (1956) proposed, it is possible to transform this histogram to obtain a density function where the most decisive step is to smooth over the edges using a kernel weight function. The Kernel method is the most mathematically studied and commonly used non-parametric density estimation method. The Kernel estimation of $f(x)$ or the smoothed histogram is defined as:

[1] The Singh-Maddala and the Dagum distribution are known in statistics literature respectively as Burr 12 and Burr 3.

[2] Software developed by Duclos *et al.* (1999), http://www.pep-net.org

$$\hat{f}(x)=\frac{1}{N}\sum_{i=1}^{N}\frac{1}{h}K\left(\frac{x-x_i}{h}\right) \qquad (3\text{-}1)$$

where the Kernel function *K()* is generally unimodal, symmetric, bounded density function, for instance, the standard normal density function and the *h* is called the smoothing parameter. Imagine it intuitively, a "bump" is placed on each data point, and the sum of all "bumps" reflects the overall distribution of all data points. The Kernel function determines the shape of each bump while the smoothing parameter determines the width of each bump. This function has the following properties:

- no need to know the data range in advance,
- $\hat{f}(x)$ itself forms a density function which inherits all the continuity, differentiability and integrability properties of the Kernel function,
- *K* and *h* are two factors affecting the accuracy but essentially by the smoothing parameter.

The estimation consists of measuring and minimizing the global error between the density estimation and the real underlying density function such as:

$$Mean\ Integrated\ Squared\ Error\ (\hat{f},f)=E\int(\hat{f}(x)-f(x))^2dx \ \rightarrow\ 0 \qquad (3\text{-}2)$$

where $$\hat{f}(x)=\frac{1}{N\sqrt{2\pi}}\sum_{i=1}^{N}\exp\left\{-\frac{1}{2}\left(\frac{x-x_i}{h}\right)^2\right\} \qquad (3\text{-}3)$$

and $$h=\sigma.(1.06).N^{-\frac{1}{5}} \qquad (3\text{-}4)$$

if using the Normal kernel. The DAD software uses the non-parametric method of Gaussian Kernel type.

3.2.2. The parametric method

Even if the smoothing reduces the data specificity of the histogram, many economists have approached the issues of the distribution of income with statistical parametric distribution. We summarize the distribution used, the data and parameters constraints in Table 3-1. Some characteristics of these distributions follow.

		Pdf	Data and parameters constraints
Two-parameter distributions	**Lognormal**	$f(x;\mu,\sigma)=\frac{1}{(x-\min)\sqrt{2\pi\sigma^2}}\exp\left[-\frac{(\ln(x-\min)-\mu)^2}{2\sigma^2}\right]$	x > 0
	Gamma	$f(x;\alpha,\beta)=\frac{(x-\min)^{\alpha-1}}{\beta^{\alpha}\Gamma(\alpha)}\exp\left(-\frac{x-\min}{\beta}\right)$	x > 0 ; α > 0 ; β > 0 ; Γ(α) = gamma function
	Beta	$f(x;p,q)=\frac{1}{B(p,q)}\frac{(x-\min)^{p-1}(\max-x)^{q-1}}{(\max-\min)^{p+q-1}}$	xmin < x < xmax; p > 0 ; q > 0 ; $B(p,q)=\int_{xmin}^{xmax}\frac{(x-xmin)^{p-1}.(xmax-x)^{q-1}}{(xmax-xmin)^{p+q-1}}dx$ [3]
		or $f(y;p,q)=\frac{1}{B(p,q)}y^{p-1}(1-y)^{q-1}I_{(0,1)}(y)$	0 < y < 1; p > 0 ; q > 0 $B(p,q)=\int_0^1 x^{p-1}.(1-x)^{q-1}dx$
	Champernowne	$f(x;\mu,\theta)=\theta\frac{\mu^{\theta}x^{\theta-1}}{(\mu^{\theta}+x^{\theta})^2}$	x > 0 ; μ > 0 ; θ > 0 ;
Three-parameter distributions	**Displaced lognormal**	$f(x;\lambda,\mu,\sigma)=\frac{1}{(x-\lambda)\sqrt{2\pi\sigma^2}}\exp\left[-\frac{(\ln(x-\lambda)-\mu)^2}{2\sigma^2}\right]$	x > λ
	Singh-Maddala	$f(x;a,b,q)=\frac{qax^{a-1}}{b^a\left(1+\left(\frac{x}{b}\right)^a\right)^{1+q}}$	x > 0 ; a > 0 ; b > 0 ; q > 1/a
	Dagum	$f(x;\beta,\lambda,\delta)=\beta\lambda\delta x^{-\delta-1}(1+\lambda x^{-\delta})^{-\beta-1}$	x > 0 ; β > 0 ; λ > 0 ; δ > 0

Table 3-1: *The continuous density functions and constraints on data and parameters*

The lognormal distribution exhibits three main features: asymmetry, a left humpback and a long right-hand tail. Since a lot of observations appear to be on the left side of the distribution and the highest concentration of observations seems to also be on this side, economists interested in poverty analysis often use this distribution.

The Gamma distribution is bounded at the lower side and the shape of this function will depend of parameter α. For an increasing value of this shape parameter, the peak of the distribution moves away from the minimum value of income.

Beta distribution describes a family of curves that are unique in that they are nonzero only on the interval [0; 1]. A more general version of the function assigns parameters to the end-points of the interval as indicated in Table 3-1. This distribution can approach zero or infinity at either of its bounds with *p* controlling the lower bound and *q* controlling the upper bound. If *p* and *q* are less than 1, the beta distribution approaches infinity. This case will be problematic to determine the poverty FGT measure characterized by the surface under the distribution.

[3] This function is called complete Beta function.

In his 1953 article, Champernowne proposes a model in which individual incomes were assumed to follow a random walk in the logarithmic scale and yielded a Pareto distribution over the whole range. There exist three, four and five parameter distributions, which improve the fit by incorporating extra parameters, therefore allowing for more flexibility, as we saw previously. However, in our paper, we decided to use the limiting form of the density function of the Champernowne distribution, the two-parameter distribution. The parameter μ is the income median value and θ a constant corresponding to Pareto's constant for high income.

The displaced lognormal distribution is more general than the lognormal. The location or threshold parameter, λ, is crucial to modeling the skewness of some distributions. There exist various methods of estimation but the most used is the maximum likelihood method. This consists in solving a system of three non-linear equations (Aitchinson and Brown, 1957; Cohen, 1951; Johnson and Kotz, 1970 and Abdelkhaleck and Chaoubi, 2000).

As the displaced lognormal density function, the Singh – Maddala and Dagum pdf's are a three-parameter distribution encompassing a wide range of distributional shapes. The parameters β and δ of the Dagum distribution represent the shape or *equality* parameters for the lower and upper-middle tail. There will be an improvement of the income distribution in terms of equality when at least one shape parameter will be increased. Finally, λ is a scale parameter. The Singh-Maddala distribution has two main advantages: first, it accommodates sufficient flexibility to model heterogeneous income data and secondly the estimation is easy. The parameters a and q determine the shape of the distribution whereas b is a scale parameter. These distributions are known to provide a good fit of income data in many situations (McDonald, 1984).

The estimators for each continuous distribution were obtained by maximizing the maximum likelihood function, based on individual observations and are asymptotically efficient. The same estimation method was applied for each distribution since McDonald and *al.* (1979) found that the estimates depend on the functional forms of the distribution and estimation technique selected.

3.3. The goodness-of-fit tests or do the observations come from a particular distribution?

Two approaches permit to compare the input data to the fitted distributions in a statistically manner. The first one is graphical and the second, numerical.

3.3.1. Graphical approach

This method allows us to visualize the comparison between the empirical distributions graphed with the observed data and the density function or cumulative function estimated from the observed data. The density function that is closest to the empirical distribution in the limits of the poverty line would be the form that best approximate the "real distribution" around the poverty line. This method must be used with care given its nature. To avoid "false" conclusions, we used it mainly to confirm numerical tests.

3.3.2. Statistical tests of goodness-of-fit

The goodness-of-fit test indicates whether it is reasonable to assume that a random sample comes from a specific distribution, based on consistence with observed data. This is generally used in the case of income distribution analysis[4]. Goodness-of-fit tests are a form of hypothesis testing where the null and alternative hypotheses are:

- H_0: sample data come from the stated distribution.
- H_1: sample data do not come from the stated distribution.

Then if the probability of observing the data i.e. the *p-values* is too low, the model is rejected. Three tests are generally used: Chi-square type or based on errors tests for continuous and discrete distributions, tests of Kolmogorov-Smirnov based on the empirical distribution function (EDF) and Anderson-Darling test for continuous distributions. This last test is valuable since it is sensitive to discrepancies at the tails of the distribution. However, it is not easy to find tables for critical values in cases where complex distributions are used.

To evaluate the quality of adjustment of distributions to observations, we selected the first type of test based on errors where the continuous data were separated into intervals, this tests starting with the observed data in intervals. The measures of goodness-of-fit utilized are: the sum of squared errors (SSE), the sum of absolute errors (SAE), the chi-squared goodness-of-fit (χ^2) and log-likelihood values.

The SSE, SAE and χ^2 statistics are calculated according to the following equations:

$$SSE = \sum_{i=1}^{k} \left({n_i}/{N} - p_i\left(\hat{\Theta}\right) \right)^2 \qquad \textbf{(3-5)}$$

$$SAE = \sum_{i=1}^{k} \left| {n_i}/{N} - p_i\left(\hat{\Theta}\right) \right| \qquad \textbf{(3-6)}$$

$$\chi^2 = N \sum_{i=1}^{k} \frac{\left({n_i}/{N} - p_i\left(\hat{\Theta}\right) \right)^2}{p_i\left(\hat{\Theta}\right)} \qquad \textbf{(3-7)}$$

[4] A detailed demonstration of goodness-of-fit techniques is presented in D'Agostino and Stephens (1986).

where n_i/N is the observed relative frequencies ($N=\sum_{i=1}^{k} ni$) corresponding to $p_i(\hat{\Theta})=\int_{I_i} f(x;\hat{\Theta})dx$, the predicted fraction of the population in the i^{th} of k income groups defined by $I_i=[x_{i-1},x_i)$. The χ^2 statistic has an asymptotic distribution, which is Chi-square with degrees of freedom equal to the difference between the number of income groups (k) and the number of parameters even though the criterion for both based errors statistics will be the minimization. Finally, the distribution that will maximize the log-likelihood is the one that will be chosen.

4. The SAM: The Senegalese multi-household CGE model and SAM.

4.1. The data

In the paper we used a CGE model to generate ex ante policies simulation to induce changes in the income distribution and poverty indices. The approach adopted is the integrated multi-household CGE model such as proposed by Decaluwé, Dumont and Savard (1999) and applied by Cockburn (2001). The CGE model used in this paper is the same as the one used in Boccanfuso, Cissé, Diagne and Savard (2003). This approach allows for particular mode of distribution changes as each representative household from the household survey is used in the model and has specific characteristics, which will contribute to specific changes in its income used later in the poverty analysis. As each household is represented by a specific income and expenditure structure, changes in factor payment will lead to differential income changes for each household[5]. The main features of this relatively standard model are the presence of a perfectly segmented market, small open economy with Armington (1969) assumption for import demand behaviour. Capital is supposed fixed, which generates price of capital payment specific to each production branches. This provides us with 11 prices for factor's payment, which are the main sources of heterogeneous impact on household income.

The SAM is decomposed in 10 production sectors, where 7 are tradable and 3 are non tradable. Specific accounts are used to distinguish the destination of the goods, namely for the domestic and export market. For the factor account we distinguish qualified and unqualified labour, and we also have specific accounts for each production branches' capital income, as we will perform capital income mapping between branches and households. The entire households of the 3278 household of the ESAM I (1994/1995) are integrated in the SAM. Other agents include government, rest of the world and firms. We also include saving and investment accounts. In total the SAM has 3336 accounts of which 58 are non-household accounts.

[5] See Table 7-1for descriptive statistics of the sample.

4.2. CGE simulations and interpretation of results

We simulated two scenarios to illustrate our point and generate ex ante income distribution for poverty analysis. The first one being a 20% increase in the food industry (non food oil) capital, the second is a decrease of 70% in import duties for processed food product. We present very few results from the CGE model in Table 4-1 since they are not the main focus of this paper. We focus on factor payments and aggregate household income as well as government related variables. It is still interesting to look at the main transmission mechanisms to the households' income.

- **Simulation 1**: 20% increase in capital of food industries (excluding food oil industries).

We note in this scenario a direct impact on both qualified and unqualified wage with both of them decreasing slightly. As for the capital income, as expected, we observe a strong decrease in the capital payment of the "other food industries" as capital becomes more abundant then labour in the sector. This will have an important impact on factor payments. Moreover, capital income decreases abruptly in "food oil", "other industries" and "extractive industries". The commercial sector capital payment is the one that increases the most with this external shock. The main effect on household's incomes drops down via these factor payments. We note that the aggregate effect on household income is an increase of 0.7%. We can already see from this simulation that there are no major biases introduced with this policy between the households endowed with either qualified or unqualified labour. Indeed, the negative effect on the unqualified endowed household is not as strong –0,48% compared to 0,70%. The government income is very slightly affected with an increase of 0,38% and an improvement of the government deficit of 2,81%.

- **Simulation 2**: Decrease of 70% import duties of "other food industries".

This policy represents a decrease in protection of the most protected sector of the economy 35,5% effective tariff rate. As expected, this policy will have the strongest negative impact on output in the sector concerned by the policy i.e. it decreases by 0,36% albeit this decrease is relatively small. The same holds for capital payment, as the target sector is the most negatively affected with a decrease of 2,07%. The other sectors are only slightly affected by this policy in terms of capital payment, with the exception of the "food oil industries" which sees their capital income increase by 5,09%. As for the variation in wages, the qualified wage increases slightly by 0,16% and the unqualified decreases by the same percentage rate. The decrease in tariff rates induces a strong decrease in government income 2,56% and a stronger increase in budget deficit (19,13%).

Variables	Branches	Base	**Sim 1:** +20% K Other food ind.	**Sim 2:** 70% decrease in TR other food industries
ytm		177,62	0,70	-0,03
s		1,00	-0,70	0,16
sn		0,50	-0,48	-0,16
yg		59,41	0,38	-2,56
sg		7,95	2,81	-19,13
e		1,00	-1,16	1,11
poverty	Threshold	1430,8		
Va	Agriculture	23,06	0,09	-0,10
Va	Cattle	17,50	0,03	0,00
Va	Fish Industry	4,98	1,71	-0,16
Va	Food oil	0,95	-4,74	3,10
Va	Other food Ind	14,29	12,17	-0,36
Va	Extractive	3,19	-0,15	0,11
Va	Other indus.	14,94	-0,50	0,13
Va	Telecoms.	18,54	0,92	-0,04
Va	Services	46,20	0,27	0,08
Va	Pub. Services	22,08	0,39	-0,18
r	Agriculture	1,00	-0,31	-0,35
r	Cattle	1,00	8,88	-0,79
r	Fish industry	1,00	2,71	-0,36
r	Food oil	1,00	-8,20	5,09
r	Other food Industry	1,00	-32,70	-2,07
r	Extractive Industry	1,00	-2,71	1,60
r	Other industry	1,00	-3,05	0,67
r	Commerce	1,00	10,58	-0,54
r	Services	1,00	1,16	0,63

Table 4-1: *Simulations results*

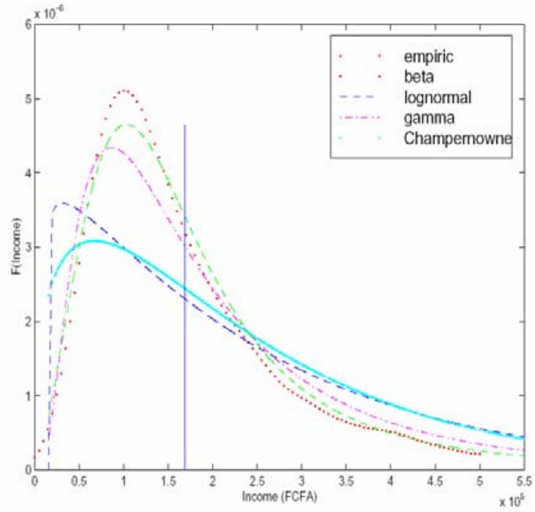
5. Results Interpretation

TheFigure 5-1 and 5-2 Figures present the graphs for income distribution function adjusted to the income vector for the whole sample of household included in the ESAM 94/95. The first graph compares the empiric distribution based on observed data with the estimated two-parameter density functions. In this case, the beta distribution seems to be the better fit compared to the empiric density as the Champernowne and the lognormal under-estimate the empirical distribution. The displaced lognormal three parameter function is the nearest of the empiric distribution.

Poverty Line (Fcfa) / adult equivalent / year		
Base	Sim 1	Sim 2
168 500,00	172 003,09	167 565,07

Table 5-1: *The poverty line*

At first glance, we see that if we draw a poverty line on these graphs and compute the area below the distribution function below (to the left of) the poverty line, we will get a different result for poverty indices, head count, depth and severity. Table 5-1 summarizes the three poverty lines obtained after simulation with the procedure described previously.



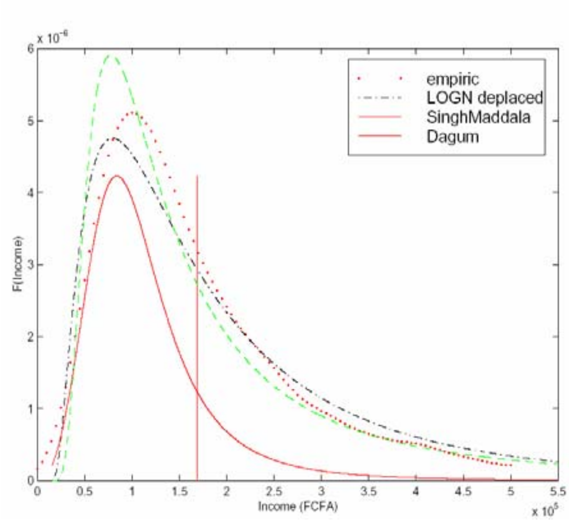


Figure 5-1 : *Senegal 2-parameter density functions*

Figure 5-2 : *Senegal 3-parameter density functions*

5.1. Poverty analysis for the benchmark case

With a poverty line evaluated at 168 500 Fcfa, poverty indices are computed and presented in Table 5-2. First, if we look at the aggregated data for Senegal, the Singh-Maddala and Gamma distributions provide different evaluations of the head count ratio obtained with the other six distributions. With a Singh-Maddala distribution, the head count ratio is 40,01% whereas the Gamma evaluates the number of poor at 43,55 % of the total population. The other six distributions provide values between 52,64%, for the lognormal distribution and 58,27% for the Dagum distribution. The non-parametric DAD approach estimates the poverty rate to 57,93%. The FGT1 and FGT2 do not exhibit the same characteristics just described for FGT0 since the Singh-Maddala distribution provides the lowest estimation of the severity of poverty at 17,64, and the Beta distribution the highest evaluation at 28,55. To measure the severity of poverty, the Beta distribution provides the highest level with 18,08 and the Singh-Maddala the lowest with 9,47.

If we look at the same computation of decomposable poverty indices for the different sub categories of households, the general picture is slightly different. The gamma distribution globally underestimated the level of poverty compared to other distributions except for the Dakar Educated group. For example, for the other-urban non-educated group (AUNE) the gamma distribution evaluates the number of poor at 55,83% and the Champernowne density provides the highest level of poor, with 62,78%. The same applies with the rural educated households, the head count ratio is evaluated at 46,65% with the gamma distribution, compared to 71,01% with the non-parametric DAD approach.

DAD	BETA	LOGN	GAMMA	CHAMP	LOGN3	SM	DAG
SENEGAL							

FGT0	57,93	56,74	52,64	43,55	55,53	54,30	40,01	58,27
FGT1	22,67	28,55	22,04	20,10	22,26	22,66	17,64	24,10
FGT2	11,43	18,08	12,01	12,10	11,84	12,11	9,47	12,42
	RNE							
FGT0	86,64	83,78	84,18	82,05	86,05	84,12	85,60	80,20
FGT1	39,20	37,64	38,50	36,74	39,01	38,65	38,30	39,63
FGT2	21,32	20,96	21,30	20,57	21,23	21,41	20,24	22,79
	RE							
FGT0	71,01	57,01	61,27	46,65	66,31	67,22	46,73	70,10
FGT1	33,56	29,32	28,07	21,16	29,71	32,64	24,64	33,40
FGT2	19,04	18,45	16,01	12,24	16,56	19,09	14,60	19,00
	DKRE							
FGT0	16,51	22,57	18,90	17,12	18,24	20,31	17,20	18,52
FGT1	4,41	9,03	5,87	6,22	5,61	5,75	4,57	4,66
FGT2	1,72	4,65	2,55	3,01	2,45	2,26	1,80	1,71
	DKRNE							
FGT0	37,17	34,93	34,36	29,85	34,68	38,43	28,52	40,63
FGT1	9,50	11,14	9,64	9,15	9,29	10,15	7,76	10,07
FGT2	3,34	4,74	3,77	3,85	3,56	3,64	2,87	3,40
	AUE							
FGT0	38,49	36,25	37,55	31,55	37,55	39,20	36,75	39,56
FGT1	10,12	14,61	13,46	12,62	12,86	13,95	12,07	12,93
FGT2	3,69	7,78	6,53	6,75	6,13	6,65	5,46	5,82
	AUNE							
FGT0	60,12	56,83	60,98	55,83	62,78	61,06	61,93	62,85
FGT1	19,18	22,23	22,48	21,23	21,93	22,53	20,10	24,97
FGT2	8,11	11,49	10,89	10,87	10,28	10,90	8,64	12,31

Table 5-2: *Poverty analysis for base benchmark*

Considering the depth of the poverty, the rural populations minimize it with the gamma distribution (36,74 for non-educated and 21,16 for educated). This is verified with the Singh-Maddala for the Dakar groups (7,76 for non-educated and 4,57 for educated) and DAD approach for the other non-Dakar urban populations (19,18 for non-educated and 10,12 for educated).

On the other hand, if we look at the severity of poverty, except for the rural population, the empiric DAD approach generally provides the lower evaluation compared to the other functional forms. We cannot identify a generalization for the higher severity index. This corroborates our intuition that no single distribution can fit every sub-group of the population whatever type of poverty indices used. This will be confirmed empirically in the following sub-section.

5.2. The goodness-of-fit tests results: the benchmark case

Table 5-3 shows the results of the four tests of goodness-of-fit for the benchmark case. The first statement is that no distribution function is a good fit for every group at the same time. Indeed, for both Dakar groups in the Senegalese case (educated and non-educated), the Singh-Maddala distribution is the best fit whereas for the non-educated rural and the non-educated other urban groups the Champernowne will be preferable. Moreover, it appears that some groups could be well fitted by two distributions. For the Dakar Educated group, the Singh Maddala is accepted by the all four criteria, but it could also be approximated by the Dagum distribution, which verifies three acceptance criteria. The Singh-Maddala and the Dagum distributions could also be chosen for the other educated urban case.

	BETA**	LOGN**	GAMMA**	CHAMP**	LOGN3*	SM*	DAG*	χ^2 Critical values
	SENEGAL							
SSE	0,01	0,00	0,01	*0,00*	*0,00*	*0,00*	*0,00*	* 30,14
SAE	0,45	0,20	0,42	0,13	0,16	*0,05*	0,14	** 31,41
Chi2	878,96	174,93	758,67	107,05	112,84	*14,33*	93,01	
LOGL	43 892,76	43 116,60	43 757,45	43 043,72	43 061,04	*46 584,59*	43 047,67	
	RNE							
SSE	*0,00*	*0,00*	0,01	*0,00*	*0,00*	*0,00*	0,01	* 22,36
SAE	0,16	0,14	0,22	*0,10*	0,14	0,11	0,27	** 23,68
Chi2	43,25	37,69	95,89	*20,46*	36,65	34,63	121,85	
LOGL	15 612,68	15 541,38	15 611,15	15 528,89	15 540,52	15 536,06	*15 627,67*	
	RE							
SSE	0,02	*0,01*	*0,01*	*0,01*	*0,01*	*0,01*	*0,01*	* 31,41
SAE	0,47	0,34	0,44	0,34	*0,33*	0,38	*0,33*	** 32,67
Chi2	51,06	66,26	123,49	59,19	50,53	47,30	43,20	
LOGL	1 772,803	1 787,358	1 816,882	1 784,039	1 778,606	*1 929,593*	1 776,267	
	DKRE							
SSE	0,03	*0,00*	0,01	*0,00*	*0,00*	*0,00*	*0,00*	* 27,59
SAE	0,23	0,25	0,46	0,21	0,21	*0,13*	*0,13*	** 28,87
Chi2	31,81	57,87	171,49	45,18	37,03	*26,98*	*14,16*	
LOGL	7 616,08	7 550,66	7 648,14	7 543,22	7 536,09	*7 736,07*	7 523,03	
	DKRNE							
SSE	0,01	*0,00*	0,01	*0,00*	*0,00*	*0,00*	*0,00*	* 30,14
SAE	0,31	0,23	0,37	0,17	0,16	*0,14*	0,16	** 31,41
Chi2	69,45	45,56	108,20	35,75	*21,49*	*22,96*	*21,90*	
LOGL	7 576,58	7 557,72	7 630,45	7 547,54	7 533,19	*8 318,49*	7 530,03	
	AUE							
SSE	0,01	*0,00*	0,03	*0,00*	*0,00*	*0,00*	*0,00*	* 3,84
SAE	0,34	0,10	0,27	0,03	0,08	0,03	*0,02*	**5,99
Chi2	53,10	*3,61*	25,67	*0,72*	*3,02*	*0,32*	*0,14*	
LOGL	4 368,76	4 384,49	4 433,34	4 379,48	4 382,13	*4 477,60*	4 375,26	
	AUNE							
SSE	0,01	*0,00*	0,01	*0,00*	*0,00*	*0,00*	0,02	* 5,99
SAE	0,16	0,13	0,19	*0,03*	0,13	0,10	0,26	** 7,81
Chi2	25,59	10,93	28,93	*0,81*	11,10	9,15	57,17	
LOGL	8 096,41	8 091,61	8 130,80	8 080,47	8 091,59	8 122,49	*8 143,62*	

Table 5-3: *Goodness-of-fit test, benchmark case*

Finally, the Rural Educated group distribution could be approximated by the three-parameter distributions that verify at least two acceptance criteria. This ambiguity could be caused by the small size of this group, composed of only 138 households (4,2% of the sample). These results tend to demonstrate that the three-parameter distribution is generally a better fit to the real distribution. For larger groups, like for instance the rural non-educated group (1265 observations, 38,6% of the sample) and the other non-educated urban household (635 observations, 19,4% of the sample), the two-parameter Champernowne distribution seems to be more appropriate. An important result from this analysis is that the most commonly used distribution to approximate income distribution such as the Gamma, Lognormal or Beta, are the ones that fit the worst in our application.

The general results presented seem to be confirmed with the graphical comparison (see in Figure 5-3). Indeed around the poverty line[6], the Singh-Maddala is very close of the empiric distribution followed by the Dagum and the Champernowne for the Senegalese case.

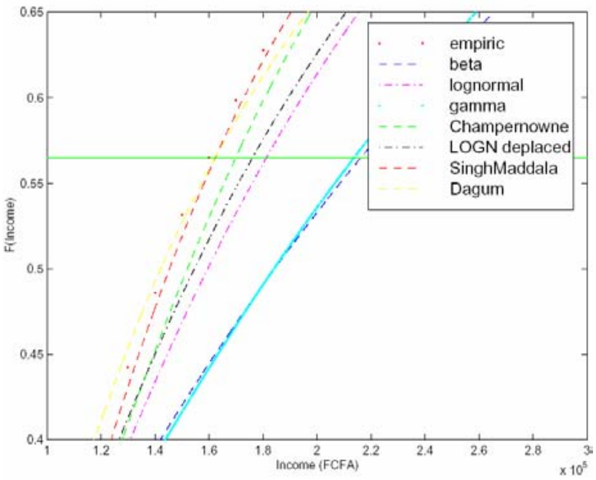


Figure 5-3 : *Senegal CDF (around the poverty line)*

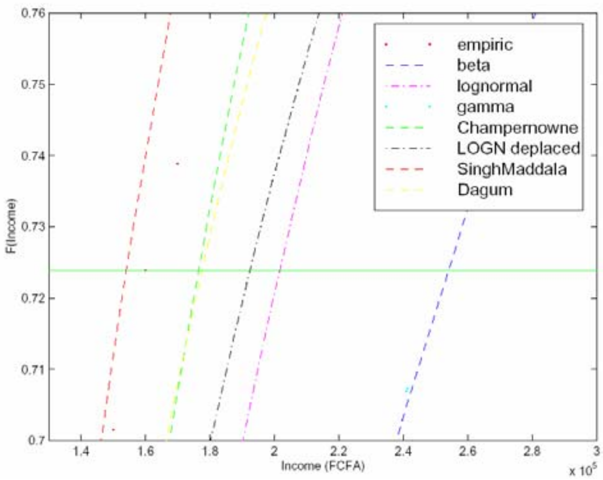


Figure 5-4 : *Rural educated CDF (around the poverty line)*

The graphical analysis facilitates the choice when other numerical tests do not provide a consensus. Consider the case of the educated rural household. In Table 5-3, each of the three-parameter distributions verifies two criteria. It is difficult to conclude what distribution best approximates the true distribution. The graphical analysis allows us (see figure 5-4) to see which distribution is closest to the empirical distribution around the poverty line. Indeed, the Singh-Maddala distribution is the nearest to the empiric one in the region of the poverty line. Therefore, in the case of the rural educated, we conclude that the Singh-Maddala is the best approximation of the empiric distribution.

5.3. The goodness-of-fit tests results: the simulation cases

[6] This zoom allows us to see which distribution is the nearest; possibly one of the best continuous distribution to represent the empirical distribution and therefore the most efficient one to compute poverty indices; which one is the best continuous distribution to approximate the empiric distribution and finally to calculate the best indices of poverty.

In most cases, the "best" distributions fitting the data do not change after simulation. However, in some cases, the parametrical distribution is modified and therefore fitting properties also change[7]. This is the case for the non-educated households of Dakar and rural educated shown in simulation 1.

	BETA**	LOGN**	GAMMA**	CHAMP**	LOGN3*	SM*	DAG*	χ^2 Critical values
				RE (benchmark)				
SSE	0,02	*0,01*	*0,01*	*0,01*	*0,01*	*0,01*	*0,01*	* 31,41
SAE	0,47	0,34	0,44	0,34	*0,33*	0,38	*0,33*	** 32,67
Chi2	51,06	66,26	123,49	59,19	50,53	47,30	43,20	
LOGL	1 772,803	1 787,358	1 816,882	1 784,039	1 778,606	*1 929,593*	1 776,267	
				RE (Sim1)				
SSE	0,02	0,01	0,01	0,01	0,01	0,01	*0,00*	
SAE	0,49	0,31	0,45	0,30	0,29	0,32	*0,27*	
Chi2	57,50	49,55	99,61	44,02	33,90	33,01	*27,41*	
LOGL	1 774,71	1 789,60	1 819,19	1 786,20	1 780,69	*1 916,89*	1 778,42	
				DKRNE (benchmark)				
SSE	0,01	*0,00*	0,01	*0,00*	*0,00*	*0,00*	*0,00*	* 30,14
SAE	0,31	0,23	0,37	0,17	0,16	*0,14*	0,16	** 31,41
Chi2	69,45	45,56	108,20	35,75	*21,49*	*22,96*	*21,90*	
LOGL	7 576,58	7 557,72	7 630,45	7 547,54	7 533,19	*8 318,49*	7 530,03	
				DKRNE (sim1)				
SSE	0,01	*0,00*	0,01	*0,00*	*0,00*	*0,00*	*0,00*	
SAE	0,32	0,25	0,39	0,21	0,17	0,16	*0,15*	
Chi2	69,19	54,70	118,08	44,83	31,28	32,29	31,61	
LOGL	7 579,64	7 560,65	7 632,95	7 550,41	7 537,14	*8 289,19*	7 534,26	

Table 5-4: *Goodness-of-fit: cases with a change in the distribution after simulation*

In Table 5-4, the distribution for the rural educated households becomes less ambiguous after simulation 1. Indeed, the parametric tests combined to the graphical analysis led to select the Singh-Maddala distribution. After simulation, the income distribution fitting best becomes the Dagum distribution. For the non-educated households of Dakar, where the Singh-Maddala distribution was the best approximation for the benchmark case, the choice is not as clear after the simulation. The Singh-Maddala and the Dagum distribution both fit relatively well. This result is interesting since it shows that the shape of the income distribution can change after simulation and highlights the drawback of postulating that one distribution represents all groups and is invariant across policy simulations or external shocks. The consequence of this postulate could contribute to misleading conclusions with respect to poverty analysis.

5.4. Poverty analysis for the simulation cases

[7] See the Table 7-2 and Table 7-3 in annex.

Table 5-5 puts in relation the results obtained from the DAD non-parametric approach and those obtained with the continuous distributions best fitting the observed data. Table 7-4 in annex summarizes the variation of the three FGT poverty indices for Senegal and the six sub-groups considered.

First, the empirical approach called DAD in the tables has some interesting characteristics. Indeed, for the educated rural households and non-educated Dakar group the variation of FGT0 is null. The discrete character of the empirical approach and the fact that the rural group is relatively small with only 138 households could explain this result. The variations of the poverty line and of the shape of the distribution (intra-group variation of income) are not sufficient to contribute to a change in the poverty level. However, this does not verify when the α parameter (of FGT) increases. In these cases the FGT variations increase with the increase in α and this is verified for all groups.

Another result of simulation 1 concerning the Dakar educated households is interesting to mention. The increase of 20% of capital in the food industry sector causes great dispersion of FGT indices between the chosen distributions. The empirical approach (DAD) overestimates the impact of this external shock whereas the gamma two-parameter distribution seems to underestimate the variation on incidence, depth and severity of poverty. This result for the empirical distribution can be explained by the fact that there is a concentration of households around the poverty line and therefore the variation of the poverty line (from the endogenous poverty line discussed earlier) and the income distribution leads to an important increase of the poverty indices.

Furthermore, Table 7-4 confirms the importance of the choice of distribution. Indeed, considering the Senegalese case, the results show that the behavior of the Dagum distribution differs from other distribution (continuous and empiric). The impact of the first simulation has an inverse effect on poverty with the Dagum distribution since the poverty indices decrease (-0.72% for incidence, -5.77 for depth and -7.75 for severity) whereas the poverty increases with the other distributions, except for severity with DAD and displaced lognormal distributions. The best fitting distribution, in this case the Singh-Maddala as well as the DAD approach produce an increase in poverty but the variation with the benchmark decreases when the poverty aversion rate increases.

Finally, another interesting result of the simulation 1 is that the impacts on poverty indices by groups are not all in the same direction. This result is crucial in the context of policy analysis and targeting. If the objective is to target the rural group where more than 80% of the population is poor, then the first policy (simulation 1) will have the expected effects. For all the distributions, the poverty indices decrease for the poorest group (rural non-educated). In the case of the Rural Educated, considering the best fitting distribution (Dagum), we have the headcount that marginally increases (0,03%) and decreases for the depth and severity[8]. But for the other groups, (less poor) we observe an increase of the FGT poverty indices. The global effect on Senegalese population will be an increase of poverty except for the Dagum distribution choice as we mentioned before.

	RNE		**RE**		**DKRNE**		**DKRE**		**AUNE**		**AUE**		**SENEGAL**	
	Sim1													
	Champ	**DAD**	***DAG***	**DAD**	***SM / DAG***	**DAD**	**SM / DAG**	**DAD**	**Champ**	**DAD**	**SM/ DAG**	**DAD**	**SM**	**DAD**
Mean Variation	*Decrease*		*Decrease or increase*		*Increase*		*Increase*		*Increase*		*Increase*		*Increase*	
FGT0	-0,41	-0,37	0,03	0,00	3,75 / 3,35	3,79	3,49 / 4,64	9,09	0,64	2,33	2,23 / 2,48	-0,94	0,55	0,95
FGT1	-1,15	-1,20	-0,24	-0,39	5,54 / 5,96	5,37	3,28 / 4,94	5,44	1,73	2,55	2,98 / 3,09	5,04	0,34	0,26
FGT2	-1,60	-1,64	-0,47	-0,37	6,97 / 7,94	7,49	2,22 / 4,09	5,23	2,53	4,07	3,48 / 3,44	7,05	0,00	-0,09
	Sim2													
	Champ	**DAD**	***SM / Logn3 / DAG***	**DAD**	***SM***	**DAD**	**SM / DAG**	**DAD**	**Champ**	**DAD**	**SM / DAG**	**DAD**	**SM**	**DAD**
Mean Variation	*Decrease*		*Decrease*		*Decrease*		*Decrease*		*Decrease*		*Decrease*		*Decrease*	
FGT0	-0,27	-0,18	-0,62 / -0,24 / -0,21	0,00	-1,02	0,00	-1,34 / -1,35	-0,51	-0,59	-0,33	-0,71 / -0,86	-2,81	0,05	-0,41
FGT1	-0,67	-0,30	-0,65 / -0,37 / -0,39	-0,39	-1,68	-1,79	-1,31 / -1,72	-0,41	-0,91	-1,09	-1,08 / -1,16	-1,68	-0,28	-0,57
FGT2	-0,94	-0,38	-0,68 / -0,42 / -0,42	-0,47	-2,09	-2,40	-1,67 / -2,34	-0,41	-1,07	-1,36	-1,28 / -1,37	-2,17	-0,53	-0,61

Table 5-5: *FGT's variations for DAD and" best" continuous distribution (%)*

For the second simulation, a decrease of 70% of the import tariff in the "other food" sector will reduce the poverty in Senegal and for each sub-group. The groups benefiting to most of this policy are the educated and urban households (Dakar and other urban areas) considering the poverty incidence index whereas the depth and severity decrease more for the other sub-groups. This result is interesting since this policy seems to be less beneficial to the poorer groups of the Senegalese population in light of the

[8] The FGT0 variation is positive but close to zero. This ambiguity could be explained by the small sample (138 observations).

FGT_0. However, when we look at the depth and severity, the groups that benefit the most are the poorest groups.

When there is an ambiguity on the choice of the continuous distribution, the poverty measure variations are generally close. This result runs for both simulations.

*** faire un commentaire sur le lissage

6. Conclusion

In this paper we attempted to illustrate what can be the implication of using a single distribution function to estimate poverty indices and income distribution in a CGE. We used seven different continuous distribution functions and one non-parametric method to estimate poverty indices from the household income generated by policy simulation/external shocks of an integrated multi-household CGE model of the Senegal economy. This model allows us to generate household specific income changes for 3278 households. Thus, we have in inter-group and intra-group changes in income distribution endogenously determined by the CGE model. We note that the changes in shapes generated by simulations can be significant enough that the fitting properties change before and after simulations. We also see that we can have a relatively big variation in poverty indices measurement depending on the choice of the functional form used. We used three statistical and one graphical methods to compare the fitting properties of the distributions. Consequently, we found that there is no single "best fitting" functional form for all groups, but the most flexible ones seems to be more efficient most often. We also show that results obtained from the non-parametric DAD approach are often at the extremes of what is obtained from smooth parametric forms. Moreover, when samples are relatively small, the non-parametric approach is not as sensitive as with functional forms, which contributes to smooth the distribution. This is especially true for the headcount ratio index. We think there is value in testing appropriateness of fit when analyst use these types of methods (CGE modelling) or other analytical methods that will change the nature of the distribution of income in an ex-ante situation. For instance, we can imagine that fixed income distribution between policy simulations will certainly lead to misleading conclusion and that the use of inappropriate functional forms can also bias the results. The richer the modelling approach is in providing insight on income distribution dispersion following a policy simulation or external shock, the stronger should be the concern over choosing an appropriate method to approximate the "true" income distribution of household to perform poverty analysis.

Performing rigorous work could involve using more than one functional form in an integrated multi-household CGE modelling exercise to analyse the impact of policy/external shocks on poverty and income distribution. Two characteristics of CGE modelling for poverty analysis might underlie these conclusions, the fact that when modellers undertake this work, they want to compare impact of policy on different household groups having different characteristics and therefore their distribution of income might well exhibit different properties. Moreover, policy simulation in CGE modelling can have small effects on income; therefore, functional forms that are more flexible and have more parameters might be the most appropriate choices. According to Metcalf (1972), three and four parameter functions might be better suited to capture economic fluctuations or policy simulations.

6. Bibliography

Abdelkhalek T. and A. Chaoubi (2000), « *Croissance économique et pauvreté au Maroc : une analyse des distributions des dépenses de consommation* », Cahier de recherche MIMAP/INSEA-Maroc, Janvier 2000, Etude # C2.

Adelman, I. and S. Robinson (1979), *Income Distribution Policy: A Computable General Equilibrium Model of South Korea*, in Adelman, I, The selected essays of Irma Adelman. Volume 1. Dynamics and income distribution. Economists of the Twentieth Century Series. Aldershot, U.K., pp. 256-89.

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

Aitchison, J. and J. A. C. Brown (1957), « *The lognormal distribution* », Cambridge University Press, Cambridge.

Ammon, (1895), « *Die Gesellschaftsordnung und ihre Natürlichen Grundlagen* », Jena

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

Bandourian, R., J. B. McDonald and R. S. Turley (2002), « *A comparison of Parametric Models of Income Distribution Across Countries and Over Time*», MIMEO Department of Economics, Brigham Young University.

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

Bordley, R. F., J. B. McDonald and A. Mantrala (1996), « *Something New, Something Old: Parametric Models for the Size Distribution of Income* », Journal of Income Distribution, 6(1), pp. 97-102.

Bourguignon, F., J. de Melo, and A. Suwa, (1991), « *Modelling the effects of adjustment programs on income distribution* », World Development, Vol. 19. no 11. pp. 1527-1544.

Champernowne, D.G. (1953), « *A model for income distribution* », Economic Journal, no. 53, pp. 318-351.

Champernowne, D.G. (1953), « *The graduation of income distributions* », Econometrica, vol. 20, no. 4, pp 591-615.

Chia, N. -C., S. Wahba and J. Whalley (1994), « *Poverty-Reduction Targeting Programs: a General Equilibrium Approach* », Journal of African Economies, 3(2), pp. 309-338.

Cockburn, J. (2001), « *Trade liberalization and Poverty in Nepal: A Computable General Equilibrium Micro-simulation Analysis* », Working paper 01-18. CREFA, Université Laval.

Cogneau D. and A.-S Robillard. (2000), « *Income Distribution, Poverty and Growth in Madagascar: Micro simulations in a General Equilibrium Framework* », IFPRI TMD Discussion Paper No.61.

Cohen, A. C. Jr. (1951), « *Estimating parameters of logarithmic-normal distribution by maximum likelihood* » Journal of the American statistical association, 46, pp. 206-212.

Dagum, C. (1977), « *A New Model of Personal Income Distribution: Specification and Estimation* », Économie Appliquée, Paris, Tome XXX, No. 3, pp. 413-437.

D'Agostino R. B. and M. A. Stephens, (1986), « *Goodness-of-fit techniques* », Marcel Dekker, INC, New-York and Basel.

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

Decaluwé, B., J.C. Dumont and L. Savard (1999), « *How to Measure Poverty and Inequality in General Equilibrium Framework* », Laval University, CREFA Working Paper #9920.

Decaluwé, B., A. Martens and L. Savard (2001), *La politique Économiques du Développement*, Université Francophone-Presse de l'Université de Montréal, Montréal. pp. 1-509.

de Janvry, A., E. Sadoulet and A. Fargeix (1991), « *Adjustment and Equity in Ecuador* » OECD Development Center, Paris.

de Janvry, A. and E. Sadoulet (1995), « *Quantitative Development Policy Analysis* », The Johns Hopkins University Press, Baltimore and London.

Dervis, K, J. de Melo and S. Robinson (1982), « *General Equilibrium Models for Development Policy* », Cambridge University Press, London, pp. 1-526.

Duclos J.Y., A. Arrar and C. Fortin (1999), « *DAD 4.01: Distributional analysis/Analyse distributive* », MIMAP Project, International Development Research centre, Government of Canada.

Fisk, P.R. (1961), « The graduation of income distributions », Econometrica, vol. 29, no. 2 pp. 171-185.

Foster, J., J. Greer and E. Thorbecke (1984), « *A Class of Decomposable Poverty Measures* », Econometrica, 52(3), pp. 761-766.

Gertel, H. R, R. Giuliodori, P. F. Auerbach and A. F. Rodriguez (2001), « *Unemployment and Income Distribution Analysis: New Evidence using a Dagum Parametric Income Distribution Model* », Facultad de Ciencias Económicas Universidad Nacional de Córdoba,

Gibrat (1931), « *Les Inégalités Economiques* », Paris, Sirey.

Gordy, M.B. (1998), « *A Generalization of the Generalized Beta Distribution* », Board of Governors of the Federal Reserve System, Washington.

Gunning, W.J. (1983), « *Income Distribution and Growth: A Simulation Model for Kenya* », in D.G. Greene (principal author), Kenya: Growth and Structural Change, 2 vols., Washington, DC: World Bank, pp. 487-621.

Harter, H. L. and Moore, A. H. (1966), « *Local Maximum likelihood estimation of the parameters of the three parameter lognormal distribution* », Journal of the American statistical association, 61, pp. 842-851.

Johnsons, N. L. and S. Kotz (1970), « *Continuous Univariate Distribution* », Vol. II, Wiley, New York.

Livny, M., R. Ramakrishnan and T. Zhang (1996), « *Fast Density and Probability Estimation Using CF-Kernel Method for Very Large Databases* » Computer Sciences Department, University of Wisconsin-Madison.

Mcdonald, J.B., and A. Mantrala (1995), « *The Distribution of Personnel lncome: Revisited* », Journal of applied Econometrics, vol. 10, pp. 201-204.

Mcdonald, J.B., and M.R. Ransom (1979), « *Functional formes, estimation techniques and the distribution of income* », Econometrica, vol. 47(6), pp. 1513-1525.

Mcdonald, J.B., and Y.J. Xu (1995), « *A Generalization of the Beta Distribution with applications* », Journal of Econometrics, vol. 66, no. 1-2. pp. 133-152-204.

Mcdonald, J.B. (1984), « *Some generalized functions for the size distribution of income* », Econometrica, vol 52, pp. 647-663.

Metcalf, E.C., (1972), *An Econometric Model, of the Income Distribution,* Markham Publishing Co. Chicago.

Salem, A.B.Z. and T. D. Mount (1974), « *A convenient descriptive model of income distribution* », Econometrica, vol. 42, pp. 1115-1127.

Singh, S.K. and G. S. Maddala (1976), « *A function for the size distribution of incomes* », Econometrica, vol. 44, pp. 963-973.

Majumber, A. and S. B. Chakravarty (1990), « *Distribution of personal income: development of a new model and its application to US income data* », Journal of applied econometrics. Vol 5. pp 189-196.

Morrisson, C. (1991), « *Adjustment incomes and Poverty in Morocco* ». World Development, Vol. 19. no 11. pp.

Ravallion, M. (1994), *Poverty Comparisons* , Harwood Academic Publisher.

Rosenblatt (1956), « *Remarks on some non-parametric estimates of a density function* », Annals of mathematical statistics, no. 27 pp. 832-835.

Rutherford, R.S.G. (1955), « *Income distribution a new model* », Econometrica, vol 23, pp 277-294.

Sahota, G. S. (1978), « *Theories of personal income distribution: A survey* », Journal of Economic Literature, vol 16, pp 1-55.

Spanos, A. (1999), *Probability theory and statistical inference: Econometric modelling with observational data*, Cambridge University Press, Cambridge.

Thorbecke, E. (1991), « *Adjustment growth and income distribution in Indonesia. and Equity in Indonesia* », World Development, Vol. 19. no 11. pp. 1595-1614.

7. Annex

This annex reports some tables discussed in the text and completes the illustration of the results.

Groups	Abbreviation	Frequencies	% of sample	Min income (Fcfa)	Max income (Fcfa)	Mean income (Fcfa)
Senegal	-	3278	100	15 670,52	7 524 305,42	237 903,40
Educated rural	RE	1265	38,60	27 350,18	2 237 762,38	194 629,22
Non-educated rural	RNE	138	4,20	15 670,52	3 073 462,67	116 810,14
Educated other urban	AUE	278	8,50	29 568,87	5 699 360,50	300 637,98
Non-educated other urban	AUNE	499	15,20	22 852,08	1 462 550,28	171 222,19
Educated Dakar	DKRE	533	16,30	44 912,58	7 524 305,42	528 869,00
Non-educated Dakar	DKRNE	565	17,20	54 685,59	2 364 345,00	255 297,18

Table 7-1: *Descriptive statistics for Senegal and 6 groups*

	BETA**	LOGN**	GAMMA**	CHAMP**	LOGN3*	SM*	DAG*	χ^2 Critical values
				SENEGAL				
SSE	0,01	*0,00*	0,01	*0,00*	*0,00*	*0,00*	*0,00*	* 30,14
SAE	0,45	0,20	0,42	0,14	0,17	*0,05*	0,06	** 31,41
Chi2	861,45	185,66	774,94	117,45	121,15	17,48	*16,45*	
LOGL	43 934,62	43 134,75	43 773,17	43 059,30	43 080,57	*46 592,13*	42 983,78	
				RNE				
SSE	*0,00*	*0,00*	0,01	*0,00*	*0,00*	*0,00*	0,01	* 22,36
SAE	0,16	0,14	0,22	*0,11*	0,14	0,13	0,29	** 23,68
Chi2	51,67	44,86	102,80	26,51	44,06	41,80	132,26	
LOGL	15 653,86	15 578,66	15 646,28	15 565,64	15 578,11	15 575,73	*15 669,60*	
				RE				
SSE	0,02	0,01	0,01	0,01	0,01	0,01	*0,00*	* 31,41
SAE	0,49	0,31	0,45	0,30	0,29	0,32	*0,27*	** 32,67
Chi2	57,50	*49,55*	99,61	44,02	33,90	33,01	*27,41*	
LOGL	1 774,71	1 789,60	1 819,19	1 786,20	1 780,69	*1 916,89*	1 778,42	
				DKRE				
SSE	0,02	*0,00*	0,01	*0,00*	*0,00*	*0,00*	*0,00*	* 27,59
SAE	0,22	0,24	0,45	0,20	0,20	0,12	*0,11*	** 28,87
Chi2	*28,36*	51,64	158,75	40,36	30,19	*24,12*	*10,27*	
LOGL	7 606,92	7 548,83	7 646,22	7 541,57	7 532,84	7 772,48	7 520,77	
				DKRNE				
SSE	0,01	*0,00*	0,01	*0,00*	*0,00*	*0,00*	*0,00*	* 30,14
SAE	0,32	0,25	0,39	0,21	0,17	0,16	*0,15*	** 31,41
Chi2	69,19	54,70	118,08	44,83	31,28	32,29	31,61	
LOGL	7 579,64	7 560,65	7 632,95	7 550,41	7 537,14	*8 289,19*	7 534,26	

	BETA**	LOGN**	GAMMA**	CHAMP**	LOGN3*	SM*	DAG*	χ^2 Critical values
				AUE				
SSE	0,01	0,00	0,02	0,00	0,00	0,00	0,00	* 3,84
SAE	0,33	0,08	0,25	0,04	0,06	0,05	0,02	**5,99
Chi2	45,95	2,28	22,00	0,97	1,59	0,89	0,20	
LOGL	4 369,49	4 385,30	4 434,24	4 380,41	4 382,74	4 478,64	4 376,07	
				AUNE				
SSE	0,01	0,00	0,01	0,00	0,00	0,00	0,01	* 5,99
SAE	0,16	0,12	0,19	0,04	0,12	0,10	0,25	** 7,81
Chi2	24,89	10,98	29,28	1,12	10,99	9,87	54,95	
LOGL	8 110,75	8 104,22	8 143,64	8 094,66	8 104,22	8 124,18	8 153,34	

Table 7-2: *Goodness-of-fit test, Simulation 1 case*

	BETA**	LOGN**	GAMMA**	CHAMP**	LOGN3*	SM*	DAG*	χ^2Critical values
				SENEGAL				
SSE	0,01	0,00	0,01	0,00	0,00	0,00	0,00	* 30,14
SAE	0,45	0,20	0,42	0,13	0,16	0,06	0,14	** 31,41
Chi2	866,43	174,57	756,30	108,43	111,79	15,02	91,35	
LOGL	43 888,99	43 116,79	43 757,59	43 044,38	43 061,16	46 553,72	43 048,22	
				RNE				
SSE	0,00	0,00	0,01	0,00	0,00	0,00	0,01	* 22,36
SAE	0,16	0,14	0,22	0,10	0,14	0,11	0,27	** 23,68
Chi2	51,67	37,69	95,89	20,46	36,65	33,57	121,85	
LOGL	15 612,68	15 541,38	15 611,15	15 528,89	15 540,52	15 540,31	15 627,67	
				RE				
SSE	0,02	0,01	0,01	0,01	0,01	0,01	0,01	* 31,41
SAE	0,47	0,35	0,46	0,35	0,34	0,38	0,34	** 32,67
Chi2	53,33	73,36	134,44	65,58	56,16	51,63	47,90	
LOGL	1 772,84	1 787,26	1 816,79	1 783,96	1 778,53	1 931,49	1 776,19	
				DKRE				
SSE	0,03	0,00	0,01	0,00	0,00	0,00	0,00	* 27,59
SAE	0,23	0,25	0,45	0,21	0,21	0,13	0,12	** 28,87
Chi2	32,09	57,34	170,08	45,16	36,50	27,13	13,70	
LOGL	7 617,87	7 551,04	7 648,57	7 543,60	7 536,57	7 733,88	7 523,46	
				DKRNE				
SSE	0,01	0,00	0,01	0,00	0,00	0,00	0,00	* 30,14
SAE	0,31	0,23	0,37	0,18	0,17	0,15	0,17	** 31,41
Chi2	68,37	48,46	111,45	38,55	24,35	26,02	24,85	
LOGL	7 576,36	7 557,48	7 630,26	7 547,26	7 532,92	8 323,21	7 529,71	
				AUE				
SSE	0,01	0,00	0,03	0,00	0,00	0,00	0,00	* 3,84
SAE	0,34	0,11	0,27	0,04	0,11	0,03	0,02	**5,99
Chi2	53,89	3,99	26,70	1,11	4,00	0,53	0,32	
LOGL	4 369,04	4 384,73	4 433,54	4 379,73	4 384,73	4 476,02	4 375,53	
				AUNE				
SSE	0,01	0,00	0,01	0,00	0,00	0,00	0,02	* 5,99
SAE	0,17	0,13	0,19	0,03	0,13	0,09	0,26	** 7,81
Chi2	25,95	11,22	30,06	0,80	11,22	8,75	57,16	
LOGL	8 096,48	8 091,42	8 130,62	8 080,22	8 091,42	8 121,97	8 143,60	

Table 7-3: *Goodness-of-fit test, Simulation 2 case*

		DAD	BETA	LOGN	GAMMA	CHAMP	LOGN3	SM	DAG
		RNE							
SIM 1	ΔFGT0	-0,37	-0,67	-0,49	-0,49	-0,41	-0,48	-0,50	-0,56
	ΔFGT1	-1,20	-1,28	-1,12	-1,03	-1,15	-1,19	-1,31	-1,03
	ΔFGT2	-1,64	-1,58	-1,50	-1,36	-1,60	-1,59	-1,87	-1,36
Sim2	ΔFGT0	-0,18	-0,35	-0,30	-0,37	-0,27	-0,30	-0,38	-0,24
	ΔFGT1	-0,30	-0,69	-0,68	-0,68	-0,67	-0,65	-0,47	-0,56
	ΔFGT2	-0,38	-0,86	-0,89	-0,88	-0,94	-0,89	-0,54	-0,79
		RE							
Sim1	ΔFGT0	*0,00*	0,16	-0,15	-0,24	-0,11	0,10	4,04	0,03
	ΔFGT1	-0,39	0,14	-0,57	-0,76	-0,54	-0,09	2,48	-0,24
	ΔFGT2	-0,37	-0,03	-0,94	-1,31	-0,91	-0,26	1,58	-0,47
Sim2	ΔFGT0	*0,00*	-0,33	-0,28	-0,34	-0,27	-0,24	-0,62	-0,21
	ΔFGT1	-0,39	-0,37	-0,39	-0,38	-0,40	-0,37	-0,65	-0,39
	ΔFGT2	-0,47	-0,35	-0,44	-0,41	-0,42	-0,42	-0,68	-0,42
		DKRE							
Sim1	ΔFGT0	9,09	1,93	2,43	0,18	2,58	3,99	3,49	4,64
	ΔFGT1	5,44	0,03	1,70	-1,93	1,43	3,83	3,28	4,94
	ΔFGT2	5,23	-1,92	0,39	-3,65	0,00	3,54	2,22	4,09
Sim2	ΔFGT0	-2,24	-0,51	-0,95	-0,64	-1,04	-1,08	-1,34	-1,35
	ΔFGT1	-1,36	-0,41	-1,19	-0,80	-1,07	-1,39	-1,31	-1,72
	ΔFGT2	-1,74	-0,41	-1,57	-0,66	-1,22	-1,33	-1,67	-2,34
		DKRNE							
Sim1	ΔFGT0	3,79	2,67	3,43	3,28	3,81	3,10	3,75	3,35
	ΔFGT1	5,37	4,16	4,98	4,59	5,27	5,32	5,54	5,96
	ΔFGT2	7,49	5,44	6,10	5,97	6,46	7,42	6,97	7,94
Sim2	ΔFGT0	*0,00*	-0,85	-1,08	-0,97	-1,18	-1,01	-1,02	-1,08
	ΔFGT1	-1,79	-1,23	-1,56	-1,31	-1,51	-1,67	-1,68	-1,79
	ΔFGT2	-2,40	-1,52	-1,86	-1,56	-1,97	-1,92	-2,09	-2,35
		AUE							
Sim1	ΔFGT0	-0,94	2,20	1,97	1,58	2,26	2,14	2,23	2,48
	ΔFGT1	5,04	2,64	2,38	1,51	2,64	2,65	2,98	3,09
	ΔFGT2	7,05	2,94	2,60	1,33	2,77	3,01	3,48	3,44
Sim2	ΔFGT0	-2,81	-0,70	-0,75	-0,63	-0,85	-4,92	-0,71	-0,86
	ΔFGT1	-1,68	-0,85	-0,97	-0,79	-1,01	-4,44	-1,08	-1,16
	ΔFGT2	-2,17	-0,92	-1,07	-0,89	-1,14	-2,86	-1,28	-1,37
		AUNE							
Sim1	ΔFGT0	2,33	0,36	0,64	0,61	0,64	0,51	1,44	0,70
	ΔFGT1	2,55	1,14	1,42	1,27	1,73	1,24	2,60	1,24
	ΔFGT2	4,07	1,79	2,02	1,75	2,53	1,93	3,62	1,62
Sim2	ΔFGT0	-0,33	-0,57	-0,54	-0,61	-0,59	-0,67	-0,56	-0,41
	ΔFGT1	-1,09	-0,75	-0,80	-0,80	-0,91	-1,02	-1,03	-0,76
	ΔFGT2	-1,36	-0,82	-1,01	-1,01	-1,07	-1,10	-1,31	-0,97
		SENEGAL							
Sim1	ΔFGT0	0,95	0,89	0,65	1,08	0,79	0,50	0,55	-0,72
	ΔFGT1	0,26	0,54	0,36	0,90	0,54	0,18	0,34	-5,77
	ΔFGT2	-0,09	0,28	0,08	0,83	0,25	-0,08	0,00	-7,65
Sim2	ΔFGT0	-0,41	-0,48	-0,44	-0,46	-0,49	-0,39	0,05	-0,34
	ΔFGT1	-0,57	-0,51	-0,54	-0,55	-0,58	-0,53	-0,28	-0,54
	ΔFGT2	-0,61	-0,49	-0,67	-0,58	-0,68	-0,58	-0,53	-0,64

Table 7-4: *FGT variations for Senegal and the six sub-groups (%)*